Developing a psychological model of users' experience of public-service Web sites

Esen Moğulkoç

Master of Science Thesis

August 26th, 2008

Communication Studies

Faculty of Behavioural Science

University of Twente, Enschede, the Netherlands

Supervisors:

Dr. S. A. de Vries

Dr. Ir. P. W. de Vries

Abstract

This research aims to study the influence of individual-difference variables (need for cognition as a trait, self-efficacy as a trait, self-efficacy as a state and intrinsic motivation) on user experience and technology acceptance, moderated by artefact and task characteristics. User experience outcomes include mental effort, perceptions of aesthetics (classic aesthetics and expressive aesthetics), enjoyment and disorientation. In addition, technology acceptance variables such as perceived usefulness, perceived ease of use and intention to use are other important outcome measures in this study. The current research addresses the following research question:

What is the effect of individual-difference variables on user experience outcomes of the use of public-sector Web sites and how is the effect moderated by artefact characteristics and task characteristics?

This question is answered by developing a psychological model of users' experience of public-service Web sites, which provides an insight for the effect of individual-difference variables moderated by artefact characteristics and task characteristics. This model depends on The Person-Artefact-Task (PAT) model as the frame of the current study. The PAT model emphasises three interacting components contributing to the flow experience within an Internet-based context: person, artefact and task. Individual-difference variables used in the current study are related to the first component of this model. Artefact characteristics and task characteristics, which are characterised by their level of complexity, are the other components of the model.

For testing psychological model of users' experience of public-service Web sites, both a larger-scale quantitative experiment and a smaller-scale qualitative procedure have been employed to answer the research question. For quantitative experiment, three experimental conditions were generated by a combination of (public-service) Web site and task complexity. The first condition represented a high artefact complexity with high task complexity. The second condition represented a high artefact complexity with low task complexity. The third condition represented a low artefact complexity with low task complexity. Artefact complexity was formed by using two different UK council Web sites, while task complexity was varied within the same (complex) Web site. Tasks required either one or two steps (low task complexity) or four or five steps (high task complexity), respectively, to complete tasks. Think-aloud method was employed as a qualitative method, which requires participants to verbalise their thoughts while performing tasks.

The findings indicated that moderation of individual-difference variables by high artefact complexity and low task complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was

bigger compared to those with lower individual-difference variables. While people with a high need for cognition experienced higher perceived enjoyment and lower perceptions of expressive aesthetics than those with low need for cognition, the relation between those with high need for cognition and other outcome measures could not be confirmed. Besides, people with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state. However, the relation between those with high self-efficacy (trait and state) and other outcome measures also could not be proven. No association of people with a high intrinsic motivation with outcome measures was found. On the other hand, while test subjects performing in complex Web site experienced lower mental effort, test subjects performing complex tasks experienced higher mental effort. Regarding these quantitative findings, the final psychological model of users' experience of public-service Web sites (PAT-UX) has been developed (Figure I). This model enhances the understanding of user experience and moderated effect of user characteristics by Web site characteristics and task characteristics.

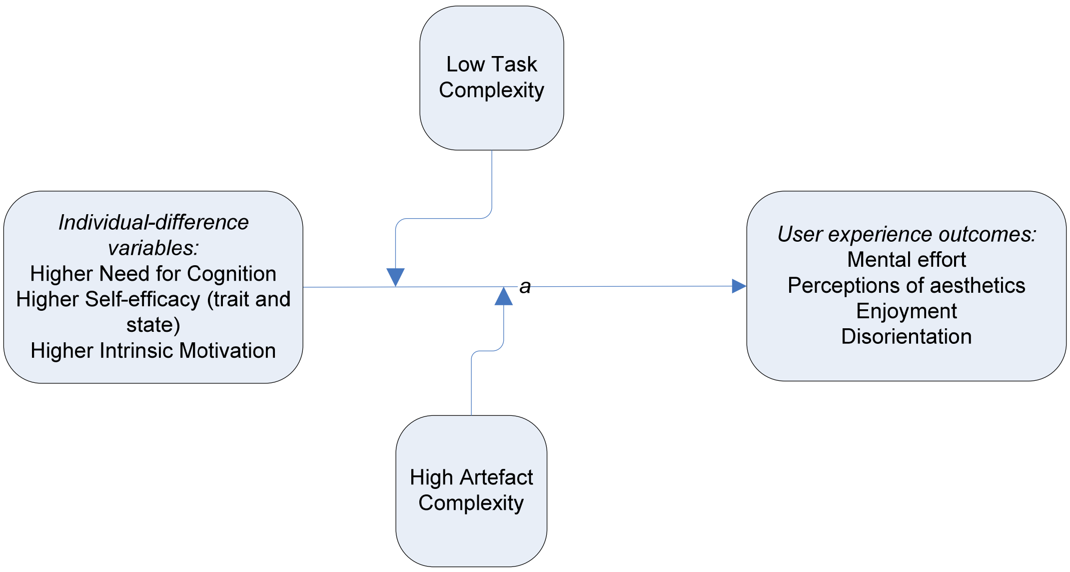

Figure I The final model of psychological model of users' experience of public-service Web sites (PAT-UX model)
[a] Higher effect of moderators

In addition, qualitative findings, which are used to give some design recommendations to designers and future studies, are highly contributive in comprehending users' experience and useful design in the context of e-government. Even though, this study has some limitations, it can guide both designers and e-government practitioners.

Acknowledgements

I would like to start by expressing my special gratitude directly to Paul van Schaik, for his guidance through my work and help in each and every step of it with great patience. Without his advices, help and support, this thesis would not have been realised in this version.

I wish to continue by sending my appreciations to Sjoerd de Vries and Peter de Vries, my thesis supervisors, for the advice and support in the development of this research. I extend my gratitude to Joyce Karreman, for being my mentor and every professor of the New Media Research and Design master programme, who provided me guidance in previous stages of the development of this thesis.

I would also like to thank Egbert Woudstra, Susanna Hartzheim, Dionysia Loman and other members of BOZ for their help and advice through admission to master programme, as well as progress and completion of it.

Other than those whose names are presented above, I would like to thank to the students who participated in my experiment. Last but not the least, I will conclude by sending my thanks to my husband and to our families for their continuous support.

Table of Contents

1. Introduction

Nowadays, the importance of user-centred relationship, which includes approaching to target audiences on a personal level, is being acknowledged by many governments all over the world in order to segment their audiences and to provide them a more efficient communication, service delivery or transactions. The overwhelming growth of ICTs is certainly the main reason of this emphasis on a client/citizen/user-centric relationship. Furthermore, even though recent research investigated the acceptance and use of e-government services by Dutch citizens (Van Dijk et al, 2007), the impact of users' characteristics in the use of technology within an e-government context still deserves a further examination. There is an indisputable influence of personal needs of users on their perceptions and intentions for the purpose of establishing an effective relationship between individuals and services in a computer-based environment, there appears to be a lack of research on these characteristics distinguishing individuals from each other and the external effects such as characteristics of the Web sites moderating this Human-Computer Interaction in more specifically public-service Web sites. It is important to address this important issue in order to comprehend whether individual differences affect user experience in e-government Web sites and what the influence of both site and task characteristics is. Therefore, the aim of this research is to study the influence of individual-difference variables (need for cognition as a trait, self-efficacy as a trait, self-efficacy as a state and intrinsic motivation) on user experience and technology acceptance, moderated by artefact and task characteristics. The research addresses the following question:

What is the effect of individual-difference variables on user experience outcomes of the use of public-sector Web sites and how is the effect moderated by artefact characteristics and task characteristics?

The current study investigates the research question from The Person-Artefact-Task (PAT) model's perspective. The PAT model developed by Finneran and Zhang (2003) is, therefore, the frame of this research in response to the need for illuminating the main issues of the study. In a Web-based environment, the PAT model emphasises three interacting components contributing to the flow experience such as person, artefact and task. Besides PAT model, some other studies such as paradox of active user, technology acceptance, design principles, information scent and perceptions of aesthetics are considered because research on user experience has a broad range beginning from the study of usability and aesthetics to technology use (Hassenzahl & Tractinsky, 2006).

The concept of segmentation is associated with the first component of the PAT model, namely person. In the marketing literature there are many definitions for the concept of segmentation. Rao and Steckel (1998) defined segmentation as ‘identifying groups of consumers who behave differently in response to a given marketing strategy’ (p. 23). In this definition they emphasised not only the homogeneity within groups, but also dissimilarities between groups. According to Weinstein (2004), segmentation is ‘the process of partitioning markets into groups of potential customers with similar needs and/or characteristics who are likely to exhibit similar purchase behaviour’ (p. 4). Van Dijk et al. (2005) defined segmentation as ‘the process of dividing a population into groups (segments) on the basis of similarities in user-related information of individuals’ (p. 14). While Weinstein (2004) was stressing homogeneity within groups in terms of their purchase behaviour, the definition of Van Dijk et al. (2005) comprised all user-related information. The purpose of segmentation is to gain an understanding and knowledge of customers, in order to give them what they want, build relationships with them and communicate with them via targeted channels (Weinstein, 2004). Rao and Steckel (1998) alleged that lifestyles of consumers have become much more complex because of expanded disposable incomes and higher educational levels. As a result, the changing needs of individuals as clients, customers, or citizens do not only influence segmentation strategies in the marketing domain, but also those in e-government activities. Therefore, selecting the right strategy which embraces varying needs of audiences is vital for the services of electronic government.

Recent research (Doornbos, 2004) yielded three main segmentation strategies for marketing, which can be applied to e-government: undifferentiated, differentiated, and concentrated. In the undifferentiated segmentation strategy all segments are treated in the same way without concerning their differences. In the differentiated segmentation strategy separate segments are approached differently. The concentrated segmentation strategy focuses on one particular segment. Because of the nature of e-government which must serve all citizens, the concentrated segmentation strategy is highly inappropriate. As a successful relationship between government and citizens depends heavily on focusing unique needs of different segments, the most appropriate segmentation strategy seems the differentiated segmentation.

As mentioned previously, users’ characteristics in the context of e-government are important for user experience of public-service Web sites. Each individual has a unique pattern of mental abilities and the variance in characteristics of individuals describes how individuals interact with their environment (Jonassen & Grabowski, 1993). Research on individual differences is often considered by differential psychology. Differential psychology explores the impact of individual differences, which are firmly related to intelligence, cognitive styles and personality, on human behaviour. This literature review indicated three frequently elaborated individual-difference variables: need for cognition, self-efficacy and intrinsic motivation. Need for cognition is useful for understanding individuals’ behaviour because it is closely related with general theoretical frameworks (Haugtvedt et al., 1992). Self-efficacy is an appropriate construct to use in order to investigate the research question of the current study because self-efficacy is a direct determinant of an individual's behaviour supported by the literature (Bandura, 1982). Intrinsic motivation is an important construct because it can encourage the exploration of users in Web-based environments (David et al., 2007). Although the constructs of need for cognition and intrinsic motivation overlap partly by their definitions, several authors note their unique significance. While need for cognition is people's tendency to engage in and enjoy effortful cognitive endeavours (Cacioppo & Petty, 1996), intrinsic motivation retains the positive potential of individuals through novelty and challenges in order to learn and explore (Ryan & Deci, 2000). On the other hand, self-efficacy is the key construct of social cognitive theory of Bandura (1977) and reflects the confidence of individuals in accomplishing a particular task.

Electronic government, or e-government, is related to the second component of the PAT model, namely artefact. This can be defined simply as 'the use of technology to enhance access and delivery of governmental services to benefit citizens, business partners and employees' (Silcock, 2001, p. 88). Each of these customer groups, or market segments, adopts and uses e-government services differently, and for different reasons. Schedler and Summermatter (2007) proposed three statements to comprehend e-government better: e-government uses information technology, e-government deals with organisational issues of public administration and e-government considers the interaction of public administration with its environment.

As mentioned previously, segmentation is an important tool to acquire an understanding and knowledge of customers for e-government as well as for e-commerce. The most prominent difference of e-government from e-commerce is the fact that e-government should provide access to the entire population, including people who might be hard to reach, such as people with lower educational levels and/or disabilities (Carter & Bélanger, 2005), whilst e-commerce can choose its customers and exclude some of them. Nevertheless, segmentation enables e-government to address different audiences differently and therefore can be more user-centred, more service-oriented, more effective and more cost-efficient. Accordingly, segmentation in the context of e-government effectively augments the communication of government with its citizens and dissemination of governmental services by means of applying the right strategy. In this context, right strategy requires acknowledging both similarities within groups and differences between groups and addressing these citizens in terms of their needs.

Artefact characteristics include the application of user-interface design principles such as the use of colour, font type, organisation of information on a Web page, characteristics of classic aesthetics design and expressive aesthetics design (Lavie & Tractinsky, 2004) and information scent as a match between task goals and (textual) information presented on a Web page (Blackmon et al., 2002). More specifically, Nadkarni and Gupta's (2007) and Ivory et al.'s (2001) objective-Web-site-complexity metrics can be used in order to define artefact characteristics.

Task is the third component of the PAT model. Here, in the current study task characteristics include aspects of task structure, such as length of task sequence and number of decision points. There is an undeniable association between individual differences and task and artefact characteristics. Attitudes of users towards a Web site are highly affected by their characteristics and information cues on the Web site influence a user's experience of this Web site (Nadkarni & Gupta, 2007). In addition, the PAT model examines this relationship extensively.

User experience outcomes include mental effort (Zijlstra, 1993), perceptions of aesthetics (Lavie & Tractinsky, 2004), enjoyment (Jackson & Marsh, 1996) and disorientation (Ahuja & Webster, 2001). In addition, technology acceptance variables such as perceived usefulness, perceived ease of use and intention to use are important outcome measures in this study as potential gains in user experience as well as efficiency and effectiveness may be lost if interactive systems meet with a lack of acceptance.

Depending on the main interest of the research, individual-difference variables are focal variables and artefact characteristics and task characteristics are moderators or the roles of these three are exchanged.

2. Theoretical background and model

The purpose of this chapter is to elaborate the Person-Artefact-Task (PAT) model with respect to several theories on user experience and technology acceptance. These theories are important because they give an explanation toward user experience and technology acceptance of public-service Web site users. The PAT model is used to build a relation between its three components and user experience for e-government Web sites. After explaining the framework of this study, the characteristics of person, artefact and task are revealed with a deep examination. At the end of this chapter, a psychological model of users' experience will be proposed with several hypotheses. This model is developed to test whether or not there is an effect of individual-difference variables moderated by artefact and task complexity on user experience outcomes. Within the frame of the current study, the PAT model is going to be used to model users' experience of public-service Web sites.

2.1 The Person-Artefact-Task (PAT) model in perspective

This section reviews the Person-Artefact-Task (PAT) model in order to elucidate user experience and technology acceptance. As mentioned above, the PAT model is the frame of the current study. According to this model, user experience is affected by the interaction of three dimensions: person, artefact and task. After explaining what user experience and its relation with the PAT model are, both Technology Acceptance Model (TAM) and Universal Theory of the Acceptance and Use of Technology (UTAUT) will be explained because these two are important to comprehend technology acceptance of Web users as a part of user experience. Moreover, paradox of the active user is also a relevant concept to user experience of public-service Web sites because of its focus toward learning new procedures (new experiences) in a computer-mediated environment.

2.1.1 User experience

User experience is a frequently-used term in the field of Human-Computer Interaction. This concept will be explained before presenting the Person-Artefact-Task (PAT) model in the next section in response to need for understanding what user experience is before understanding how the dimensions of the model affect users' experience.

There are many definitions of user experience in the literature, emphasising different aspects of the concept such as pragmatic and hedonic aspects of the concept. Among these definitions, Hassenzahl and Tractinsky's (2006) definition of user experience deserves attention. They studied user experience broadly and defined the concept of user experience as follows:

'A consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.)' (p. 95)

There are two reasons why their definition is an important attempt. On the one hand, their definition is pretty contributive because it includes two of the aspects of user experience addressed by the current study: person and artefact. (Indeed, it can be seen that Hassenzahl and Tractinsky's definition of the term of user experience fits the theoretical background of the current study with task as an aspect of the context of use). On the other hand, the research on the concept of user experience ranges from the study of usability and aesthetics to technology use. A problem so far has been that there is a lack empirical research on user experience. The current research addresses this important gap and contributes to the emerging knowledge base in this research field.

2.1.2 The Person-Artefact-Task (PAT) model

The Person-Artefact-Task (PAT) model is the frame of reference used in the current study (Finneran & Zhang, 2003). This research investigates the effect of individual-difference variables moderated by task and artefact characteristics on user experience from the PAT model's perspective. The PAT model was designed to systematically study the construct of flow, as in Csikszentmihalyi's (1990) Flow Theory, but the PAT model can be applied more widely to the study of user experience – as demonstrated by the current study. Flow is defined as 'holistic sensation that people feel when they act with total involvement' (Csikszentmihalyi, 1990, p. 477). Another definition to the concept of flow from literature is 'a psychological state in which the person feels simultaneously cognitively efficient, motivated, and happy' (Moneta & Csikszentmihalyi, 1996, p. 277).

Flow occurs when all levels of consciousness are consistent with each other (Csikszentmihalyi, 1988). Individual who are experiencing flow, or in other words achieving an optimal experience within an activity, are likely to have clear goals, exercise control, lose their self-consciousness and experience a distortion of time (Finneran & Zhang, 2003). This optimal experience occurs by means of peoples' perceptions of challenges and skills in given situations. According to Csikszentmihalyi's (1990) conceptualisation, flow experience yields positive effect, high levels of arousal and perceived freedom. In his view, flow, or optimal experience, is a kind of purpose of people's life. Flow state is substantially associated with how people become absorbed in their activities and irrelevant thoughts are screened out (Chen et al., 1999). Literature displayed that flow leads to a number of favourable experiences such as freedom from self-consciousness and enjoyment of the activity (Jackson & Marsh, 1996). Jackson and Marsh (1996) defined enjoyment as an autotelic experience, an intrinsically rewarding activity which is done for its own sake without any further benefit. Enjoyment is conceptualised by Csikszentmihalyi (1990) as the final result of being in flow.

Recent research (Chen et al., 1999) on flow experience in an interactive computer system frames flow experience via the following stages: flow antecedents, flow experience and flow consequences. Problems in the conceptualisation of flow experience are the 'inconsistent flow models, different uses of constructs and ambiguous operationalisations' in the literature (Finneran & Zhang, 2002; cited in Finneran & Zhang, 2003, p. 476). Finneran and Zhang (2003) aimed to address these problems through their PAT model, emphasising flow antecedents leads.

The PAT model emphasises three interacting components of an interactive computer system, contributing to the flow experience: person, artefact and task (Figure 1). The PAT model distinguishes these three concepts from each other in order to comprehend better their effects on flow experience. Thus, the current study focuses on the effect of users' (persons') characteristics on flow as moderated by task- and artefact characteristics.

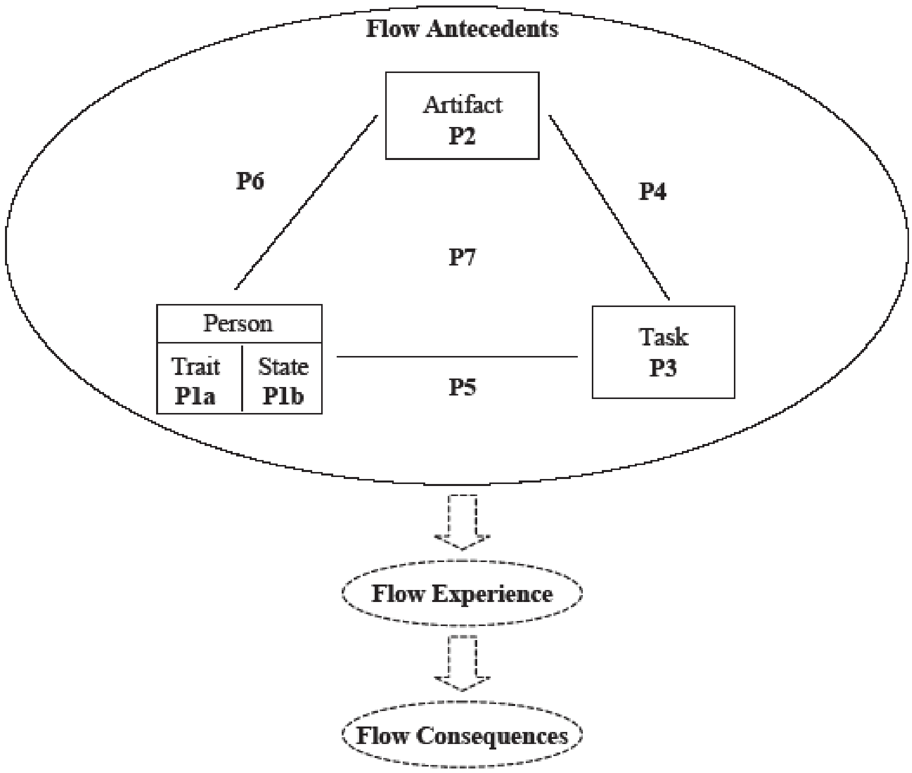

Figure 1 Stages of flow and the person-artefact-task model of flow antecedents (Finneran & Zhang, 2003).

A person (computer user) is the individual performing a task using an interactive computer system. The tendency of people to experience flow is highly associated with their personal characteristics. Some individuals, for instance, are more likely to experience flow than others (Finneran & Zhang, 2003). Therefore, flow experience, as a trait, is relatively stable across occasions (Hong & O'Neil, 2001). However, a person with high tendency of flow experience, for instance, in a Web-based environment may not experience flow while searching and seeking information within an uninteresting or frustrating Web site. An artefact is a tool, such as a search engine or a Web site, for accomplishing an activity. Artefacts as a tool in the context of Human-Computer Interaction are not completely controlled by users (Finneran & Zhang, 2003). Therefore, the characteristics of artefacts should be considered in a Web-based environment because they likely yield a marked influence on user experience. A task is the main goal of an activity such as using e-mail and voice-mail. Task is an important facet of flow experience because flow state of a person is driven by the task at hand (Finneran & Zhang, 2003).

There is a diverse interaction between the three components of the PAT model. The artefact-task consistency can be considered as the support of a specific task (structure) by technology. Artefact-person consistency can be seen as the artefact's clarity of comprehension and usability for the person using the artefact. Person-task consistency can be regarded as a task's challenge to the person and a person's skill (in terms of procedural and declarative knowledge required) at performing a task.

The PAT model contributes to the literature by means of re-conceptualizing flow for individuals experiencing an activity when using an interactive computer system. However, the model can be used to model user experience (of which flow is a component) more generally. Therefore, in the current study this model is called the Person-Artefact-Task User Experience (PAT-UX) model.

2.1.3 The Technology Acceptance Model (TAM)

Considering the PAT model as the frame of the study, the Technology Acceptance Model (TAM) proposed by Davis (1989) also can contribute to the current study because of its firm association with the acceptance of individuals through technology. TAM explains potential users' behavioural intentions to use a technological system (e.g., innovation). TAM was designed to predict technology acceptance and usage in a job-based environment (Venkatesh et al., 2003), but more recently other (hedonic) types of system have been studied (van der Heijden, 2004). According to van der Heijden (2004), hedonic versus utilitarian nature of the interactive computer systems are distinguished in the basis of the purpose of users. Whilst individuals utilise Web sites for playing computer games and messaging in order to seek their hedonic purposes, they use utilitarian Web sites for work arrengements to accomplish their job-related purposes. His study supports that both perceived ease of use and perceived enjoyment is tightly associated with behavioural intention to use of an hedonic system than perceived usefulness. Furthermore, the impact of the determinants of technology acceptance model changes as the goal is to provide a hedonic value rather than aiming to provide a utilitarian value.

TAM states that user adoption of a new technological system is determined by both their intention to use the system and their beliefs about the system. Two core constructs of TAM are perceived usefulness and perceived ease of use. Perceived usefulness is defined as 'the extent to which a person believes that using a particular system will enhance his or her job performance' (Davis, 1989, p.320). Perceived ease of use is defined as 'the extent to which a person believes that using a particular system will be free of effort' (Davis, 1989, p.320).

Regarding the validity of TAM, King and He (2006) conducted a meta-analysis depending on 88 TAM studies. Their research displayed that TAM is one of the most widely used models in information technologies.

2.1.4 Universal Theory of the Acceptance and Use of Technology (UTAUT)

Universal Theory of the Acceptance and Use of Technology (UTAUT) was developed by Venkatesh et al. (2003), in order to improve on the explanatory of its predecessor TAM in explaining users' acceptance of new technology. UTAUT is derived from eight existing theories of the determinants of intention and usage of information technology, and synthesises the main constructs of these models. Theory was formulated by four core determinants of intention to use of new technology such as performance expectancy, effort expectancy, social influence and facilitating conditions. These four play a significant role as direct determinants of user acceptance and usage behaviour and their effects are moderated by gender, age, experience and voluntariness.

Performance expectancy is defined by Venkatesh et al. (2003) as 'the degree to which an individual believes that using the system will help him or her to attain gains in job performance' (p. 447). Performance expectancy is formulated via five interacting constructs from the literature: perceived usefulness (from Technology Acceptance Model), extrinsic motivation (from Motivational Model), job-fit (from Model of PC Utilisation), relative advantage (from Innovation Diffusion Theory) and outcome expectations (from Social Cognitive Theory). Venkatesh et al.'s findings support performance expectancy as the most important determinant of intention among the four determinants.

Venkatesh et al. (2003) defined effort expectancy as 'the degree of ease associated with the use of the system' (p. 450). Three concepts from the existing theories compose effort expectancy: perceived ease of use (from Technology Acceptance Model), complexity (from Model of PC Utilisation) and ease of use (from Innovation Diffusion Theory).

The definition of social influence by Venkatesh et al. (2003) is 'as the degree to which an individual perceives that important others believe he or she should use the new system' (p. 451). Social influence construct of UTAUT comprises the concepts of subjective norm (from Technology Acceptance Model), social factors (from Model of PC Utilisation) and image (from Innovation Diffusion Theory).

Facilitating conditions are defined as 'the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system' (p. 453). Facilitating conditions refer to the constructs of perceived behavioural control (from Theory of Planned Behaviour), facilitating conditions (from Model of PC Utilisation) and compatibility (from Innovation Diffusion Theory).

2.1.5 Paradox of the active user

Paradox of the active user is a related concept because it focuses on users' activities (experiences) with a system. The active user as a term of computer-mediated systems embraces people who are engaging an unknown learning-to-use-a-computer situation (Carroll & Rosson, 1987). In Carroll and Rosson's view active user tend to use the functions they already know and are comfortable using in order to accomplish a new goal instead of exploring new methods - in other words 'striking out into the unknown' – for the same goal (p. 81). Here, this statement points out a paradox of the active user in a computer-mediated environment and is considered an issue for both researchers of the field and designers of the computer systems. The paradox of the active user is the persistence of both new and experienced users in using inefficient familiar procedures in spite of the existence of more efficient methods (Fu & Gray, 2004). There are two specific paradoxes: the production paradox and the assimilation paradox.

The production paradox captures not only the idea of 'one would only ever want to learn to use a new tool if one wanted first to get something done' (Carroll & Rosson, 1987, p. 82), but also the necessity of learning something for using which is sometimes a challenge for users. New users, for instance, may ignore to read the instruction or guide of the new tool since they focus on executing their real work. This brings about unwillingness and frustration in novice users in learning the new tool rather than performing a real task using the tool, regardless of how important and efficient to learn using the tool. On the other hand, experienced users, which are encountering in learning, face a production paradox as a balance between investment of time in learning new procedures and already-known procedures. While already-learned procedures in the system enable users to acquire their existing/regular outputs, new procedures in the system may lead users to gain better outputs in the future.

The assimilation paradox represents the idea that 'if we knew nothing at all, it is difficult to imagine how we could make any progress at learning anything at all' (Carroll & Rosson, 1987, p. 88). This paradox emphasises the association of new things and already-learned things. New users have an inclination of comprehending new systems by using their prior knowledge about the other (interactive computer and other) systems. Nevertheless, what novice users already know about another system may be disadvantageous by misleading them about how a new system works as novice users encounter a new situation. Experienced users may also face problems, such as negative transfer.

Carroll and Rosson (1987) recommended some strategies to overcome each of two paradoxes for both novice and experienced users, such as making the system easy to learn, making the system similar to something familiar, making the system intrinsically rewarding, and reducing apparent connections to prior knowledge.

2.2 The Person-Artefact-Task characteristics

This section aims to explain three interacting components of the Person-Artefact-Task model with regard to the constructs selected to define these dimensions. Need for cognition, self-efficacy (trait and state) and intrinsic motivation characterise the person whilst aesthetics, design principles, information scent and Web-site complexity defines artefact. Task complexity is chosen to characterise task.

2.2.1 Person

The first dimension of the PAT model is person represented by individual-difference variables in this paper. Individual-difference variables selected for this research are need for cognition, self-efficacy (trait and state) and intrinsic motivation. The current study argues whether or not there is an influence of these variables moderated by task characteristics and artefact characteristics on user experience outcomes.

Need for cognition

The study of Cacioppo and Petty (1982) started the contemporary research on individual differences and presented the psychological construct of Need for Cognition (Cacioppo & Petty, 1996). Cacioppo and Petty (1996) state the construct of need for cognition as a 'stable individual difference in people's tendency to engage in and enjoy effortful cognitive activity' (p. 198). In other words, need for cognition refers to a stable intrinsic motivation towards cognitive activities.

According to Cacioppo and Petty's (1996) conceptualisation, individuals with high or low need for cognition respond to the world in different ways. For instance, individuals with a high need for cognition have a tendency to seek, acquire, think about, and reflect on information to make sense of stimuli, relationships, and events in their world. However, individuals with a low need for cognition are more likely to rely on others - such as celebrities and experts, cognitive heuristics or social comparison processes. In other words, people with low need for cognition are more likely to have a relative absence of an inclination to engage in and enjoy effortful cognitive activities compared to people with high need for cognition. Nevertheless, people with a high need for cognition are likely to a have motivation for effortful cognitive activities, and characterised generally as active and exploring minds. They are thought to be more likely to spend ‘effort on information acquisition, reasoning, and problem solving to cope with a wide variety of predicaments in their world’ (Cacioppo & Petty, 1996, p. 199). Therefore, compared to those with low need for cognition, individuals with a high need for cognition likely have more positive attitudes towards stimuli or tasks requiring reasoning and problem-solving. Furthermore, they likely have richer cognitive endeavours compared to individuals with low need for cognition, and have a chronic tendency to process information effortfully. The concept of need for cognition reflects a cognitive motivation of individuals; in particular, this cognitive motivation may lead to more knowledgeable responding to the information.

The study of Haugtvedt and Petty (1992) found that individuals with a high need for cognition persisted over time in their favourable attitudes towards a product to a greater extent than the attitudes of individuals with a low need for cognition. The study demonstrated that individuals with high need for cognition are likely to have a high persistence level in their attitudes towards products or services. Therefore, this study assumes people with high need for cognition to have a higher persistence level compared to those with low need for cognition.

Cacioppo and Petty (1984) elaborated a scale for the construct need for cognition, which is still used nowadays. They revised the original 34-item Need for Cognition Scale (NCS) developed by Cacioppo and Petty (1982) and presented this short form of NCS in their study. This 18-item NCS measures people’s tendencies to engage in effortful cognitive endeavours. This scale is described as measuring individual differences in intrinsic motivation for effortful cognitive endeavours (Thompson et al., 1993).

Self-efficacy

Self-efficacy is a key construct of social cognitive theory of Bandura (1977). The concept can be defined as ‘people’s judgements of their capabilities to perform a given task’ (Yi & Hwang, 2003, p. 434). In other words, self-efficacy reflects the confidence of individuals in accomplishing a particular task. Self-efficacy as an individual characteristic is an appropriate construct to use in order to investigate the research question of the current study because self-efficacy is a direct determinant of an individual’s behaviour supported by the literature (See Bandura, 1982). Bandura

(1989) states that ‘self-efficacy beliefs function as an important set of proximal determinants of human motivation, affect, and action.’ (p. 1175). Therefore, these self-efficacy beliefs likely affect thought patterns of individuals. This means that people’s beliefs in their self-efficacy influence the goals that they set for themselves. People with high self-efficacy perceive difficult tasks as some challenges to be accomplished rather than as threats to be eluded (Bandura, 1993). While computer self-efficacy as a state is assessed by using Compeau and Higgins’s (1995) 10-item scale, self-efficacy as a trait is measured by Generalised Self-efficacy Scale of Hong and O’Neil (2001).

The literature review on self-efficacy has demonstrated that self-efficacy as an individual-difference variable has been elaborated by researchers both as a trait (Hong & O’Neil, 2001) and as a state (Compeau & Higgins, 1995). These two aspects of the concept of self-efficacy are distinguished from each other in how they elaborate personality of individuals. Whilst states are attributes of individuals that are relatively changeable over time, traits are attributes of individuals that are relatively stable across occasions (Hong & O’Neil, 2001). Thus, it can be said that the degree of stability creates the distinction between these state and trait constructs. The present study takes into account both aspects of the self-efficacy construct in order to enhance the use of this individual difference variable. Therefore, both trait and state measurement scales are used in order to investigate participants’ personal differences better.

The impact of computer self-efficacy (CSE) as a state on task performance of users within a computer-mediated environment was investigated by Compeau and Higgins (1995). Their self-efficacy research demonstrated that CSE was a determinant of system use. Marakas et al. (1998) allege that computer self-efficacy (CSE) is a multi-level construct operating at two distinct levels, namely the general computing level (general CSE) and the specific application level (application-specific self-efficacy). In their paper, general CSE is defined as an individual judgement of efficacy across multiple computer domains, while application-specific self-efficacy is defined as an individual perception of efficacy in using a specific application or system. Although recent research on user acceptance of technology emphasised on general computer self-efficacy, the study of Yi and Hwang (2003) affirmed that application-specific self-efficacy and behavioural intention are determinants of actual system use. David et al. (2007) also defined computer self-efficacy which was used in their study as ‘perceived confidence in accomplishing the task at hand.’ (p. 174).

The study of Venkatesh and Davis (1996) empirically supported that self-efficacy is a key antecedent of perceived ease of use. In addition, the studies of Agarwal and Karahanna (2000) and Agarwal et al. (2000) also acknowledge the self-efficacy construct as a predictor of perceived ease of use.

Prior research agreed on the existence of the relation between the construct of self-efficacy and persistence of individuals (Jacobs et al., 1984; Bandura, 1989). Their findings, which are consistent with the statements of social cognitive theory, demonstrated that self-efficacy expectancies were the best predictor of persistence. Besides, Bandura (1989) asserted that self-efficacy of individuals are firmly related to their persistence. Therefore, the construct of self-efficacy is another sign of persistence of individuals besides the construct of need for cognition.

Intrinsic motivation

Csikszentmihalyi and Nakamura (1989) describe intrinsic motivation as ‘being associated with many of the factors characteristic of the mastery-oriented motivational pattern’ (cited in Heyman & Dweck, 1992, p. 239). Frey (1997) defined intrinsic motivation within an economical domain as ‘being motivated to do something without being forced by commands and without being (non-routinely) paid to do it’ (cited in Lindenberg, 2001, p. 319). A more operational definition of the construct is ‘the freely chosen continuation of an activity in free time, measured in seconds’ (Lindenberg, 2001, p. 318). Ryan and Deci (2000) asserted that individuals are intrinsically motivated only as activities hold intrinsic interest and appealing for them. The Situational Motivation Scale (SIMS) developed by Guay et al. (2000) is used to determine the intrinsic motivation of the participants. SIMS comprises 16 items.

This literature review indicated that some definitions of the constructs need for cognition and intrinsic motivation coincided on. The study of Cacioppo and Petty (1996), for instance, refers to a stable intrinsic motivation towards cognitive activities while defining need for cognition. In addition, Thompson et al. (1993) confirmed that there is a relation between the constructs of need for cognition and intrinsic motivation. Zhang and Buda (1999) also related the concept of need for cognition to intrinsic motivation. In their study, it was stated that individual differences among consumers in their desire to engage in, which holds the same meaning with need for cognition, was governed by their intrinsic motivation to process the message.

Even though some authors did not make a sharp distinction between these two concepts and connect those to each other, literature on intrinsic motivation indicated that these two have been treated separately by many researchers (e.g. Ryan & Deci, 2000; David et al., 2007; Haugtvedt and Petty, 1992; Lindenberg, 2001). Therefore, these concepts precisely carry a different significance. For instance, as claimed in the study of Ryan and Deci (2000), the construct of intrinsic motivation as an individual-difference variable is highly important:

‘perhaps no single phenomenon reflects the positive potential of human nature as much as intrinsic motivation, the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn’ (p. 70).

David et al. (2007) also states that intrinsic motivation deserves attention since it can encourage the exploration in computer-mediated environments. In the light of these statements, the last individual difference variable of this research is the construct of intrinsic motivation.

2.2.2 Artefact

Artefact, which is represented by artefact characteristics in this paper, is the second dimension of the PAT model. This section examines artefact characteristics with regard to both its components as aesthetics, design principles and information scent, and its complexity criteria. These artefact components likely have an influence on user experience of public-service Web sites because they affect perceptions of users. Moreover, regarding the research question, individual characteristics moderated by artefact complexity likely have an effect on user experience.

Aesthetics

Research on aesthetics in Human-Computer Interaction has considered the concept of aesthetics as an integral part of effective interaction design (Alben, 1996). Nevertheless, there are few, but increasing, number of research studies on the crucial role of aesthetics in a computer-mediated environment (Lavie & Tractinsky, 2004). Therefore, the current study is highly important for the field of Human-Computer Interaction as an exploratory research.

The concept of aesthetics is defined as ‘an artistically beautiful or pleasing appearance’ (The American Heritage Dictionary of the English Language), or as ‘concerned with beauty and art’ (Oxford Wordpower Dictionary of the English Language). Historically, ‘aesthetic values appeared as a reformulation of ideas about beauty, subsequently replacing them’ (Lavie & Tractinsky, 2004, p.271). Hassenzahl and Tractinsky (2006) see the concept of aesthetics as an aspect of user experience. In addition to this, Alben (1996) sees the aesthetics as an important part of the quality if technology.

In Lavie and Tractinsky’s (2004) conceptualisation, there are two dimensions of aesthetic design: classical aesthetics and expressive aesthetics. While the dimension of classical aesthetics emphasises orderly and clear design, the dimension of expressive aesthetics breaks (traditional) design conventions. In other words, classical aesthetics is substantially related to the design rules and reduces ambiguity whilst expressive aesthetics is reflected by designers' creativity and originality and increases arousal of users.

The findings of Lindgaard et al (2006) pointed out the importance of aesthetics for users' perceptions of Web sites, especially in making a good first impression. These authors concluded that a Web page's visual appeal, which was used to represent the concept of aesthetics and its dimensions as classic and expressive, can be assessed within 50 milliseconds by Web site users. However, they did not focus on the aesthetics perceptions after use.

Aesthetic dimensions are likely to have different effects on users' perceptions of aesthetics and finally may have an effect on users' acceptance of a specific service. This suggests that when a Web site is designed according to the characteristics of classical aesthetics, users' perceptions of these aesthetics should be higher than those of a site not designed in this way. For instance, the findings of van Schaik and Ling (2008b) demonstrated that Web pages designed based upon the rules of classic aesthetics were rated as more attractive than the pages designed according to the rules of expressive aesthetics. Moreover, a user's perceived aesthetic value of an artefact may have an impact on the user's flow experience or intention to use. For example, the study of Csikszentmihalyi (1990) found a clear association between aesthetics and flow experience.

Regarding the relation between usability and aesthetics, the study of Tractinsky et al. (2000) found that there is a strong relationship between aesthetics perceptions of users on Web pages and their perceptions of the usability. Besides, Lavie and Tractinsky (2004) found that usability of a system was highly associated with classic aesthetics over expressive aesthetics. Moreover, Tractinsky et al. (2000) stated an enduring influence of aesthetics perceptions even after the actual interface use. They claimed that 'What is beautiful is usable'. However, their study suffered from various methodological shortcomings (see Hassenzahl, 2004).

On the other hand, Hassenzahl's (2004) model of user experience takes a different perspective from that of Tractinsky et al. (2000). According to Hassenzahl, beauty is more related to self-oriented, hedonic attributes (pleasure-producing product qualities) of a product rather than to its goal-oriented, pragmatic attributes (user-perceived usability). In contrast to Tractinsky et al.'s (2000) conclusions, his findings supported this perspective and also indicated that there is only a weak relationship between beauty and pragmatic attributes. Hassenzahl's (2004) paper also revealed that perceived usability as well as goodness (i.e. an overall judgement of the quality of interaction with an artefact) was affected by experience, such as actual use and usability of the system, while beauty and hedonic attributes remained stable over time. This finding contradicts with the statements of Tractinsky et al. (2000), proposing an enduring influence of aesthetics perceptions even after the actual system use. Consistent with the findings of Hassenzahl (2004), van Schaik and Ling (2008) found evidence for the study of Hassenzahl (2004) and their study also did not support the Tractinsky et al. (2000)'s claim 'What is beautiful is usable'.

Lavie and Tractinsky (2004) found that usability of the Web sites was firmly associated with the classic aesthetics than with the expressive aesthetics. Although classical and expressive aesthetic dimensions compose the general aesthetic judgements, they are easily distinguishable from each other. Lavie and Tractinsky (2004) developed a Perceptions of Aesthetics Scale in order to measure these dimensions of aesthetics within the field of Human-Computer Interaction.

Design principles

According to van Schaik and Ling (2008a), the characteristics of classically aesthetics design and its purposes (e.g. reducing ambiguity) overlap with the principles of usable system designs in Human-Computer Interaction. Nonetheless, van Schaik and Ling (2008a) stated that there are some additional requirements of usable system design, such as their interaction structure, not provided by classically aesthetics design. Accordingly, even though classically aesthetics design makes a favourable contribution to design principles, it is not an adequate condition for a usable system design. On the other hand, the study of van Schaik and Ling demonstrated that there is not an overlap between the characteristics of expressively aesthetics design and the principles of usable system design.

van Schaik and Ling (2007) investigated the effect of design principles on users' task performance, navigation behaviour and perceptions. In their research, several user interface design principles were either complied with a Web site or violated in different versions of a Web site. The design principles comprise preserving the context of information units, using higher-order information units, avoiding gratuitous animations and providing consistency. Preserving the context of information refers to the headings of information units within a single page. Using higher-order information units is 'to enable readers to identify important components of the document and their relationships' (Dalal et al., 2000, p. 611). Using higher-order information units is a way of helping the user visualise the structure of a hyperdocument by organising them. An example of avoiding gratuitous animations is not to use unnecessary logo on a page. Providing consistency refers to the use of colour and type faces. The results of van Schaik and Ling supported that design principles had an effect on task performance of users.

Dalal et al. (2000) demonstrated that home pages designed in terms of cognitive design principles led to a better comprehension of information. In other words, Web pages providing cognitive design principles are comprehended better by users compared to Web pages designed with regard to non-cognitive design principles. Cognitive design represents coherence and low cognitive overhead. A lack of coherence and high cognitive overhead in a Web-based environment may bring about disorientation (Dalal et al., 2000) and thereby influence user experience.

Information scent

Information scent is an important concept in order to comprehend the factors affecting task performance of users. This concept is about how people use perceptual cues such as World Wide Web links to make an information-seeking decision (Pirolli, 2003). While choosing the U.K. e-government Web sites to use in this study, it was considered by the researcher that these two characterise a similar information scent and aesthetics.

Information scent was defined by Chi et al. (2001) as ‘the imperfect, subjective perception of the value and cost of information sources obtained from proximal cues, such as Web links, or icons representing the content sources.’ (p. 491). In a computer-mediated environment, the concept of information scent refers to users’ actions on the hyperlink, image or other screen objects that perceived by these users as the most semantically similar to their goals (Blackmon et al., 2002). Interface users such as Web users are heavily driven by information scent (Pirolli, 1997). They appraise and evaluate the environmental cues while judging information among Web pages and searching through information sources. When a word or phrase has little meaning for these users, this word or phrase offers inadequate information scent for them. This degree of semantic similarity, indeed, influences users’ Web-based behaviours such as navigating and information seeking, in particular as they are encountering a novel Web site or performing a task within a familiar Web-based context. Moreover, the findings of Blackmon et al. (2002) indicate that insufficient information scent yields an unfavourable effect on user performance due to users’ decreasing comprehension abilities with the unfamiliar information.

Web-site complexity

The literature on Web-site complexity distinguishes between objective and perceived Web-site complexity. On the one hand, objective complexity is defined by Nadkarni and Gupta (2007) as ‘a universal set of design characteristics that encompass the technological aspects of a Web site’ (p. 503) such as presentation formats, multimedia and search tools, hierarchical menu structure and download time for Web pages. On the other hand, perceived complexity is defined within Human-Computer Interaction literature and is based on users’ personal perceptions of a certain Web site and their interaction with it. However, literature also indicates that difference between the perceived Web site complexity and objective Web site complexity is not elaborated in depth by the authors. In spite of this shortage in the differentiation of the two Web site complexity aspects, the current study takes into account the objective Web-site complexity factors of both Nadkarni and Gupta (2007) and Ivory et al. (2001) in order to define artefact complexity. Objective Web site complexity represents the technological aspects of Web sites and provides this research to obtain more concrete results because its dimensions depend on the objective criteria of Web site evaluation.

This literature review presents several approaches on the basis of Web-site complexity metrics. These metrics make it possible to measure the level of complexity of a particular Web site. According to Tarasewich (2005), for instance, complexity of a Web site includes both some quantitative measures such as symmetry and balance and some Web specific metrics, such as the number of links in a site and their average depth, the number of graphics and their size, page length and width and the number of pages. He proposes that the time spent on a given task and error rate must also be measured besides complexity of Web sites. Tarasewich (2005) found that the Web pages which are redesigned with regard to the complexity metrics allowed users to perform information retrieval tasks better.

Geissler et al. (2006) state some factors that affect perceived home page complexity: home page length, number of graphics, number of links, amount of text and use of animation. The results of Geissler et al.’s (2006) study showed that number of graphics, number of links, and home page length were the key drivers of home page complexity.

Nadkarni and Gupta (2007) presented 13 elements from the recent research in their study in order to define objective Web site complexity: percentage of white space, graphics count, graphics size, word count, colour count, average number of different presentation forms used on a Web page (text, graphics, video, audio, animation), average internal and external links on the Web page, number of Web pages configuring a Web site, average depth of pages, coefficient of variation in the number of different presentation forms (e.g., text, graphics, video, audio, animation) used across Web pages, average pop-up advertisements per Web page, average Web page download time and number of support tools (e.g., site map, search option, help links).

As stated by Nadkarni and Gupta (2007), Ivory et al. (2001) also used a similar metric, including 11 elements as the determinants of objective Web-site complexity (see Appendix A): word count, percentage of body text, percentage of emphasised body text, text positioning count, text cluster count, link count, page size, percentage of graphics, graphics count, colour count and font count.

Juvina and Oostendorp (2006), collected Web-logging data in real time and logged interaction events such as page downloads, view time and use of buttons during a navigation session. Besides, some data on the Web structure being navigated was recorded (e.g. page title/ URL, number of words per page and number of outgoing/ incoming links). Results indicated that predictions of task outcomes based on user characteristics, interface and context factors appeared to be more accurate than those based on navigation metrics.

Banker et al. (1998) used Wood's dimensions of Web-site complexity in order to define their software complexity dimensions: software component complexity, software coordinative complexity and software dynamic complexity. Banker et al. (1998) measured data density by means of Halstead's N2 software science metric (Halstead, 1977). This metric counts the number of data variables in a programme. *Decision density* was measured by using McCabe's cyclomatic complexity (McCabe, 1976) which was also validated by the study of Rauterberg (1996). This metric count the number of decision paths through software programme and depends on graph theory. Cyclomatic complexity, v(G), is simply found by determining the number of decision statements in a program and is calculated as:

$v(G)$ = number of decision statements + 1

The cyclomatic complexity of McCabe and Butler (1989) uses three metrics such as module design complexity, design complexity, and integration complexity. Decision volatility was counted by the number of decision paths that were dynamically altered at software execution time.

2.2.3 Task

Task, which is represented by task characteristics in this paper, is the last dimension of the PAT model. In this section, task characteristics are demonstrated in terms of its complexity. Regarding the research question, individual characteristics moderated by task complexity likely have an effect on user experience.

Task complexity

Wood (1986) defined task complexity as ‘a function of the number of distinct acts that must be completed and the number of distinct information cues about the attributes of the task-related stimulus object an individual has to process when performing a task.’ (cited in Jiang & Benbasat, 2007, p. 481).

The literature on task complexity evaluates objective and perceived complexity of tasks as two distinct constructs (Nadkarni & Gupta, 2007). Whilst the objective complexity is defined by the number of information cues in the task stimulus, perceived complexity emphasises on the person-stimulus interaction. Objective complexity of a task influences the cognitive load related to task performance since users have limited cognitive resources (Lindsay & Norman, 1977). This statement is consistent with Miller’s (1956) study. According to Miller (1956), in information process people have some limits in their working memory. Miller (1956) argued the number seven, plus or minus two, as a limit to be hold in people’s working memory. High task complexity may cause too much cognitive effort of individuals leading to a simplification of task execution strategies (Jiang & Benbasat, 2007).

Jiang and Benbasat (2007) used Miller’s (1956) rule in their study in order to investigate the moderator effect of task complexity on the effectiveness of online product presentation formats, namely static pictures, video-without-narration, video-with-narration and virtual product experience. Their findings revealed that both videos (–with and -without narration) and virtual product experience, enabling users virtually touch, feel and try products, led to higher perceived Web site diagnosticity (i.e., the extent to which consumers believe a Web site is helpful for them to understand products) than static pictures; in a high task complexity condition, the superiority of videos (–with and -without narration) and virtual product experience shrunk; and in a moderate task complexity condition, virtual product experience and videos (–with and -without narration) were more effective than static pictures. The study of Jiang and Benbasat (2007) highlights the significant influence of task complexity on product understanding.

According to Wang et al. (2005), task complexity is ‘an objective concept describing the relationship among elements (e.g., task requirements, skills required, resources needed, steps to be taken, etc.) involved in completing an information task in this study’ (p.6). These authors presented two aspects describing the complexity of a task: the number of elements involved in completing the task and the interrelationship among these elements. The more elements involved in completing a task, the more interrelationships will exist among them, and the more likely these relationships become complicated. In addition, Ebbers et al. (2008) defined task complexity as ‘the extent of multiple interrelated actions that have to be taken to solve one problem’ (p.191). Task complexity is likely to depend on the degree of uncertainty on the task inputs, process, and outcomes (Ebbers et al., 2008).

Campbell (1988) defined task complexity as a function of the task performer's psychological states, the interaction between the abilities of the task performer and the characteristics of task, and the objective attributes of the task such as the number of sub-tasks or the uncertainty of the task outcome. In this definition, task complexity is a combination of task performer's subjective characteristics with objective attributes of the task itself. Besides, Vakkari (1999) argued that task complexity is not an objective measure since personal factors can affect an individual's assessment of the complexity of a task. Therefore, subjective complexity should be considered besides objective complexity of the task.

March and Simon (1958) defined some characteristics of objective task complexity. According to them, complex tasks include unknown or uncertain alternatives and consequences, inexact means-ends connections, and number of sub-tasks. Consistent with March and Simon (1958), Campbell (1988) proposed four attributes increasing the complexity of the task: multiple potential paths to a desired end-result, the presence of multiple desired outcomes, the presence of conflicting interdependencies between paths and uncertainty regarding paths. According to Campbell (1988), a simple task contains none of these attributes. Frese (1987) stated that complexity is determined by both the number of decisions that have to be made and the relation among these decisions.

Lazonder et al. (2000) used a 'locate Web site' task differing in complexity levels (low/simple, medium and high/complex) in their study. While URL can be easily inferred from the task description in medium complexity level, in low and high complexity levels URL is either given in the task description or cannot be inferred from the task description. Kuiken and Vedder (2008) used six requirements in complex task condition while using three requirements in low task condition in order to specify cognitive task complexity levels of their study. However, their findings did not present any significant difference between complex and non-complex tasks.

Bell and Ruthven (2004) used the classification of Byström and Järvelin (1995) to define the levels of task complexity. However, they narrowed five complexity levels into three levels in order to make the difference more clear. The first task complexity level of Bell and Ruthven (2004) (low task complexity level) comprises the tasks which are *a priori determinable*. It is clear what information is required for accomplishing these tasks. The second task complexity level of authors (medium task complexity level) comprises the tasks in which the desired information may be clear. However, the individual must make decisions in this level. The last task complexity level of authors (high task complexity level) comprises the tasks in which both the desired information and how to obtain this information are unclear. To detect the factors affecting task complexity, they used three questions with a 5-point scale ranging from 1 (lowest level of agreement) to 5 (highest level of agreement): ‘Useful information was provided by the task’, ‘The type of information to be retrieved was clear’ and ‘The amount of information to be retrieved was clear’. To measure the subjective factors affecting task complexity, authors used three different questions with a 5-point scale ranging from 1 (lowest level of agreement) to 5 (highest level of agreement): ‘This task was easy to understand’, ‘The task was interesting’ and ‘The task was relevant to me’. The findings of Bell and Ruthven indicated that their model borrowed from Byström and Järvelin (1995) was validated as a means of predicting and manipulating task complexity. Authors recommended that assessing task complexity in pilot or pre-testing can be a useful method for determining whether tasks are appropriate for individual evaluations.

Horvath et al. (2006) used four questions ranged from easy to difficult in order to measure perceived task difficulty: ‘How difficult do you feel this class is for you’, ‘How difficult do you feel this class is for the average class member’, ‘How difficult do you think the professor is’ and ‘How difficult is this subject for you.

According to Kieras and Polson (1999), the structure of a task is characterised by the hierarchy of the goals and subgoals fulfilled during the task performance. On the other hand, Vakkari (1999) the structure of a task refers to the elements of the task and their interrelations.

In the light of literature review on task complexity, measures of task complexity used in the current study include length of task sequence (number steps in a task or a goal-oriented or an interaction task) and number of decision points. Regarding this decision, simple tasks require 1-2 steps from participants, while complex tasks require 4-5 steps from them.

2.3 Psychological model of users’ experience (PAT-UX)

This section aims to elaborate several outcome measures as a consequence of users’ experience through Web sites. At the end of this chapter, the psychological model of users’ experience of public-service Web sites (PAT-UX) developed by the current study will be demonstrated with regard to the hypotheses proposed.

2.3.1 User experience outcomes

Several user experience outcomes such as mental effort, enjoyment, perceptions of aesthetics and disorientation will be used in the current study. In addition, technology acceptance variables such as perceived usefulness, perceived ease of use and intention to use are important outcome measures in this study as potential gains in user experience.

Mental effort

This literature review indicated that mental effort is one of the user experience outcomes through a Web-based environment. It is defined by Arnold (1999) as the amount of energy a user has to activate to meet perceived task demands (cited in Hassenzahl & Ullrich, 2007). Because experienced mental effort is a significant predictor of usability problems, it is highly related to experienced barriers in goal attainment of users (Hassenzahl & Ullrich, 2007). In addition, mental effort is a subjective and goal-related concept. Hassenzahl and Ullrich's (2007) results indicate that mental effort is affected by the active instrumental goal which increases the mental effort. Furthermore, it is experienced negatively when a user is pursuing the instrumental goal. van Schaik and Ling (2008a) also found that mental effort and task performance are predictors of 'goodness', or overall quality. Thus, it can be said that mental effort influences the user experience. Mental effort of participants can be measured by the Subjective Mental Effort Questionnaire (SMEQ) of Zijlstra (1993) ranging from 0 (hardly effortful) to 220 (exceptionally effortful).

Enjoyment

Enjoyment was elaborated by the literature as a result of user experience through Internet-based systems. According to Csikszentmihalyi (1990), enjoyment occurs:

'When a person has not only met some prior expectation or satisfied a need or a desire but also gone beyond what he or she has been programmed to do and achieved something unexpected, perhaps something even unimagined before' (p. 46).

He argues that enjoyment requires from users a cognitive effort and attention, and uses the concept of optimal experience to define two concepts: flow and enjoyment. Csikszentmihalyi also uses the terms flow and enjoyment interchangeably. Enjoyment is often used by him as the result of flow experience of individuals.

Some researchers argue that enjoyment which is derived from an activity is central to intrinsic motivation (Deci & Ryan, 1985; Jackson & Marsh, 1996). This means that enjoyment used as autotelic experience in Jackson and Marsh's (1996) study is an intrinsically performed experience. A person then undertakes the activity for its own sake. As recommended by Moneta and Csikszentmihalyi (1996), enjoyment should be considered while studying user experience.

Perceptions of aesthetics

As mentioned previously, aesthetics is an important aspect of Web site evaluation and aesthetics perceptions of individuals can be determined by user experience. Beauty has a positive effect on the perceptions of individuals in terms of their attitudes toward products. Hassenzahl (2004) found that beautiful products or products with aesthetics were perceived more stimulating than others. In addition, Lavie and Tractinsky (2004) developed a measurement model in order to explore users' perceptions of aesthetics of Web sites. According to van der Heijden (2003), beauty affects perceptions of other Web site qualities as the most important determinant of preferring a Web site.

Disorientation

The concept of disorientation is firmly related to user experience through a Web site. Disorientation can be defined as 'the tendency to lose one's sense of location and direction in a non-linear document (Ahuja & Webster, 2001, p. 16) and it relates to task performance. For instance, when a Web user is disoriented, she or he is likely to have some difficulties in finding Web pages known to exist or finding pages already visited (Pitkow & Kehoe, 1996; cited in Ahuja & Webster, 2001). The study of Ahuja and Webster (2001) indicates that perceived disorientation is related to differing Web designs and therefore the measurement of perceived disorientation may be useful to evaluate the usability of Web pages. Furthermore, disorientation can have an impact on the acceptance of Web sites or can be affected by these Web sites (van Schaik & Ling, 2007). van Schaik and Ling (2003, 2005, and 2007) found that disorientation, which can be seen a more specific lack of perceived ease of use, is a better measure than perceived ease of use in terms of sensitivity to design principles.

Perceived usefulness and perceived ease of use

As mentioned before, the concepts of perceived usefulness and perceived ease of use are basis of technology acceptance of individuals. These two are likely to be affected by user experience and indicate the level of technology acceptance of individuals. Perceived usefulness is defined as 'the extent to which a person believes that using a particular system will enhance his or her job performance', while perceived ease of use is defined as 'the extent to which a person believes that using a particular system will be free of effort' (Davis, 1989, p.320).

Perceived behavioural intention

Perceived behavioural intention is also one of user experience and technology acceptance outcomes. Fishbein and Ajzen (1975) defined behavioural intention as 'the subjective probability that an individual will take a particular action' (cited in Gotlieb et al., 1994, p. 875), such as purchasing a product. People are expected to carry out their intentions when the opportunity arises (Ajzen, 2002). Therefore, perceived behavioural intention is assumed to be the immediate antecedent of behaviours and used to predict behaviour of users (Venkatesh et al., 2003).

2.3.2 Model and hypotheses

The current study developed a psychological model of users' experience of public-service Web sites (PAT-UX model) in order to find an answer to research question (Figure 2). In general, this psychological model suggests that individual-difference variables are moderated by artefact complexity and task complexity and this moderation influences user experience outcomes of the use of public-sector Web sites.

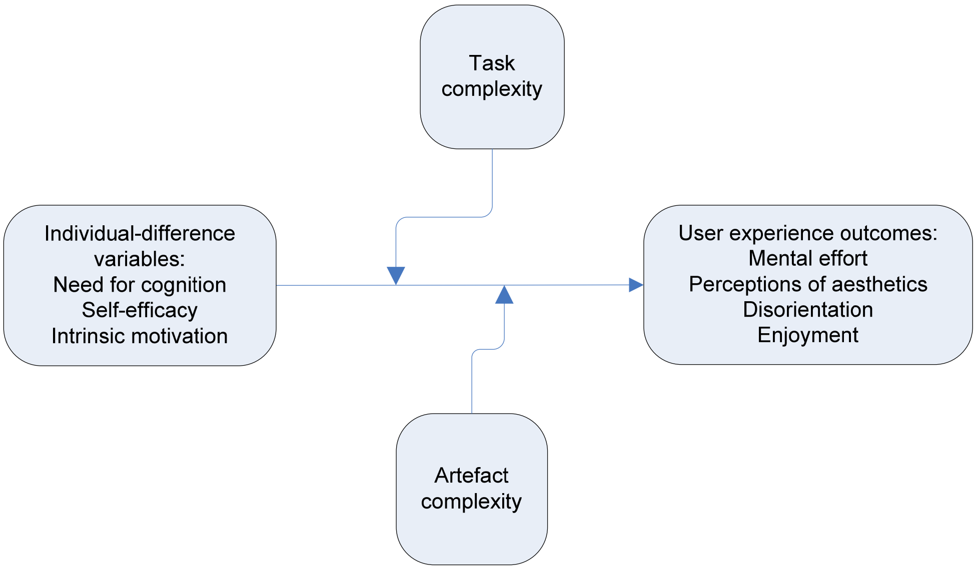

Figure 2 Suggested psychological model of users' experience of public-service Web sites (PAT-UX model)

Regarding the theoretical framework of the current study presented, the following hypotheses are proposed:

> H1: People with a high need for cognition experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceived enjoyment, higher perceptions of aesthetics (classic and expressive), higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low need for cognition.
>
> H2: People with a high self-efficacy (trait and state) experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceived enjoyment, higher perceptions of aesthetics (classic and expressive), higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low self-efficacy (trait and state).
>
> H3: People with a high intrinsic motivation experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceived enjoyment, higher perceptions of aesthetics (classic and expressive), higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low intrinsic motivation.
>
> H4: High task complexity leads to more mental effort, higher perceived disorientation, lower perceived usefulness, lower perceived ease of use, lower perceived enjoyment, lower perceptions of aesthetics (classic and expressive), lower task performance and navigation behaviour, and lower perceived behavioural intention than low task complexity.

H5: High artefact complexity leads to more mental effort, higher perceived disorientation, lower perceived usefulness, lower perceived ease of use, lower perceived enjoyment, lower perceptions of aesthetics (classic and expressive), lower task performance and navigation behaviour, and lower perceived behavioural intention than low artefact complexity.

H6: The effect of individual-difference variables on quality of services is moderated by artefact complexity and task complexity.

H6a: the effect of individual-difference variables (need for cognition, self-efficacy and intrinsic motivation) is moderated by task complexity: for individuals with a high need for cognition, high self-efficacy (trait and state) and high intrinsic motivation the effect of task complexity is smaller than that for individuals with a low need for cognition, low self-efficacy (trait and state) and low intrinsic motivation.

H6b: the effect of individual-difference variables (need for cognition, self-efficacy and intrinsic motivation) is moderated by artefact complexity: for individuals with a high need for cognition, high self-efficacy (trait and state) and high intrinsic motivation the effect of artefact complexity is smaller than that for individuals with a low need for cognition, low self-efficacy (trait and state) and low intrinsic motivation.

2.4 Conclusion

In this chapter, the Person-Artefact-Task (PAT) model with several perspectives from literature on user experience and technology acceptance was elaborated. In addition, each dimension of the PAT model was characterised by some constructs chosen from the literature, contributing to user experience. For instance, person was characterised by need for cognition, self-efficacy (trait and state) and intrinsic motivation, while the characteristics of an artefact (Web site) consists of aesthetics, design principles and information scent. Task was characterized by its complexity level. Moreover, a psychological model of users' experience of public-service Web sites (PAT-UX) model was developed with several hypotheses to test the relation between these dimensions and user experience. In the next chapter, the influence of these three will be empirically investigated in response to testing both the hypotheses and the model.

3. Method

In this chapter, method of the study will be elucidated, including the elaboration of the experimental design, participants, materials of the experiment, procedure, reliabilities of the scales used in the experiment and relations between individual-difference variables and user experience outcomes.

3.1 Experimental design

A between subjects experimental design was used with two independent variables (Figure 2). The two independent variables were used as moderators of focal variables. The focal variables were the individual-difference variables, namely need for cognition, computer self-efficacy and intrinsic motivation. The first moderator was artefact complexity with two levels (high and low). The second moderator was task complexity with two levels (high and low). It was not possible to create a (2×2) full factorial design because the combination of high task complexity and low artefact complexity was not possible, given the well-considered choice of Web sites made. The dependent variables included perceived mental effort, perceived disorientation, perceived usefulness, perceived ease of use, perceived enjoyment, perceptions of aesthetics, measures of task performance and navigation behaviour, and perceived behavioural intention.

Participants were randomly assigned to one of three experimental conditions generated by a combination of (public-service) Web site and task complexity. The first condition represented a high artefact complexity with high task complexity. The second condition represented a high artefact complexity with low task complexity. The third condition represented a low artefact complexity with low task complexity. Artefact complexity was formed by using two different Web sites with simple tasks (requiring one or two steps to complete). Task complexity was varied within the same (complex) Web site. Tasks required either one or two steps (low task complexity) or four or five steps (high task complexity), respectively, to complete tasks. Two UK council Web sites used in the experiment: a simple Web site and a complex Web site (Figure 3 and Figure 4). They were chosen with regard to a set of Web site complexity metrics borrowed from the studies of Nadkarni and Gupta (2007) presented previously and Ivory et al. (2001) presented in Appendix A. Regarding the number of steps which makes a task either complex or simple, ten information retrieval questions were formulated in the two public-service Web sites by the researcher. All information retrieval questions used in the experiment were presented in Appendix B.

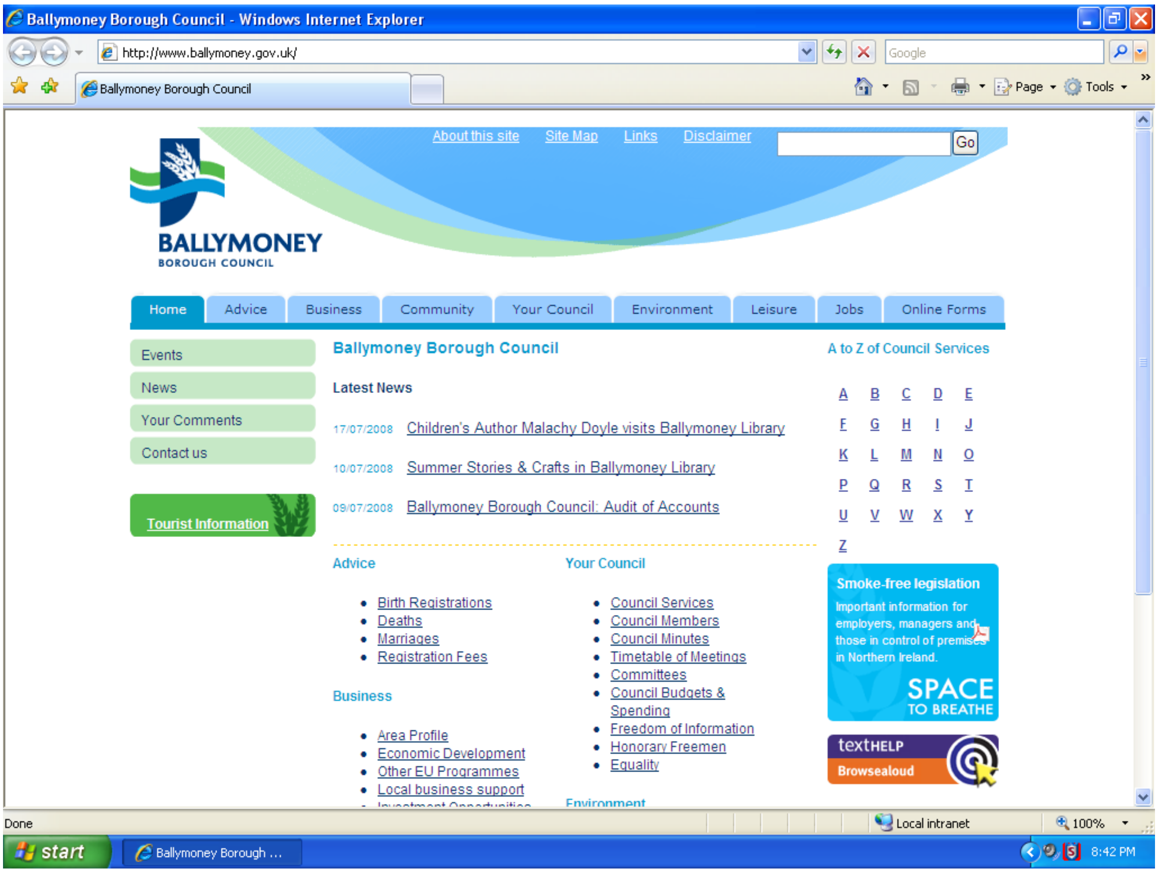

Figure 3 Home page of simple Web site

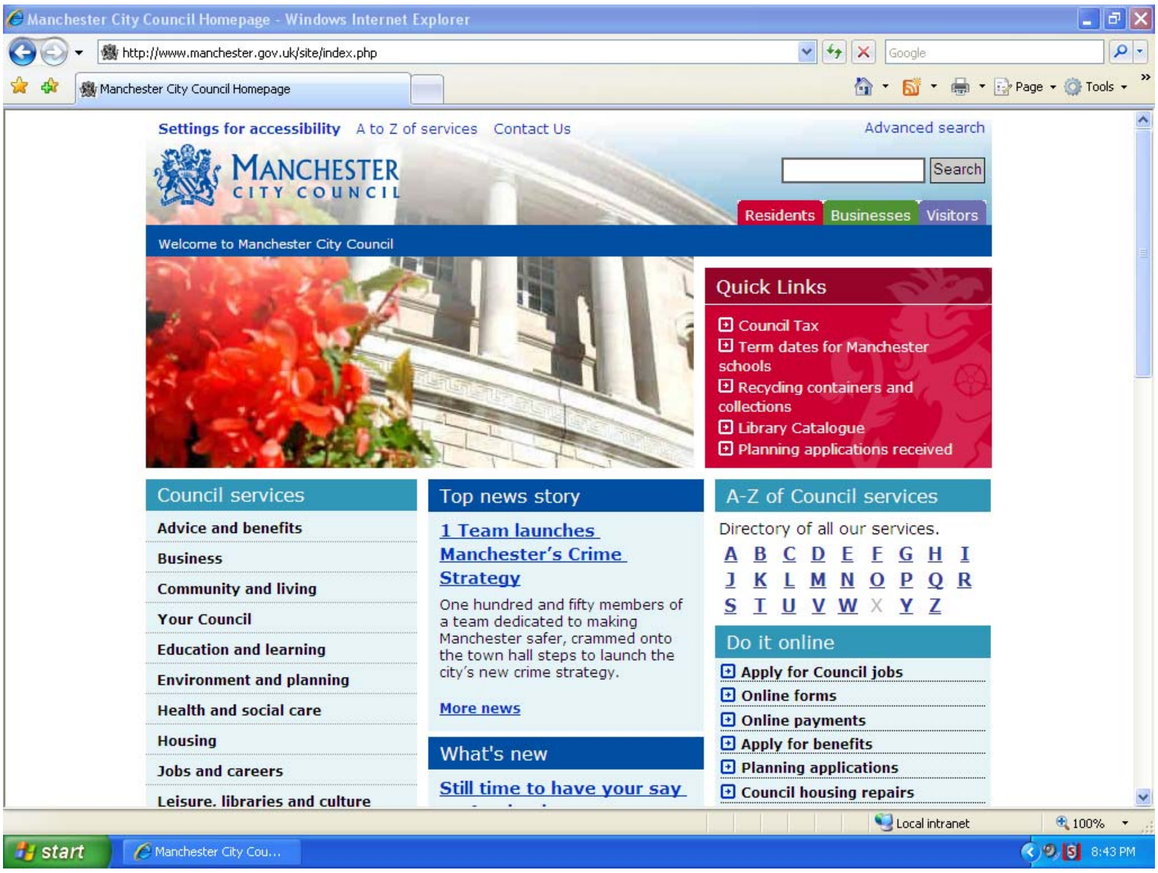

Figure 4 Home page of complex Web site

Besides this experimental design, a think-aloud method was executed in order to investigate the research question. This qualitative method to assess an artefact requires participants to verbalise their thoughts while performing tasks (van den Haak et al., 2004). All participants (those who were required to think aloud and those who were not) were randomly assigned to each of the three experimental conditions.

3.2 Participants

Power analysis reveals that in order for a t test to achieve a statistical power of 0.80 for a large effect size (d = 0.80) a sample size of 52 (= 2×26) was required for each of the independent variables task complexity and artefact complexity (comparing high and low complexity for each variable). Based on this power analysis, the total sample size would be 2×26 + 2×26 - 26 (the condition of low task complexity and high artefact complexity would be used once only).

Sixty-one undergraduate students (36 females, 10 males and 15 missing genders) enrolled in the Faculty of Communication and Behavioural Sciences, in University of

Twente, The Netherlands, participated in the experiment for one course credit. Demographics of 15 participants were missing. The ages of remaining 46 participants had a mean of 22.54, while the standard deviation was 4.42. Participants' nationalities were as follows: Dutch (66%), German (18%), and Other (16%). Mean first year of using the Web sites was 1999.26 (SD = 3.01), mean first year of feeling confident of using the Web sites was 2000.89 (SD = 3.14), mean time per week spent using the Web sites was 19.50 hours (SD = 13.56) and mean frequency of Web site use per week was 18.17 times (SD = 12.60).

Another eight undergraduate students (five females and three males; eight Dutch and one German) enrolled in the Faculty of Communication and Behavioural Sciences, in University of Twente, The Netherlands, participated for 5 euros in the experiment using the think-aloud method and one male student of applied physics took part. Their mean age was 22.11 (SD = 1.62). Mean first year of using the Web sites was 1998.56 (SD = 2.01), mean first year of feeling confident of using the Web sites was 2000.44 (SD = 2.55), mean time per week spent using the Web sites was 13.11 hours (SD = 8.81) and mean frequency of Web site use per week was 18.22 times (SD = 9.74).

3.3 Equipment and materials

The 18-item Need for Cognition Scale (NCS) borrowed from the study of Cacioppo and Petty (1984) measures people's tendencies to engage in effortful cognitive endeavours. This scale was used in the experiment in order to measure participants' need for cognition level. The scale range was 1 (extremely uncharacteristic) to 7 (extremely characteristic). Level of self-efficacy was measured by using both the 10-item computer self-efficacy scale of Compeau and Higgins (1995) and the 10-item self-efficacy scale of Hong and O'Neil (2001). The range of both self-efficacy scales was 1 ('not at all confident') to 4 ('moderately confident') to 7 ('totally confident'). The Situational Motivation Scale (SIMS) of Guay et al. (2000) including 16 items was used to measure intrinsic motivation of participants. The scale ranged from 1 ('corresponds not all') to 7 ('corresponds exactly').

In order to measure the dependent variables of the experiment, several measurement scales were used. The Subjective Mental Effort Questionnaire (SMEQ) of Zijlstra (1993) ranging from 0 ('hardly effortful') to 220 ('exceptionally effortful') was to measure mental effort of participants. All the following questionnaires used 7-point scales. The Disorientation Scale of Ahuja and Webster (2001) was used to measure disorientation. The perceived usefulness and perceived ease of use measures of Davis (1989) was used to measure perceived usefulness and ease of use the conditions. The Flow State Scale of Jackson and Marsh (1996) was to measure the perceived enjoyment of participants. To measure perceive behavioural intention, the measurement scale in Venkatesh et al. (2003) was used. Perceptions of aesthetics were measured by Lavie and Tractinsky's (2004) aesthetics scale, with dimensions classical aesthetics and expressive aesthetics. All the questionnaires used in the experiment were presented in Appendix C.

A bespoke experimental program was developed and employed to control the experiment. The program recorded participants' use of the Web site they were using in the experiment, including each page and time spent on the page. The program also administered the measurement scales and demographics including age, gender, experience with using the Web. Philips Digital Voice Tracer 7655 was used to record nine participants' voices in think-aloud protocols.

3.4 Procedure

The experiment was held in a computer laboratory. Before the experiment, participants were given general instructions regarding the scenario of the experiment and the approximate duration of the experiment.

They then answered three questionnaires on their computer screen orderly in order to measure their levels of need for cognition, computer self-efficacy and self-efficacy as a treat, and intrinsic motivation.

Then, a series of three practice tasks with use of a UK e-government Web site followed, presented in fixed order. Each task was an information retrieval question. Participants were told that each of a series of questions would appear at the top of the screen. After reading the question they had to click on the button labelled 'Show Web site'. The home page of the first site was then displayed in the browser window. Participants were instructed to find the answer to each question using the site and were told to take the most direct route possible to their answers. Once they found the answer they had to click a button labelled 'Your answer'. A dialog box then appeared into which participants entered their answer. Then they were shown the next question. After the practice tasks then main tasks followed, using the same procedure. The recording of participants' online navigation behaviour allowed measures of task performance and navigation behaviour to be calculated after data collection.

Mental effort of participants was measured after completing each task. After finishing the main tasks, participants completed the measurement scales for the following constructs: perceived disorientation, perceived usefulness, perceived ease of use, perceived enjoyment, perceptions of aesthetics and perceived behavioural intention. The think-aloud procedure was also done in a computer laboratory. Before the experiment, participants were given think-aloud instructions and were rewarded by 5 Euros after the sessions.

Nine think-aloud sessions for all three versions of the experiment were conducted in the study to explore any usability problems and other aspects of participants' user experience (e.g. perceptions of aesthetics and enjoyment) from their verbalisations and interaction with the Web sites. Participants were randomly assigned to one of the three experimental conditions: two participants for experimental condition 1 (high artefact complexity with high task complexity), three participants for experimental condition 2 (high artefact complexity with low task complexity) and four participants for experimental condition 3 (low artefact complexity with low task complexity). Although equal number of participants had been planned in think-aloud experiment before starting the sessions, there were unequal number of participants in each experimental condition because of an unknown breakdown in the experimental programme. Before starting the experiment, participants were informed about how to act in the think-aloud experiment (for instructions see Appendix D). As in the main experiment, participants first answered psychological questionnaires by themselves; then they started thinking aloud whilst they were performing the tasks on the Web site. During each session, an audio recorder was recording participants' voice. After finishing their tasks on the Web site, they answered the questionnaires without thinking aloud. When nine sessions were completed, verbal transcripts were made along with the researcher's comments on the participants' actions during the sessions.

Data from think-aloud sessions and participants' answers to three open-ended questions in the main experiment were utilised to explore the qualitative results of the experiment. Think-aloud data included the combination of the data through verbalisation by the participant and researcher's comments through observation. The analysis was focused on two main issues such as classification of usability problems detected and several aspects of participants' user experience such as favourable aspects of user experience, perceptions of aesthetics and enjoyment. These issues were categorised on the basis of general terms of usability and aspects of user experience, such as enjoyment, aesthetics and flow. Terminology problems, problems in understanding English, were also considered in the analysis of the data.

3.5 Reliabilities

This section demonstrates the reliability scores of questionnaires used in the experiment. Because the current study has two kinds of data, namely non-think-aloud and think-aloud, reliability scores of both data are presented below.

Values above 0.7 are considered acceptable. Published results indicated that all the scales borrowed from the previous studies had good internal consistency with a sufficient Cronbach's alpha coefficient. In addition, it is clear that the scales had a high degree of reliability in the current study, except the subscale of 'identified motivation' of SIMS and the subscale of 'transformation of time' of Flow State Scale. The reliability of scales from non-think-aloud data was presented inTable 1.

Table 1 Reliability of scales (non-think-aloud data)

Scales	Cronbach's alpha coefficients
Need for Cognition Scale	0.75
Computer Self-efficacy Scale	0.84
Generalised Self-efficacy Scale	0.76
Situational Motivation Scale (SIMS): subscale 'intrinsic motivation'	0.80
Situational Motivation Scale (SIMS): subscale 'identified regulation'	0.61
Situational Motivation Scale (SIMS): subscale 'external regulation'	0.88
Situational Motivation Scale (SIMS): subscale 'amotivation'	0.83
Perceptions of Aesthetics: subscale classic aesthetics	0.84
Perceptions of Aesthetics: subscale expressive aesthetics	0.74
Flow State Scale: subscale 'challenge-skill balance'	0.83
Flow State Scale: subscale 'action-awareness merging'	0.82
Flow State Scale: subscale 'clear goals'	0.81
Flow State Scale: subscale 'unambiguous feedback'	0.83
Flow State Scale: subscale 'concentration on task at hand'	0.83
Flow State Scale: subscale 'paradox of control'	0.81
Flow State Scale: subscale 'loss of self-conscious'	0.79
Flow State Scale: subscale 'transformation of time'	0.64
Flow State Scale: subscale 'autotelic experience'	0.85
Perceived Ease of Use Scale	0.90
Disorientation Scale	0.91
Perceived Usefulness Scale	0.83
Behavioural Intention Scale	0.94

Table 2 Reliability of scales (think-aloud data)

Scales	Cronbach's alpha coefficients
Need for Cognition Scale	0.86
Computer Self-efficacy Scale	0.82
Generalised Self-efficacy Scale	0.85
Situational Motivation Scale (SIMS): subscale 'intrinsic motivation'	0.43
Situational Motivation Scale (SIMS): subscale 'identified regulation'	0.02
Situational Motivation Scale (SIMS): subscale 'external regulation'	0.36
Situational Motivation Scale (SIMS): subscale 'amotivation'	0.11
Perceptions of Aesthetics: subscale classic aesthetics	0.48
Perceptions of Aesthetics: subscale expressive aesthetics	0.79
Flow State Scale: subscale 'challenge-skill balance'	0.94
Flow State Scale: subscale 'action-awareness merging'	0.67
Flow State Scale: subscale 'clear goals'	0.71
Flow State Scale: subscale 'unambiguous feedback'	0.76
Flow State Scale: subscale 'concentration on task at hand'	0.91
Flow State Scale: subscale 'paradox of control'	0.93
Flow State Scale: subscale 'loss of self-conscious'	0.77
Flow State Scale: subscale 'transformation of time'	0.87
Flow State Scale: subscale 'autotelic experience'	0.74
Perceived Ease of Use Scale	0.88
Disorientation Scale	0.89
Perceived Usefulness Scale	0.97
Behavioural Intention Scale	0.98

When item 10 in the subscale of ‘identified motivation’ of SIMS was deleted, the alpha score was 0.754. Besides, the alpha score was 0.742 as item 8 in the subscale of ‘transformation of time’ of Flow State Scale. Therefore, data analysis was repeated by removing these two items in order to both increase reliability of scales and compare the results with the non-think-aloud data analysis. The results of this re-conducted data analyses were also presented in the rest of the study and called ‘repeated non-think-aloud data analyses’ accompanying ‘non-think-aloud data analyses’.

The Cronbach’s alpha coefficient scores of think-aloud data was presented inTable 2. Reliability analysis showed that there were several alpha scores lower than 0.7. However, none of the items was removed in order to repeat the analyses because the number of participants was extremely low to analyse. Therefore, the current study focused on the qualitative results of think-aloud data instead of quantitative results of this data.

3.6 Relationships of variables and data

In this section, correlation coefficients of individual-difference variables and outcome measures are demonstrated.

3.6.1 Relations of individual-difference variables and outcome measures in non-think-aloud data

The relationship between individual-difference variables and outcome measures for original data was explored by means of using Pearson product-moment correlation coefficient. These correlation coefficients were presented inTable 3. The strength and direction of these variables were as follows:

There was a weak, negative correlation between need for cognition and intrinsic motivation, r = -0.11; a weak, negative correlation between need for cognition of individuals and mental effort, r = -0.18; a weak, negative correlation between need for cognition of individuals and perceptions of aesthetics subscale of classic aesthetic, r = -0.15; a weak, negative correlation between need for cognition of individuals and perceptions of aesthetics subscale of expressive aesthetic, r = -0.25; a weak, positive correlation between need for cognition of individuals and enjoyment, r = 0.22; a highly weak, negative correlation between need for cognition of individuals and perceived ease of use, r = -0.04; a highly weak, negative correlation between need for cognition of individuals and disorientation, r = -0.07; a highly weak, negative correlation between need for cognition of individuals and perceived usefulness, r = -0.09; a highly weak, negative correlation between need for cognition of individuals and behavioural intention, r = -0.09. None of these correlations was statistically significant at 0.05 significance level.

There was a weak, positive correlation between need for cognition of individuals and time-on-practice tasks, r = 0.15; a highly weak, negative correlation between need for cognition of individuals and number of main tasks completed, r = -0.07; a highly weak, positive correlation between need for cognition of individuals and time-on-main tasks, r = 0.09; a highly weak, positive correlation between need for cognition of individuals and percentage of correct main answers, r = 0.01; a highly weak, negative correlation between need for cognition of individuals and average pages loaded for correct answers, r = -0.08; a weak, negative correlation between need for cognition of individuals and average time-on-task for correctly answered main questions, r = -0.14; a weak, negative correlation between need for cognition of individuals and average number of pages loaded for incorrectly answered main questions, r = -0.12; a weak, negative correlation between need for cognition of individuals and average time-on-task for incorrectly answered main questions, r = -0.12; a highly weak, positive correlation between need for cognition of individuals and average number of correct answers per answered question, r = 0.09; a highly weak, negative correlation between need for cognition of individuals and average number of visited pages per answered question, r = -0.09; a highly weak, negative correlation between need for cognition of individuals and average number of revisited pages per answered question, r = -0.04; a weak, negative correlation between need for cognition of individuals and average number of times of visiting the homepage per answered question, r = -0.15; a weak, positive correlation between need for cognition of individuals and average number of times of visiting a page with search results per answered question, r = 0.10; a highly weak, positive correlation between need for cognition of individuals and average number of times of visiting the site map per answered question, r = 0.01; a weak, negative correlation between need for cognition of individuals and average time-on-task per answered question, r = -0.12. Again, none of these correlations was statistically significant at 0.05 significance level.

Table 3 Correlations (non-think-aloud data)

	Need for Cognition	Self-efficacy as Trait	Self-efficacy as State	Intrinsic Motivation
mental effort	-0.18	-0.06	0.10	0.06
perceptions of aesthetics subscale of classic aesthetic	-0.151	0.06	0.06	0.12
perceptions of aesthetics subscale of expressive aesthetic	-0.25	0.20	-0.07	0.15
enjoyment	0.22	0.12	*0.29	0.02
perceived ease of use	-0.04	0.21	0.13	0.02
disorientation	-0.07	-0.10	-0.15	-0.13
perceived usefulness	-0.09	-0.02	0.11	-0.12
behavioural intention	-0.09	0.19	0.09	0.05
time-on-practice tasks	0.15	-0.09	-0.01	-0.004
number of main tasks completed	-0.07	-0.02	0.08	-0.09
time-on-main tasks	0.09	-0.09	-0.05	0.09
percentage of correct main answers	0.01	-0.01	0.10	-0.15
average pages loaded for correct answers	-0.08	0.10	-0.06	0.14
average time-on-task for correctly answered main questions	-0.14	0.07	-0.15	0.22
average number of pages loaded for incorrectly answered main questions	-0.12	-0.03	-0.04	0.05
average time-on-task for incorrectly answered main questions	-0.12	0.0001	-0.07	0.22
average number of correct answers per answered question	0.09	-0.01	-0.08	-0.10
average number of visited pages per answered question	-0.09	0.04	0.10	0.05
average number of revisited pages per answered question	-0.04	-0.01	0.12	0.04
average number of times of visiting the homepage per answered question	-0.15	-0.04	-0.02	0.15
average number of times of visiting a page with search results per answered question	0.10	-0.14	-0.10	0.05
average number of times of visiting the site map per answered question	0.01	0.02	-0.01	0.03
average time-on-task per answered question	-0.12	0.05	-0.02	*0.29

*p < 0.05

There was a highly weak, negative correlation between self-efficacy of individuals as a trait and mental effort, r = -0.06; a highly weak, positive correlation between self-efficacy of individuals as a trait and perceptions of aesthetics subscale classic aesthetic, r = 0.06; a weak, positive correlation between self-efficacy of individuals as a trait and perceptions of aesthetics subscale expressive aesthetic, r = 0.20; a weak, positive correlation between self-efficacy of individuals as a trait and enjoyment, r = 0.12; a weak, positive correlation between self-efficacy of individuals as a trait and perceived ease of use, r = 0.21; a weak, negative correlation between self-efficacy of individuals as a trait and disorientation, r = -0.10; a highly weak, negative correlation between self-efficacy of individuals as a trait and perceived usefulness, r = -0.02; a weak, positive correlation between self-efficacy of individuals as a trait and behavioural intention, r = 0.19. None of these correlations was statistically significant at 0.05 significance level.

There was a highly weak, negative correlation between self-efficacy of individuals as a trait and time-on-practice tasks, r = -0.09; a highly weak, negative correlation between self-efficacy of individuals as a trait and number of main tasks completed, r = -0.02; a highly weak, negative correlation between self-efficacy of individuals as a trait and time-on-main tasks, r = -0.09; a highly weak, negative correlation between self-efficacy of individuals as a trait and percentage of correct main answers, r = -0.01; a weak, positive correlation between self-efficacy of individuals as a trait and average pages loaded for correct answers, r = 0.10; a highly weak, positive correlation between self-efficacy of individuals as a trait and average time-on-task for correctly answered main questions, r = 0.07; a highly weak, negative correlation between self-efficacy of individuals as a trait and average number of pages loaded for incorrectly answered main questions, r = -0.03; an extremely weak, positive correlation between self-efficacy of individuals as a trait and average time-on-task for incorrectly answered main questions, r = 0.001; a highly weak, negative correlation between self-efficacy of individuals as a trait and average number of correct answers per answered question, r = -0.01; a highly weak, positive correlation between self-efficacy of individuals as a trait and average number of visited pages per answered question, r = 0.04; a highly weak, negative correlation between self-efficacy of individuals as a trait and average number of revisited pages per answered question, r = -0.01; a highly weak, negative correlation between self-efficacy of individuals as a trait and average number of times of visiting the homepage per answered question, r = -0.04; a weak, negative correlation between self-efficacy of individuals as a trait and average number of times of visiting a page with search results per answered question, r = -0.14; a highly weak, positive correlation between self-efficacy of individuals as a trait and average number of times of visiting the site map per answered question, r = 0.02; a highly weak, positive correlation between self-efficacy of individuals as a trait and average time-on-task per answered question, r = 0.05. None of these correlations was statistically significant at 0.05 significance level.

There was a weak, positive correlation between self-efficacy of individuals as a state and mental effort, r = 0.10; a highly weak, positive correlation between self-efficacy of individuals as a state and perceptions of aesthetics subscale classic aesthetic, r = 0.06; a highly weak, negative correlation between self-efficacy of individuals as a state and perceptions of aesthetics subscale expressive aesthetic, r = -0.07; a moderate, positive correlation between self-efficacy of individuals as a state and enjoyment, r = 0.29; a weak, positive correlation between self-efficacy of individuals as a state and perceived ease of use, r = 0.13; a weak, negative correlation between self-efficacy of individuals as a state and disorientation, r = -0.15; a weak, positive correlation between self-efficacy of individuals as a state and perceived usefulness, r = 0.11; a highly weak, positive correlation between self-efficacy of individuals as a state and behavioural intention, r = 0.09. Correlations between self-efficacy of individuals as a state and enjoyment were merely statistically significant at 0.05 significance level, r = 0.29, p = 0.023.

There was a highly weak, negative correlation between self-efficacy of individuals as a state and time-on-practice tasks, r = -0.01; a highly weak, positive correlation between self-efficacy of individuals as a state and number of main tasks completed, r = 0.08; a highly weak, negative correlation between self-efficacy of individuals as a state and time-on-main tasks, r = -0.05; a weak, positive correlation between self-efficacy of individuals as a state and percentage of correct main answers, r = 0.10; a highly weak, negative correlation between self-efficacy of individuals as a state and average pages loaded for correct answers, r = -0.06; a weak, negative correlation between self-efficacy of individuals as a state and average time-on-task for correctly answered main questions, r = -0.15; a highly weak, negative correlation between self-efficacy of individuals as a state and average number of pages loaded for incorrectly answered main questions, r = -0.04; a highly weak, negative correlation between self-efficacy of individuals as a state and average time-on-task for incorrectly answered main questions, r = -0.07; a highly weak, negative correlation between self-efficacy of individuals as a state and average number of correct answers per answered question, r = -0.08; a weak, positive correlation between self-efficacy of individuals as a state and average number of visited pages per answered question, r = 0.10; a weak, positive correlation between self-efficacy of individuals as a state and average number of revisited pages per answered question, r = 0.12; a highly weak, negative correlation between self-efficacy of individuals as a state and average number of times of visiting the homepage per answered question, r = -0.02; a highly weak, negative correlation between self-efficacy of individuals as a state and average number of times of visiting a page with search results per answered question, r = -0.10; a highly weak, negative correlation between self-efficacy of individuals as a state and average number of times of visiting the site map per answered question, r = -0.01; a highly weak, negative correlation between self-efficacy of individuals as a state and average time-on-task per answered question, r = -0.02. None of these correlations was statistically significant at 0.05 significance level.

There was a highly weak, positive correlation between intrinsic motivation of individuals and mental effort, r = 0.06; a weak, positive correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale classic aesthetic, r = 0.12; a weak, positive correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale expressive aesthetic, r = 0.15; a highly weak, positive correlation between intrinsic motivation of individuals and enjoyment, r = 0.02; a highly weak, positive correlation between intrinsic motivation of individuals and perceived ease of use, r = 0.02; a weak, negative correlation between intrinsic motivation of individuals and disorientation, r = -0.13; a weak, negative correlation between intrinsic motivation of individuals and perceived usefulness, r = -0.12; a highly weak, positive correlation between intrinsic motivation of individuals and behavioural intention, r = 0.05. None of these correlations was statistically significant at 0.05 significance level.

There was a highly weak, negative correlation between intrinsic motivation of individuals and time-on-practice tasks, r = -0.004; a highly weak, negative correlation between intrinsic motivation of individuals and number of main tasks completed, r = -0.09; a highly weak, positive correlation between intrinsic motivation of individuals and time-on-main tasks, r = 0.09; a weak, negative correlation between intrinsic motivation of individuals and percentage of correct main answers, r = -0.15; a weak, positive correlation between intrinsic motivation of individuals and average pages loaded for correct answers, r = 0.14; a weak, positive correlation between intrinsic motivation of individuals and average time-on-task for correctly answered main questions, r = 0.22; a highly weak, positive correlation between intrinsic motivation of individuals and average number of pages loaded for incorrectly answered main questions, r = 0.05; a weak, positive correlation between intrinsic motivation of individuals and average time-on-task for incorrectly answered main questions, r = 0.22; a weak, negative correlation between intrinsic motivation of individuals and average number of correct answers per answered question, r = -0.10; a highly weak, positive correlation between intrinsic motivation of individuals and average number of visited pages per answered question, r = 0.05; a highly weak, positive correlation between intrinsic motivation of individuals and average number of revisited pages per answered question, r = 0.04; a weak, positive correlation between intrinsic motivation of individuals and average number of times of visiting the homepage per answered question, r = 0.15; a highly weak, positive correlation between intrinsic motivation of individuals and average number of times of visiting a page with search results per answered question, r = 0.05; a highly weak, positive correlation between intrinsic motivation of individuals and average number of times of visiting the site map per answered question, r = 0.03; a moderate, positive correlation between intrinsic motivation of individuals and average time-on-task per answered question, r = 0.29. Correlations between intrinsic motivation of individuals and average time-on-task per answered question were statistically significant at 0.05 significance level, r = 0.29, p = 0.025.

3.6.2 Relations of individual-difference variables and outcome measures in repeated non-think-aloud data

The relationship between individual-difference variables and outcome measures was also explored in re-analysed data by means of using Pearson product-moment correlation coefficient. The strength and direction of these variables were almost same as the original data analyses. Although the differences between original data analyses and re-data analyses for correlation did not yield any statistically significant results, those were presented below:

There was a highly weak, positive correlation between need for cognition of individuals and enjoyment, $r = 0.22$. There was a weak, positive correlation between self-efficacy of individuals as a trait and enjoyment, $r = 0.12$. There was a highly weak, positive correlation between intrinsic motivation of individuals and mental effort, $r = 0.08$; a highly weak, positive correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale classic aesthetic, $r = 0.08$; a weak, positive correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale expressive aesthetic, $r = 0.12$; a highly weak, positive correlation between intrinsic motivation of individuals and enjoyment, $r = 0.02$; a highly weak, negative correlation between intrinsic motivation of individuals and perceived ease of use, $r = -0.02$; a highly weak, negative correlation between intrinsic motivation of individuals and disorientation, $r = -0.10$; a weak, negative correlation between intrinsic motivation of individuals and perceived usefulness, $r = -0.15$; a highly weak, negative correlation between intrinsic motivation of individuals and behavioural intention, $r = 0.005$.

3.6.3 Relations of individual-difference variables and outcome measures in think-aloud data

The relationship between individual-difference variables and outcome measures for think-aloud data was explored by means of using Pearson product-moment correlation coefficient. The strength and direction of these variables were as follows:

There was a strong, negative correlation between need for cognition of individuals and mental effort, $r = -0.51$; a weak, negative correlation between need for cognition of individuals and perceptions of aesthetics subscale of classic aesthetic, $r = -0.11$; a moderate, positive correlation between need for cognition of individuals and perceptions of aesthetics subscale of expressive aesthetic, $r = 0.43$; a highly strong, positive correlation between need for cognition of individuals and enjoyment, $r = 0.77$;

a weak, negative correlation between need for cognition of individuals and perceived ease of use, r = -0.19; a weak, negative correlation between need for cognition of individuals and disorientation, r = -0.24; a highly weak, negative correlation between need for cognition of individuals and perceived usefulness, r = -0.15; a highly weak, negative correlation between need for cognition of individuals and behavioural intention, r = -0.01. Correlations between need for cognition of individuals and enjoyment were statistically significant at 0.05 significance level, r = 0.77, p = 0.015.

There was a weak, negative correlation between need for cognition of individuals and time-on-practice tasks, r = -0.22; a highly weak, positive correlation between need for cognition of individuals and number of main tasks completed, r = 0.07; a weak, negative correlation between need for cognition of individuals and time-on-main tasks, r = -0.23; a weak, positive correlation between need for cognition of individuals and percentage of correct main answers, r = 0.13; a highly weak, positive correlation between need for cognition of individuals and average pages loaded for correct answers, r = 0.05; a highly weak, positive correlation between need for cognition of individuals and average time-on-task for correctly answered main questions, r = 0.09; a moderate, negative correlation between need for cognition of individuals and average number of pages loaded for incorrectly answered main questions, r = -0.38; a moderate, negative correlation between need for cognition of individuals and average time-on-task for incorrectly answered main questions, r = -0.38; a moderate, positive correlation between need for cognition of individuals and average number of correct answers per answered question, r = 0.38; a highly weak, negative correlation between need for cognition of individuals and average number of visited pages per answered question, r = -0.10; a weak, positive correlation between need for cognition of individuals and average number of revisited pages per answered question, r = 0.10; a moderate, negative correlation between need for cognition of individuals and average number of times of visiting the homepage per answered question, r = -0.34; a highly weak, positive correlation between need for cognition of individuals and average number of times of visiting a page with search results per answered question, r = 0.08; a weak, negative correlation between need for cognition of individuals and average number of times of visiting the site map per answered question, r = -0.21; a moderate, negative correlation between need for cognition of individuals and average time-on-task per answered question, r = -0.37. None of these correlations was statistically significant at 0.05 significance level.

There was a weak, positive correlation between self-efficacy of individuals as a trait and mental effort, r = 0.22; a weak, positive correlation between self-efficacy of individuals as a trait and perceptions of aesthetics subscale classic aesthetic, r = 0.26; a moderate, negative correlation between self-efficacy of individuals as a trait and perceptions of aesthetics subscale expressive aesthetic, r = -0.31; a highly weak, negative correlation between self-efficacy of individuals as a trait and enjoyment, r = -0.04; a strong, negative correlation between self-efficacy of individuals as a trait and perceived ease of use, r = -0.51; a moderate, positive correlation between self-efficacy of individuals as a trait and disorientation, r = 0.47; a moderate, negative correlation between self-efficacy of individuals as a trait and perceived usefulness, r = -0.31; a highly weak, positive correlation between self-efficacy of individuals as a trait and behavioural intention, r = 0.004. None of these correlations was statistically significant at 0.05 significance level.

There was a highly weak, positive correlation between self-efficacy of individuals as a trait and time-on-practice tasks, r = 0.08; a strong, negative correlation between self-efficacy of individuals as a trait and number of main tasks completed, r = -0.48; a highly weak, positive correlation between self-efficacy of individuals as a trait and time-on-main tasks, r = 0.03; a strong, negative correlation between self-efficacy of individuals as a trait and percentage of correct main answers, r = -0.47; a weak, negative correlation between self-efficacy of individuals as a trait and average pages loaded for correct answers, r = -0.17; a highly weak, negative correlation between self-efficacy of individuals as a trait and average time-on-task for correctly answered main questions, r = -0.07; a weak, positive correlation between self-efficacy of individuals as a trait and average number of pages loaded for incorrectly answered main questions, r = 0.10; a highly weak, positive correlation between self-efficacy of individuals as a trait and average time-on-task for incorrectly answered main questions, r = 0.02; a weak, negative correlation between self-efficacy of individuals as a trait and average number of correct answers per answered question, r = -0.10; a weak, negative correlation between self-efficacy of individuals as a trait and average number of visited pages per answered question, r = -0.15; a highly weak, negative correlation between self-efficacy of individuals as a trait and average number of revisited pages per answered question, r = -0.06; a weak, positive correlation between self-efficacy of individuals as a trait and average number of times of visiting the homepage per answered question, r = 0.24; a weak, positive correlation between self-efficacy of individuals as a trait and average number of times of visiting a page with search results per answered question, r = 0.14; a strong, positive correlation between self-efficacy of individuals as a trait and average number of times of visiting the site map per answered question, r = 0.49; a highly weak, negative correlation between self-efficacy of individuals as a trait and average time-on-task per answered question, r = -0.04. None of these correlations was statistically significant at 0.05 significance level.

There was a moderate, negative correlation between self-efficacy of individuals as a state and mental effort, r = -0.44; a highly weak, negative correlation between self-efficacy of individuals as a state and perceptions of aesthetics subscale classic aesthetic, r = -0.01; a weak, negative correlation between self-efficacy of individuals as a state and perceptions of aesthetics subscale expressive aesthetic, r = -0.23; a moderate, positive correlation between self-efficacy of individuals as a state and enjoyment, r = 0.34; a highly weak, negative correlation between self-efficacy of individuals as a state and perceived ease of use, r = -0.03; a weak, negative correlation between self-efficacy of individuals as a state and disorientation, r = -0.10;

a moderate, negative correlation between self-efficacy of individuals as a state and perceived usefulness, r = -0.33; a highly weak, negative correlation between self-efficacy of individuals as a state and behavioural intention, r = -0.04. None of these correlations was statistically significant at 0.05 significance level.

There was a moderate, positive correlation between self-efficacy of individuals as a state and time-on-practice tasks, r = 0.29; a weak, positive correlation between self-efficacy of individuals as a state and number of main tasks completed, r = 0.17; a weak, positive correlation between self-efficacy of individuals as a state and time-on-main tasks, r = 0.27; a weak, positive correlation between self-efficacy of individuals as a state and percentage of correct main answers, r = 0.21; a moderate, negative correlation between self-efficacy of individuals as a state and average pages loaded for correct answers, r = -0.47; a moderate, negative correlation between self-efficacy of individuals as a state and average time-on-task for correctly answered main questions, r = -0.34; a moderate, negative correlation between self-efficacy of individuals as a state and average number of pages loaded for incorrectly answered main questions, r = -0.30; a moderate, negative correlation between self-efficacy of individuals as a state and average time-on-task for incorrectly answered main questions, r = -0.35; a moderate, positive correlation between self-efficacy of individuals as a state and average number of correct answers per answered question, r = 0.30; a strong, negative correlation between self-efficacy of individuals as a state and average number of visited pages per answered question, r = -0.66; a moderate, negative correlation between self-efficacy of individuals as a state and average number of revisited pages per answered question, r = -0.47; a moderate, negative correlation between self-efficacy of individuals as a state and average number of times of visiting the homepage per answered question, r = -0.40; a weak, negative correlation between self-efficacy of individuals as a state and average number of times of visiting a page with search results per answered question, r = -0.12; a highly weak, positive correlation between self-efficacy of individuals as a state and average number of times of visiting the site map per answered question, r = 0.08; a highly strong, negative correlation between self-efficacy of individuals as a state and average time-on-task per answered question, r = -0.80. Correlations between self-efficacy of individuals as a state and average time-on-task per answered question were statistically significant at 0.05 significance level, r = -0.80, p = 0.01.

There was a weak, positive correlation between intrinsic motivation of individuals and mental effort, r = 0.20; a moderate, positive correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale classic aesthetic, r = 0.41; a moderate, negative correlation between intrinsic motivation of individuals and perceptions of aesthetics subscale expressive aesthetic, r = -0.37; a moderate, negative correlation between intrinsic motivation of individuals and enjoyment, r = -0.40; a strong, negative correlation between intrinsic motivation of individuals and perceived ease of use, r = -0.61; a highly strong, positive correlation between intrinsic motivation of individuals and disorientation, r = 0.86, p = 0.003; a moderate, negative correlation between intrinsic motivation of individuals and perceived usefulness, r = -0.30; a moderate, negative correlation between intrinsic motivation of individuals and behavioural intention, r = -0.30. Correlations between intrinsic motivation of individuals and disorientation were statistically significant at 0.05 significance level, r = 0.86, p = 0.003.

There was a highly weak, negative correlation between intrinsic motivation of individuals and time-on-practice tasks, r = -0.07; a strong, negative correlation between intrinsic motivation of individuals and number of main tasks completed, r = -0.60; a weak, negative correlation between intrinsic motivation of individuals and time-on-main tasks, r = -0.12; a strong, negative correlation between intrinsic motivation of individuals and percentage of correct main answers, r = -0.51; an almost significantly strong, positive correlation between intrinsic motivation of individuals and average pages loaded for correct answers, r = 0.66; a significantly strong, positive correlation between intrinsic motivation of individuals and average time-on-task for correctly answered main questions, r = 0.72; a weak, negative correlation between intrinsic motivation of individuals and average number of pages loaded for incorrectly answered main questions, r = -0.19; a weak, negative correlation between intrinsic motivation of individuals and average time-on-task for incorrectly answered main questions, r = -0.24; a weak, positive correlation between intrinsic motivation of individuals and average number of correct answers per answered question, r = 0.19; a statistically strong, positive correlation between intrinsic motivation of individuals and average number of visited pages per answered question, r = 0.67; a statistically strong, positive correlation between intrinsic motivation of individuals and average number of revisited pages per answered question, r = 0.67; a strong, positive correlation between intrinsic motivation of individuals and average number of times of visiting the homepage per answered question, r = 0.50; a strong, positive correlation between intrinsic motivation of individuals and average number of times of visiting a page with search results per answered question, r = 0.60; a weak, positive correlation between intrinsic motivation of individuals and average number of times of visiting the site map per answered question, r = 0.14; a moderate, positive correlation between intrinsic motivation of individuals and average time-on-task per answered question, r = 0.43. Correlations between intrinsic motivation of individuals and average pages loaded for correct answers, r = 0.66, p = 0.052; average time-on-task for correctly answered main questions, r = 0.72, p = 0.03; average number of visited pages per answered question, r = 0.67, p = 0.05; and average number of revisited pages per answered question, r = 0.67, p = 0.05 were statistically significant at 0.05 significance level.

3.7 Conclusion

In this chapter, method of the experiment was elaborated. As explained extensively before, method followed a two-step approach, consisting of an experimental design and a qualitative method, in order to test the hypotheses proposed previously.

In overall, reliability of scales from non-think-aloud data demonstrated a sufficient amount of Cronbach's alpha coefficient. Therefore, it can be said that reliability of scales was confirmed by the current study. However, two items (item 10 in the subscale of 'identified motivation' of SIMS and item 8 in the subscale of 'transformation of time' of Flow State Scale) decreased the reliability of two questionnaires. Lower than ten items in a subscale bring about a lower Cronbach's alpha coefficient. Despite all subscales of this study consisted of about four items, the data analyses were repeated by removing these two items in order to explore whether or not to acquire more significant results. On the other hand, reliability of scales from think-aloud data mostly could not demonstrate a sufficient amount of Cronbach's alpha score. However, none of the items was removed and analyses were not repeated because of the insufficient number of participants to analyse. Therefore, the current study focused on the qualitative results of think-aloud data instead of quantitative results of this data.

4. Results

In this chapter, the results of the data analyses will be presented, including both quantitative and qualitative results.

4.1 Quantitative results

4.1.1 Main effects of individual difference variables on outcomes

Standard multiple regression was carried out to test whether individual-difference variables had significant main effects on outcome measures. A significance level of 0.05 was chosen for statistical testing.

Results of non-think-aloud data analyses

According to the results of multiple regression analysis, individual-difference variables did not explain a significant amount of variability on *mental effort*, $R^2 = 0.04$, $F(4, 56) = 0.66$, $p = 0.624$. In other words, there was no significant effect of individual-difference variables on mental effort participants experienced. Need for cognition made the strongest but not significant unique contribution to explaining mental effort, $\beta = -0.15$, $t(56) = -1.13$, $p = 0.26$. There was no statistically significant association between individual-difference variables and *mental effort*.

Individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics*, $R^2 = 0.04$, $F(4, 56) = 0.55$, $p = 0.70$. Need for cognition made the strongest but not significant unique contribution to explaining *perceptions of classic aesthetics*, $\beta = -0.14$, $t(56) = -1.01$, $p = 0.32$. Individual-difference variables did not explain a significant amount of variability on *perceptions of expressive aesthetics*, $R^2 = 0.14$, $F(4, 56) = 2.25$, $p = 0.07$. Need for cognition made the strongest and unique significant contribution to explaining *perceptions of expressive aesthetics*, $\beta = -0.28$, $t(56) = -2.17$, $p = 0.034$. In contradiction with Hypothesis 1, people with a high need for cognition experienced lower *perceptions of expressive aesthetics* than those with low need for cognition.

Individual-difference variables explained a significant amount of variability on *enjoyment*, $R^2 = 16.5$, $F(4, 56) = 2.76$, $p = 0.04$. Therefore overall individual-difference variables had a significant effect on *enjoyment* participants experienced.

Self-efficacy as a state made the strongest and unique significant contribution to the prediction of *enjoyment*, $\beta = 0.34$, t (56) = 2.65, $p = 0.01$. Need for cognition also made a statistically unique contribution with a lower beta, $\beta = 0.28$, t (56) = 2.26, $p = 0.03$. In other words, participants with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state. Participants with a high need for cognition experienced higher perceived enjoyment than those with low need for cognition, too. These results were consistent with Hypothesis 2.

Individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, $R^2 = 0.05$, F (4, 56) = 0.75, $p = 0.56$. Therefore, there was no significant effect of individual-difference variables on *perceived ease of use*. Self-efficacy as a trait made the strongest but not significant unique contribution to explaining *perceived ease of use*, $\beta = 0.19$ t (56) = 1.40, $p = 0.17$.

Individual-difference variables did not explain a significant amount of variability on *disorientation*, $R^2 = 0.05$, F (4, 56) = 0.81, $p = 0.522$. Thus, there was no significant effect of individual-difference variables on *disorientation* participants experienced. Self-efficacy as a state made the strongest but not significant unique contribution to explaining disorientation, $\beta = -0.17$, t (56) = -1.22, $p = 0.23$.

Individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, $R^2 = 0.03$, F (4, 56) = 0.47, $p = 0.75$. In other words, there was no significant effect of individual-difference variables on *perceived usefulness*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *perceived usefulness*, $\beta = -0.122$, t (56) = -0.91, $p = 0.36$.

Individual-difference variables did not explain a significant amount of variability on *behavioural intention*, $R^2 = 0.05$, F (4, 56) = 0.71, $p = 0.60$. In other words, there was no significant effect of individual-difference variables on *behavioural intention* participants experienced. Self-efficacy as a trait made the strongest but not significant unique contribution to explaining *behavioural intention*, $\beta = 0.183$, t (56) = 1.35, $p = 0.18$.

Individual-difference variables did not explain a significant amount of variability on *number of main tasks completed*, $R^2 = 0.02$, F (4, 56) = 0.26, $p = 0.90$. In other words, there was no significant effect of individual-difference variables on *number of main tasks completed*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *number of main tasks completed*, $\beta = -0.09$, t (56) = -0.68, $p = 0.50$.

Individual-difference variables did not explain a significant amount of variability on *time-on-main tasks*, $R^2 = 0.03$, F (4, 56) = 0.42, $p = 0.80$. In other words, there was no significant effect of individual-difference variables on *time-on-main tasks*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *time-on-main tasks*, $\beta = 0.11$, t (56) = 0.82, $p = 0.42$.

Individual-difference variables did not explain a significant amount of variability on *number of correct main answers*, $R^2 = 0.03$, $F(4, 56) = 0.46$, $p = 0.76$. In other words, there was no significant effect of individual-difference variables on *number of correct main answers*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *number of correct main answers*, $\beta = -0.14$, $t(56) = -1.06$, $p = 0.30$.

Individual-difference variables did not explain a significant amount of variability on *average pages loaded for correct answers*, $R^2 = 0.04$, $F(4, 56) = 0.62$, $p = 0.65$. In other words, there was no significant effect of individual-difference variables on *average pages loaded for correct answers*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *average pages loaded for correct answers*, $\beta = 0.12$, $t(56) = 0.89$, $p = 0.38$.

Individual-difference variables did not explain a significant amount of variability on *average time-on-task for correctly answered main questions*, $R^2 = 0.10$, $F(4, 56) = 1.48$, $p = 0.22$. In other words, there was no significant effect of individual-difference variables on *average time-on-task for correctly answered main questions*. Self-efficacy as a state made the strongest but not significant unique contribution to explaining *average time-on-task for correctly answered main questions*, $\beta = -0.20$, $t(56) = -1.43$, $p = 0.18$.

Individual-difference variables did not explain a significant amount of variability on *average number of pages loaded for incorrectly answered main questions*, $R^2 = 0.02$, $F(4, 56) = 0.28$, $p = 0.90$. In other words, there was no significant effect of individual-difference variables on *average number of pages loaded for incorrectly answered main questions*. Need for cognition made the strongest but not significant unique contribution to explaining *average number of pages loaded for incorrectly answered main questions*, $\beta = -0.13$, $t(56) = -0.92$, $p = 0.36$.

Individual-difference variables did not explain a significant amount of variability on *average time-on-task for incorrectly answered main questions*, $R^2 = 0.06$, $F(4, 56) = 0.93$, $p = 0.46$. In other words, there was no significant effect of individual-difference variables on *average time-on-task for incorrectly answered main questions*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *average time-on-task for incorrectly answered main questions*, $\beta = 0.20$, $t(56) = 1.50$, $p = 0.14$.

Individual-difference variables did not explain a significant amount of variability on *average number of correct answers per answered question*, $R^2 = 0.02$, $F(4, 56) = 0.31$, $p = 0.87$. In other words, there was no significant effect of individual-difference variables on *average number of correct answers per answered question*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *average number of correct answers per answered question*, $\beta = -0.10$, $t(56) = -0.75$, $p = 0.46$.

Individual-difference variables did not explain a significant amount of variability on *average number of visited pages per answered question*, $R^2 = 0.02$, $F(4, 56) = 0.26$, $p = 0.91$. In other words, there was no significant effect of individual-difference variables on *average number of visited pages per answered question*. Self-efficacy as a state made the strongest but not significant unique contribution to explaining *average number of visited pages per answered question*, $\beta = 0.09$, $t(56) = 0.62$, $p = 0.54$.

Individual-difference variables did not explain a significant amount of variability on *average number of revisited pages per answered question*, $R^2 = 0.02$, $F(4, 56) = 0.24$, $p = 0.91$. In other words, there was no significant effect of individual-difference variables on *average number of revisited pages per answered question*. Self-efficacy as a state made the strongest but not significant unique contribution to explaining *average number of revisited pages per answered question*, $\beta = 0.13$, $t(56) = 0.92$, $p = 0.36$.

Individual-difference variables did not explain a significant amount of variability on *average number of times of visiting the homepage per answered question*, $R^2 = 0.04$, $F(4, 56) = 0.62$, $p = 0.65$. In other words, there was no significant effect of individual-difference variables on *average number of times of visiting the homepage per answered question*. Need for cognition made the strongest but not significant unique contribution to explaining *average number of times of visiting the homepage per answered question*, $\beta = -0.14$, $t(56) = -1.06$, $p = 0.30$.

Individual-difference variables did not explain a significant amount of variability on *average number of times of visiting a page with search results per answered question*, $R^2 = 0.04$, $F(4, 56) = 0.58$, $p = 0.68$. In other words, there was no significant effect of individual-difference variables on *average number of times of visiting a page with search results per answered question*. Self-efficacy as a trait made the strongest but not significant unique contribution to explaining *average number of times of visiting a page with search results per answered question*, $\beta = -0.14$, $t(56) = -1.06$, $p = 0.30$.

Individual-difference variables did not explain a significant amount of variability on *average number of times of visiting the site map per answered question*, $R^2 = 0.03$, $F(4, 56) = 0.01$, $p = 1.00$. In other words, there was no significant effect of individual-difference variables on *average number of times of visiting the site map per answered question*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *average number of times of visiting the site map per answered question*, $\beta = 0.03$, $t(56) = 0.18$, $p = 0.86$.

Individual-difference variables did not explain a significant amount of variability on *average time-on-task per answered question*, $R^2 = 0.10$, $F(4, 56) = 1.44$, $p = 0.23$. In other words, there was no significant effect of individual-difference variables on *average time-on-task per answered question*. Intrinsic motivation made the strongest and unique significant contribution to explaining *average time-on-task per answered question*, β = 0.27, t (56) = 2.08, p = 0.04. In consistent with Hypothesis 3, people with a high need for cognition experienced higher *average time-on-task per answered question* than those with low need for cognition.

Results of repeated non-think-aloud data analyses

Repeated non-think-aloud data analyses showed that individual-difference variables did not explain a significant amount of variability on *mental effort*, R^2 = 0.05, F (4, 56) = 0.689, p = 0.603. Therefore there was no significant effect of individual-difference variables on *mental effort* participants experienced. Need for cognition made the strongest but not significant unique contribution to explaining mental effort, β = -0.15, t (56) = -1.13, p = 0.26.

Individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics*, R^2 = 0.03, F (4, 56) = 0.47, p = 0.76. Need for cognition made the strongest but not significant unique contribution to explaining *perceptions of classic aesthetics*, β = -0.14, t (56) = -1.05, p = 0.30. Individual-difference variables did not explain a significant amount of variability *on perceptions of expressive aesthetics*, R^2 = 0.14, F (4, 56) = 2.20, p = 0.08. Need for cognition made the strongest and unique significant contribution to explaining *perceptions of expressive aesthetics*, β = -0.28, t (56) = -2.19, p = 0.032. In contradiction with Hypothesis 1, people with a high need for cognition experienced lower perceptions of expressive aesthetic than those with low need for cognition.

Individual-difference variables explained a significant amount of variability on *enjoyment*, R^2 = 0.17, F (4, 56) = 2.784, p = 0.035. Thus, overall individual-difference variables had a significant effect on *enjoyment* participants experienced. Self-efficacy as a state made the strongest and unique significant contribution to the prediction of *enjoyment*, β = 0.34, t (56) = 2.64, p = 0.01. Need for cognition also made a statistically unique contribution with a lower beta, β = 0.287, t (56) = 2.28, p = 0.026. In other words, participants with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state. Participants with a high need for cognition experienced higher perceived enjoyment than those with low need for cognition, too. These results were consistent with Hypothesis 2.

Individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, R^2 = 0.05, F (4, 56) = 0.75, p = 0.56. In other words, there was no significant effect of individual-difference variables on *perceived ease of use*. Self-efficacy as a trait made the strongest but not significant unique contribution to explaining *perceived ease of use*, β = 0.19, t (56) = 1.42, p = 0.16.

Individual-difference variables did not explain a significant amount of variability on *disorientation*, $R^2 = 0.05$, F (4, 56) = 0.72, $p = 0.58$. Thus, there was no significant effect of individual-difference variables on *disorientation* participants experienced. Self-efficacy as a state made the strongest but not significant unique contribution to explaining *disorientation*, $\beta = -0.17$, t (56) = -1.21, $p = 0.223$.

Individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, $R^2 = 0.04$, F (4, 56) = 0.57, $p = 0.68$. Therefore, there was no significant effect of individual-difference variables on *perceived usefulness*. Intrinsic motivation made the strongest but not significant unique contribution to explaining *perceived usefulness*, $\beta = -0.15$, t (56) = -1.10, $p = 0.27$.

Individual-difference variables did not explain a significant amount of variability on *behavioural intention*, $R^2 = 0.05$, F (4, 56) = 0.70, $p = 0.60$. Therefore, there was no significant effect of individual-difference variables on *behavioural intention* participants experienced. Self-efficacy as a trait made the strongest but not significant unique contribution to explaining *behavioural intention*, $\beta = 0.19$, t (56) = 1.39, p = 0.17.

Summary of the results

Results from non-think-aloud data indicated some statistically significant results. Those were as follows:

People with a high need for cognition experienced lower *perceptions of expressive aesthetics* than those with low need for cognition, which was contradictory with Hypothesis 1, while people with high need for cognition experienced higher *perceived enjoyment* than those with low need for cognition, which was consistent with Hypothesis 1. Participants with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state, which was consistent with Hypothesis 2. However, there was only one statistically significant association between intrinsic motivation of participants and outcome measures: people with a high need for cognition experienced higher *average time-on-task per answered question* than those with low need for cognition. Therefore, Hypothesis 3 was partly supported in the study.

Results from repeated non-think-aloud data indicated the same statistically significant results as results of non-think-aloud data did. Those were as follows:

Results from repeated non-think-aloud data indicated that, people with a high need for cognition experienced lower *perceptions of expressive aesthetic* than those with low need for cognition, which was contradictory with Hypothesis 1, while participants with a high need for cognition experienced higher *perceived enjoyment* than those with low need for cognition, which was consistent with Hypothesis 1. Participants with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state. This was consistent with Hypothesis 2. However, there was no statistically significant association between intrinsic motivation of participants and outcome measures. Therefore, Hypothesis 3 was not supported in the study.

4.1.2 Main effects of task complexity on outcome measures

In this part data from participants with high artefact complexity was only used in order to investigate the main effects of task complexity on outcome measures.

Results of non-think-aloud data analyses

There was a significant difference in *mental effort* scores between simple tasks (M = 21.42, SD = 17.91) and complex tasks (M = 62.19, SD = 36.57); t (27) = -4.53, p < 0.0001 (two-tailed) (Table 4). In consistent with Hypothesis 4, high task complexity led to more mental effort of participants.

There was no significant difference in p*erceptions of aesthetics subscale classic aesthetics* scores between simple tasks (M = 4.33, SD = 0.99) and complex tasks (M = 4.12, SD = 1.35); t (41) = 0.59, p = 0.56.

There was no significant difference in *perceptions of aesthetics subscale expressive aesthetics* scores between simple tasks (M = 3.23, SD = 0.75) and complex tasks (M = 3.30, SD = 0.96); t (41) = -0.28, p = 0.78.

There was no significant difference in *enjoyment* scores between simple tasks (M = 4.45, SD = 0.56) and complex tasks (M = 4.22, SD = 1.08); t (28) = 0.83, p = 0.414.

There was no significant difference in *perceived ease of use* scores between simple tasks (M = 4.86, SD = 1.15) and complex tasks (M = 4.02, SD = 1.83); t (31) = 1.77, p = 0.087.

There was no significant difference in *disorientation* scores between simple tasks (M = 2.90, SD = 1.01) and complex tasks (M = 3.39, SD = 1.75); t (30) = -1.09, p = 0.284.

There was a significant difference in *perceived usefulness* scores between simple tasks (M = 4.98, SD = 0.69) and complex tasks (M = 4.20, SD = 1.30); t (28) = 2.40, p

= 0.023 (two-tailed) (Table 4). In consistent with Hypothesis 4, high task complexity led to lower perceived usefulness of participants.

There was no significant difference in *behavioural intention* scores between simple tasks (M = 4.78, SD = 1.29) and complex tasks (M = 4.22, SD = 1.77); *t* (34) = 1.18, *p* = 0.245.

Table 4 Main effects of task complexity on user experience outcomes

	Low Task Complexity	High Task Complexity
Mental Effort	21.42 (17.91)	62.19 (36.57)
Perceptions of Classic Aesthetics	4.33 (0.99)	4.12 (1.35)
Perceptions of Expressive Aesthetics	3.23 (0.75)	3.30 (0.96)
Enjoyment	4.45 (0.56)	4.22 (1.08)
Perceived Ease of Use	4.86 (1.15)	4.02 (1.83)
Disorientation	2.90 (1.01)	3.39 (1.75)
Perceived Usefulness	4.98 (0.69)	4.20 (1.30)
Bahavioural Intention	4.78 (1.29)	4.22 (1.77)

Note. Mean values are presented with standard deviations in brackets.

There was a significant difference in *number of correct main answers* scores between simple tasks (M = 1.70, SD = 1.66) and complex tasks (M = 0.75, SD = 0.79); *t* (41) = 2.32, *p* = 0.025 (two-tailed). Therefore, high task complexity led to less number of correct main answers than low task complexity.

There was a significant difference in *average time-on-task for incorrectly answered main questions* scores between simple tasks (M = 6326.44, SD = 14276.93) and complex tasks (M = 17169.25, SD = 14208.99); *t* (41) = -2.49, *p* = 0.02 (two-tailed). Therefore, high task complexity led to more average time-on-task for incorrectly answered main questions than low task complexity.

There was a significant difference in *average number of correct answers per answered question* scores between simple tasks (M = 0.88, SD = 0.27) and complex tasks (M = 0.45, SD = 0.46); *t* (30) = 3.63, *p* = 0.001 (two-tailed). Therefore, high task complexity led to less average number of correct answers per answered question than low task complexity.

There was a significant difference in *average number of visited pages per answered question* scores between simple tasks (M = 5.24, SD = 3.51) and complex tasks (M =

7.68, SD = 3.81); t (41) = -2.19, p = 0.034 (two-tailed). Therefore, high task complexity led to more average number of visited pages per answered question than low task complexity.

There was a significant difference in *average number of times of visiting the homepage per answered question* scores between simple tasks (M = 1.07, SD = 0.21) and complex tasks (M = 1.33, SD = 0.50); t (25) = -2.17, p = 0.04 (two-tailed). Therefore, high task complexity led to more average number of times of visiting the homepage per answered question than low task complexity.

There was a significant difference in *average number of times of visiting a page with search results per answered question* scores between simple tasks (M = 0.33, SD = 0.45) and complex tasks (M = 1.45, SD = 0.87); t (27) = -5.24, p < 0.0001 (two-tailed). Therefore, high task complexity led to more average number of times of visiting a page with search results per answered question than low task complexity.

There was a significant difference in *average time-on-task per answered question* scores between simple tasks (M = 70753.94, SD = 60747.12) and complex tasks (M = 168540.39, SD = 74953.87); t (41) = -4.72, p < 0.0001 (two-tailed). Therefore, high task complexity led to more average time-on-task per answered question than low task complexity.

Results of repeated non-think-aloud data analyses

As investigated for non-think-aloud data, main effects of task complexity on outcome measures were also analysed for repeated non-think-aloud data. All the results were the same as the results of non-think-aloud data except *enjoyment* scores: There was no significant difference in enjoyment scores between simple tasks (M = 4.45, SD = 0.58) and complex tasks (M = 4.24, SD = 1.11); t (28) = 0.76, p = 0.454.

Summary of the results

Results from non-think-aloud data indicated that high task complexity led to more mental effort and lower perceived usefulness of participants. These results were consistent with Hypothesis 4. Besides, high task complexity led to less number of correct main answers, more average time-on-task for incorrectly answered main questions, less average number of correct answers per answered question, more average number of visited pages per answered question, more average number of times of visiting the homepage per answered question, more average number of times of visiting a page with search results per answered question and more average time-on-task per answered question. Results from repeated non-think-aloud data also indicated the same statistically significant results. However, there were no other statistically significant results for the main effects of task complexity on outcome measures.

4.1.3 Main effects of artefact complexity on outcome measures

In this part data from participants with low task complexity was only used in order to investigate the main effects of artefact complexity on outcome measures.

Results of non-think-aloud data analyses

There was a significant difference in *mental effort* scores between simple Web site (M = 76.42, SD = 31.08) and complex Web site (M = 21.42, SD = 17.91); t (26) = 6.689, p < 0.0001 (two-tailed) (Table 5). In contradiction with Hypothesis 5, high artefact complexity led to less mental effort of participants.

There was no significant difference in *perceptions of aesthetics subscale classic aesthetics* scores between simple Web site (M = 4.59, SD = 1.03) and complex Web site (M = 4.33, SD = 0.99); t (39) = 0.81, p = 0.421.

There was no significant difference in *perceptions of aesthetics subscale expressive aesthetics* scores between simple Web site (M = 3.66, SD = 1.04) and complex Web site (M = 3.23, SD = 0.75); t (39) = 1.53, p = 0.134.

There was no significant difference in *enjoyment* scores between simple Web site (M = 4.29, SD = 0.77) and complex Web site (M = 4.44, SD = 0.56); t (39) = -0.70, p = 0.49.

There was no significant difference in perceived ease of use scores between simple Web site (M = 4.50, SD = 1.25) and complex Web site (M = 4.86, SD = 1.15); t (39) = -0.94, p = 0.35.

There was no significant difference in *disorientation* scores between simple Web site (M = 3.48, SD = 1.20) and complex Web site (M = 2.90, SD = 1.01); t (39) = 1.66, p = 0.104.

There was no significant difference in *perceived usefulness* scores between simple Web site (M = 4.82, SD = 1.22) and complex Web site (M = 4.98, SD = 0.69); t (26) = -0.49, p = 0.625.

There was no significant difference in *behavioural intention* scores between simple Web site (M = 5.02, SD = 1.34) and complex Web site (M = 4.78, SD = 1.29); t (39) = 0.57, p = 0.57.

Table 5 Main effects of artefact complexity on user experience outcomes

	Low Artefact Complexity	High Artefact Complexity
Mental Effort	76.42 (31.08)	21.42 (17.91)
Perceptions of Classic Aesthetics	4.59 (1.03)	4.33 (0.99)
Perceptions of Expressive Aesthetics	3.66 (1.04)	3.23 (0.75)
Enjoyment	4.29 (0.77)	4.44 (0.56)
Perceived Ease of Use	4.50 (1.25)	4.86 (1.15)
Disorientation	3.48 (1.20)	2.90 (1.01)
Perceived Usefulness	4.82 (1.22)	4.98 (0.69)
Bahavioural Intention	5.02 (1.34)	4.78 (1.29)

Note. Mean values are presented with standard deviations in brackets.

There was a significant difference in *average number of correct answers per answered question* scores between simple Web site (M = 0.61, SD = 0.47) and complex Web site (M = 0.88, SD = 0.27); t (26) = -2.14, p = 0.04 (two-tailed). Therefore, high artefact complexity led to more average number of correct answers per answered question than low artefact complexity.

There was a significant difference in *average number of visited pages per answered question* scores between simple Web site (M = 14.24, SD = 8.03) and complex Web site (M = 5.24, SD = 3.51); t (22) = 4.43, p < 0.0001 (two-tailed). Therefore, high artefact complexity led to less average number of visited pages per answered question than low artefact complexity.

There was a significant difference in *average number of revisited pages per answered question* scores between simple Web site (M = 3.12, SD = 2.98) and complex Web site (M = 0.76, SD = 1.40); t (23) = 3.12, p = 0.005 (two-tailed). Therefore, high artefact complexity led to less average number of revisited pages per answered question than low artefact complexity.

There was a significant difference in *average number of times of visiting the homepage per answered question* scores between simple Web site (M = 1.82, SD = 1.19) and complex Web site (M = 1.07, SD = 0.21); t (18) = 2.65, p = 0.02 (two-tailed). Therefore, high artefact complexity led to less average number of times of visiting the homepage per answered question than low artefact complexity.

There was a significant difference in *average time-on-task per answered question* scores between simple Web site (M = 168544.4, SD = 83208.10) and complex Web site (M = 70753.94, SD = 60747.12); t (39) = 4.35, p < 0.0001 (two-tailed). Therefore, high artefact complexity led to less average time-on-task per answered question than low artefact complexity.

Results of repeated non-think-aloud data analyses

As done for non-think-aloud data, main effects of artefact complexity on outcome measures were also investigated for repeated non-think-aloud data. All the results were the same as the results of non-think-aloud data except enjoyment scores: There was no significant difference in *enjoyment* scores between simple Web site (M = 4.31, SD = 0.78) and complex Web site (M = 4.45, SD = 0.58); t (39) = -0.67, p = 0.505.

Summary of the results

Results from non-think-aloud data indicated that high artefact complexity led to less mental effort of participants. This was inconsistent with Hypothesis 5. Besides, high artefact complexity led to more average number of correct answers per answered question, less average number of visited pages per answered question, less average number of revisited pages per answered question, less average number of times of visiting the homepage per answered question and less average time-on-task per answered question. Results from repeated non-think-aloud data also indicated the same statistically significant results. However, there were no other statistically significant results for the main effects of artefact complexity on outcome measures.

4.1.4 Moderator analysis

Miles and Shevlin (2001) explained moderator analysis simply in statistical terms as 'a third variable (*Z*) is said to moderate the relationship between two other variables (*X* and *Y*) if the degree of relationship between *X* and *Y* is affected by the level of *Z*'. In this section, hierarchical multiple regression was used to test the moderator effects of both task complexity and artefact complexity.

4.1.4.1 Moderator effect of task complexity

Regarding Hypothesis 6a, the effect of individual-difference variables moderated by task complexity on user experience outcomes was investigated in this section.

Results of non-think-aloud data

In this part data from participants with high artefact complexity was used in order to investigate the moderator effect of task complexity on outcome measures.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *mental effort*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables explained a significant amount of variability on *mental effort*, $R^2 = 0.44$, $F\ (5, 37) = 5.79$, $p < 0.001$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 47%, $R^2 = 0.47$, $F\ (7, 35) = 4.42$, $p = 0.001$. The moderator effect explained an additional 3% of the variance in *mental effort*, after controlling for task complexity and individual-difference variables, R^2 change = 0.03, F change (2, 35) = 0.99, $p = 0.38$. Therefore, moderated combined effects of individual-difference variables by task complexity had a significant effect on *mental effort* of participants. Simple effect tests were conducted to elucidate the direction of this relationship. Results indicated that, when task complexity was high, the effect of individual-difference variables ($R^2 = 0.33$, $F\ (4, 15) = 1.82$, $p = 0.18$) on mental effort was higher than when task complexity was low ($R^2 = 0.04$, $F\ (4, 36) = 0.42$, $p = 0.79$). On the other hand, people with higher individual-difference variables performing simple tasks experienced higher mental effort (M = 69.87, SD = 33.88) than those with lower individual-difference variables performing simple tasks (M =

20.05, SD = 20.05). Therefore, in contradiction with Hypothesis 6a, moderation of individual-difference variables by low task complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was bigger over those with lower individual-difference variables.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceptions of classic aesthetics*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics*, $R^2 = 0.05$, $F(5, 37) = 0.41$, $p = 0.84$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 7.1%, $R^2 = 0.07$, $F(7, 35) = 0.38$, $p = 0.91$. The moderator effect explained an additional 1.8% of the variance in *perceptions of classic aesthetics*, after controlling for task complexity and individual-difference variables, R^2 change = 0.018, F change (2, 35) = 0.33, $p = 0.72$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceptions of expressive aesthetics*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of expressive aesthetics*, $R^2 = 0.06$, $F(5, 37) = 0.51$, $p = 0.77$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 7%, $R^2 = 0.07$ $F(7, 35) = 0.35$, $p = 0.92$. The moderator effect explained an additional 0.2% of the variance in *perceptions of expressive aesthetics*, after controlling for task complexity and individual-difference variables, R^2 change = 0.002, F change (2, 35) = 0.03, $p = 0.97$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *enjoyment*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *enjoyment*, $R^2 = 0.157$, $F(5, 37) = 1.38$, $p = 0.25$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 16.3%, $R^2 = 0.163$, $F(7, 35) = 0.98$, $p = 0.464$. The moderator effect explained an additional 0.6% of the variance in enjoyment, after controlling for task complexity and individual-difference variables, R^2 change = 0.006, F change (2, 35) = 0.13, p = 0.88.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceived ease of use*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, R^2 = 0.106, F (5, 37) = 0.88, p = 0.51. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 13%, R^2 = 0.13, F (7, 35) = 0.75, p = 0.63. The moderator effect explained an additional 2.4% of the variance in *perceived ease of use*, after controlling for task complexity and individual-difference variables, R^2 change = 0.024, F change (2, 35) = 0.49, p = 0.62.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *disorientation*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *disorientation*, R^2 = 0.09, F (5, 37) = 0.73, p = 0.60. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 10%, R^2 = 0.10, F (7, 35) = 0.54, p = 0.80. The moderator effect explained an additional 1% of the variance in *disorientation*, after controlling for task complexity and individual-difference variables, R^2 change = 0.01, F change (2, 35) = 0.15, p = 0.86.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceived usefulness*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, R^2 = 0.14, F (5, 37) = 1.21, p = 0.32. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 19%, R^2 = 0.19, F (7, 35) = 1.16, p = 0.35. The moderator effect explained an additional 5% of the variance in perceived usefulness, after controlling for task complexity and individual-difference variables, R^2change = 0.05, F change (2, 35) = 1.04, p = 0.36.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *behavioural intention*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *behavioural intention*, $R^2 = 0.07$, F (5, 37) = 0.60, $p = 0.70$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 13%, $R^2 = 0.13$, F (7, 35) = 0.76, $p = 0.62$. The moderator effect explained an additional 5.7% of the variance in *behavioural intention*, after controlling for task complexity and individual-difference variables, R^2 change = 0.06, F change (2, 35) = 1.15, p = 0.33.

Results of repeated non-think-aloud data

In this part data from participants with high artefact complexity was used in order to investigate the moderator effect of task complexity on outcome measures.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *mental effort*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables explained a significant amount of variability on *mental effort*, $R^2 = 0.44$, F (5, 37) = 5.74, $p = 0.001$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 52%, $R^2 = 0.52$, F (7, 35) = 5.46, $p < 0.001$. Therefore, moderated combined effects of individual-difference variables by task complexity had a significant effect on *mental effort* of participants. The moderator effect explained an additional 9% of the variance in *mental effort*, after controlling for task complexity and individual-difference variables, R^2 change = 0.09, F change (2, 35) = 3.12, $p = 0.06$. Only moderation of intrinsic motivation by task complexity was significant with the largest beta value ($\beta = -14.29$, t (35) = -2.02, $p = 0.051$).

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceptions of classic aesthetics*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics* $R^2 = 0.05$, F (5, 37) = 0.36, $p = 0.88$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 6%, $R^2 = 0.6$, $F(7, 35) = 0.34$, $p = 0.93$. The moderator effect explained an additional 2% of the variance in *perceptions of classic aesthetics*, after controlling for task complexity and individual-difference variables, R^2 change = 0.02, F change (2, 35) = 0.32, $p = 0.73$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceptions of expressive aesthetics*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of expressive aesthetics,* $R^2 = 0.06$, $F(5, 37) = 0.45$, $p = 0.81$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 7%, $R^2 = 0.07$, $F(7, 35) = 0.37$, $p = 0.91$. The moderator effect explained an additional 1% of the variance in *perceptions of expressive aesthetics*, after controlling for task complexity and individual-difference variables, R^2 change = 0.01, F change (2, 35) = 0.23, $p = 0.80$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *enjoyment*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *enjoyment*, $R^2 = 0.16$, $F(5, 37) = 1.43$, $p = 0.24$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 16%, $R^2 = 0.16$, $F(7, 35) = 0.97$, $p = 0.47$. The moderator effect explained an additional 0.1% of the variance in *enjoyment*, after controlling for task complexity and individual-difference variables, R^2 change = 0.001, F change (2, 35) = 0.02, $p = 0.98$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceived ease of use*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, $R^2 = 0.10$, $F(5, 37) = 1.43$, $p = 0.24$. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 12%, $R^2 = 0.12$, $F(7, 35) = 0.67$, $p = 0.70$. The moderator effect explained an additional 2% of the variance in *perceived ease of use*, after controlling for task complexity and individual-difference variables, R^2 change = 0.02, F change (2, 35) = 0.31, p = 0.73.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *disorientation*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *disorientation*, R^2 = 0.09, F (5, 37) = 0.68, p = 0.64. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 18%, R^2 = 0.18, F (7, 35) = 1.12, p = 0.37. The moderator effect explained an additional 10% of the variance in *disorientation*, after controlling for task complexity and individual-difference variables, R^2change = 0.10, F change (2, 35) = 2.11, p = 0.14. Only moderation of intrinsic motivation by task complexity was significant with the largest beta value (β = -18.807, t (35) = -2.03, p = 0.05).

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *perceived usefulness*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, R^2 = 0.14, F (5, 37) = 1.21, p = 0.32. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 19%, R^2 = 0.19, F (7, 35) = 1.21, p = 0.33. The moderator effect explained an additional 5% of the variance in *perceived usefulness*, after controlling for task complexity and individual-difference variables, R^2 change = 0.05, F change (2, 35) = 1.17, p = 0.32.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by task complexity (need for cognition by task complexity, self-efficacy as a trait by task complexity, self-efficacy as a state by task complexity, intrinsic motivation by task complexity) to predict *behavioural intention*, after controlling for the influence of task complexity and individual-difference variables. Task complexity and individual-difference variables did not explain a significant amount of variability on *behavioural intention*, R^2 = 0.07, F (5, 37) = 0.53, p = 0.75. After entry of moderated effect of individual-difference variables by task complexity, the total variance explained by the model as a whole was 9%, R^2 = 0.09, F (7, 35) = 0.47, p = 0.85. The moderator effect explained an additional 2% of the variance in *behavioural intention*, after controlling for task complexity and individual-difference variables, R^2change = 0.02, F change (2, 35) = 0.38, p = 0.69.

Summary of the results

There was only one statistically significant result of the moderated effect of individual-difference variables by task complexity on outcome measures: moderation of individual-difference variables by task complexity had a significant effect on *mental effort* of participants. Simple effect tests showed that when task complexity was high, the effect of individual-difference variables on mental effort was higher than when task complexity was low. On the other hand, people with higher individual-difference variables performing simple tasks experienced higher mental effort than those with lower individual-difference variables performing simple tasks. In contradiction with Hypothesis 6a, moderation of individual-difference variables by low task complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was bigger compared to those with lower individual-difference variables.

4.1.4.2 Moderator effect of artefact complexity

Regarding Hypothesis 6b, the effect of individual-difference variables moderated by artefact complexity on user experience outcomes was investigated in this section.

Results of non-think-aloud data

In this part data from participants with low task complexity was used in order to investigate the moderator effect of artefact complexity on outcome measures.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *mental effort*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables explained a significant amount of variability on *mental effort*, $R^2 = 0.57$, $F(5, 35) = 9.45$, $p < 0.001$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 60%, $R^2 = 0.60$, $F(7, 33) = 6.82$, $p < 0.001$. The moderator effect explained an additional 2% of the variance in *mental effort*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.02, F change (2, 33) = 0.67, p = 0.52. Therefore, moderated combined effects of individual-difference variables by artefact complexity had a significant effect on *mental effort* of participants. Simple effect tests were conducted to elucidate the direction of this relationship. Results indicated that the effect of individual-difference variables by high artefact complexity (R^2 = 0.05, F (4, 38) = 0.51, p = 0.73) on mental effort was equally low than those with low artefact complexity (R^2 = 0.07, F (4, 13) = 0.23, p = 0.92). On the other hand, people with higher individual-difference variables using complex Web site experienced higher mental effort (M = 30.56, SD = 23.82)- t (19) = 2.42, p = 0.03) than those with lower individual-difference variables using complex Web site (M = 20.05, SD = 17.22). Therefore, in contradiction with Hypothesis 6b, moderation of individual-difference variables by high artefact complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was barely bigger over those with lower individual-difference variables.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceptions of classic aesthetics*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics*, R^2 = 0.07, F (5, 35) = 0.53, p = 0.75. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 8%, R^2 = 0.08, F (7, 33) = 0.40, p = 0.90. The moderator effect explained an additional 0.7% of the variance in *perceptions of classic aesthetics*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.01, F change (2, 33) = 0.13, p = 0.88.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceptions of expressive aesthetics*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables explained a significant amount of variability on *perceptions of expressive aesthetics*, R^2 = 0.28, F (5, 35) = 2.69, p = 0.04. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 39%, R^2 = 0.39, F (7, 33) = 3.01, p = 0.015. Therefore, moderated combined effects of individual-difference variables by artefact complexity had a significant effect on *perceptions of expressive aesthetics* of participants. The moderator effect explained an additional 11% of the variance in *perceptions of expressive aesthetics*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.11, F change (2, 33) = 3.03, p = 0.06.

In the final model, two measures were statistically significant, with the self-efficacy as a trait recording a lower beta value ($\beta = 0.71$, t (33) = 3.51, $p = 0.001$) than self-efficacy as a trait by artefact complexity ($\beta = -24.04$, t (33) = -2.46, $p = 0.02$). Both self-efficacy as a trait and moderation of self-efficacy as a trait by artefact complexity had significant effects on *perceptions of expressive aesthetics* of participants. Simple effect tests were conducted to elucidate the direction of this relationship. Results indicated that the effect of self-efficacy as a trait by low artefact complexity ($R^2 = 0.42$, F (1, 16) = 11.36, $p = 0.004$) on perceptions of expressive aesthetics was higher than those with high artefact complexity ($R^2 = 0.001$, F (1, 41) = 0.05, $p = 0.82$). On the other hand, test-users with high self-efficacy as a trait using a simple Web site experienced higher perceptions of expressive aesthetics (M = 4.16, SD = 1.04) than those using a complex Web site (M = 3.23, SD = 0.83), – t (28) = 2.59, $p = 0.015$. However, test-users with low self-efficacy as a trait using a complex Web site experienced similar perceptions of expressive aesthetics (M = 3.29, SD = 0.88) to those using a simple Web site (M = 3.16, SD = 0.82), - t (29) = -0.40, $p = 0.69$. Therefore, in contradiction with Hypothesis 6b, the effect of artefact complexity was greater for those with higher self-efficacy as a trait bigger than for those with lower self-efficacy.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *enjoyment*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *enjoyment*, $R^2 = 0.22$, F (5, 35) = 2.01, $p = 0.10$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 26%, $R^2 = 0.26$, F (7, 33) = 1.61, $p = 0.17$. The moderator effect explained an additional 3% of the variance in *enjoyment*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.03, F change (2, 33) = 0.70, $p = 0.50$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceived ease of use*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, $R^2 = 0.04$, F (5, 35) = 0.31, $p = 0.91$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 6%, $R^2 = 0.06$, F (7, 33) = 0.30, $p = 0.10$. The moderator effect explained an additional 2% of the variance in *perceived ease of use*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.02, F change (2, 33) = 0.31, p = 0.74.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *disorientation*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *disorientation*, R^2 = 0.10, F (5, 35) = 0.80, p = 0.56. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 13%, R^2 = 0.13, F (7, 33) = 0.67, p = 0.70. The moderator effect explained an additional 2% of the variance in *disorientation*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.02, F change (2, 33) = 0.41, p = 0.67.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceived usefulness*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, R^2 = 0.08, F (5, 35) = 0.58, p = 0.71. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 11%, R^2 = 0.11, F (7, 33) = 0.57, p = 0.77. The moderator effect explained an additional 3% of the variance in *perceived usefulness*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.03, F change (2, 33) = 0.59, p = 0.56.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *behavioural intention*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *behavioural intention*, R^2 = 0.07, F (5, 35) = 0.55, p = 0.74. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 8%, R^2 = 0.08, F (7, 33) = 0.43, p = 0.87. The moderator effect explained an additional 1% of the variance in *behavioural intention*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.01, F change (2, 33) = 0.20, p = 0.82.

Results of repeated non-think-aloud data

In this part data from participants with low task complexity was used in order to investigate the moderator effect of artefact complexity on outcome measures.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *mental effort*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables explained a significant amount of variability on *mental effort*, $R^2 = 0.58$, $F\ (5, 35) = 9.49$, $p < 0.001$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 60%, $R^2 = 0.60$, $F\ (7, 33) = 7.05$, $p < 0.001$. Therefore, moderated combined effects of individual-difference variables by artefact complexity had a significant effect on *mental effort* of participants. The moderator effect explained an additional 2% of the variance in *mental effort*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.02, F change (2, 33) = 0.98, $p = 0.39$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceptions of classic aesthetics*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceptions of classic aesthetics*, $R^2 = 0.07$, $F\ (5, 35) = 0.51$, $p = 0.77$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 7%, $R^2 = 0.07$, $F\ (7, 33) = 0.35$, $p = 0.93$. The moderator effect explained an additional 0.1% of the variance in *perceptions of classic aesthetics*, after controlling for artefact complexity and individual-difference variables, R^2change = 0.001, F change (2, 33) = 0.02, $p = 0.98$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceptions of expressive aesthetics*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables explained a significant amount of variability on *perceptions of* *expressive aesthetics*, $R^2 = 0.28$, F (5, 35) = 2.67, $p = 0.04$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 40%, $R^2 = 0.40$, F (7, 33) = 3.16, $p = 0.011$. Therefore, moderated combined effects of individual-difference variables by artefact complexity had a significant effect on *perceptions of expressive aesthetics* of participants. The moderator effect explained an additional and significant 12.5% of the variance in *perceptions of expressive aesthetics*, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.13, F change (2, 33) = 3.44, $p = 0.044$. In the final model, two measures were statistically significant, with the self-efficacy as a trait recording a lower beta value ($\beta = 0.73$, t (33) = 3.67, p = 0.001) than moderation of self-efficacy as a state by artefact complexity ($\beta = -25.650$, t (33) = 3.67, $p = 0.014$). In contradiction with Hypothesis 6b, for individuals with a high self-efficacy as a state the effect of artefact complexity on *perceptions of expressive aesthetics* was bigger than that for individuals with a low intrinsic motivation.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *enjoyment*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *enjoyment*, $R^2 = 0.22$, F (5, 35) = 1.10, $p = 0.104$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 30%, $R^2 = 0.30$, F (7, 33) = 1.98, $p = 0.09$. The moderator effect explained an additional 7% of the variance in enjoyment, after controlling for artefact complexity and individual-difference variables, R^2 change = 0.07, F change (2, 33) = 1.72, $p = 0.19$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceived ease of use*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceived ease of use*, $R^2 = 0.04$, F (5, 35) = 0.32, $p = 0.90$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 6%, $R^2 = 0.06$, F (7, 33) = 0.31, $p = 0.95$. The moderator effect explained an additional 2% of the variance in *perceived ease of use*, after controlling for artefact complexity and individual-difference variables, R^2change = 0.02, F change (2, 33) = 0.31, $p = 0.73$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *disorientation*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *disorientation*, $R^2 = 0.11$, $F\ (5, 35) = 0.83$, $p = 0.54$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 13%, $R^2 = 0.13$, $F\ (7, 33) = 0.67$, $p = 0.70$. The moderator effect explained an additional 2% of the variance in *disorientation*, after controlling for artefact complexity and individual-difference variables, R squared change = 0.02, F change (2, 33) = 0.36, $p = 0.70$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *perceived usefulness*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *perceived usefulness*, $R^2 = 0.09$, $F\ (5, 35) = 0.67$, $p = 0.65$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 9%, $R^2 = 0.09$, $F\ (7, 33) = 0.46$, $p = 0.86$. The moderator effect explained an additional 0.2% of the variance in *perceived usefulness*, after controlling for artefact complexity and individual-difference variables, R^2change = 0.002, F change (2, 33) = 0.04, $p = 0.96$.

Hierarchical multiple regression was used to assess the ability of the moderated effect of individual-difference variables by artefact complexity (need for cognition by artefact complexity, self-efficacy as a trait by artefact complexity, self-efficacy as a state by artefact complexity, intrinsic motivation by artefact complexity) to predict *behavioural intention*, after controlling for the influence of artefact complexity and individual-difference variables. Artefact complexity and individual-difference variables did not explain a significant amount of variability on *behavioural intention*, $R^2 = 0.07$, $F\ (5, 35) = 0.56$, $p = 0.73$. After entry of moderated effect of individual-difference variables by artefact complexity, the total variance explained by the model as a whole was 8%, $R^2 = 0.08$, $F\ (7, 33) = 0.41$, $p = 0.89$. The moderator effect explained an additional 1% of the variance in *behavioural intention*, after controlling for artefact complexity and individual-difference variables, R^2change = 0.01, F change (2, 33) = 0.10, $p = 0.90$.

Summary of the results

Non-think-aloud data indicated that moderated combined effects of individual-difference variables by artefact complexity had a significant effect on both *mental effort* and *perceptions of expressive aesthetics* of participants. Simple effect tests showed that people with higher individual-difference variables using complex Web site experienced higher mental effort than those with lower individual-difference variables using complex Web site. In contradiction with Hypothesis 6b, moderation of individual-difference variables by high artefact complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was barely bigger compared to those with lower individual-difference variables.

Moderation of self-efficacy as a trait by artefact complexity had significant effects on *perceptions of expressive aesthetics* of participants. Simple effect tests indicated that the moderated effect of self-efficacy as a trait by low artefact complexity on perceptions of expressive aesthetics was higher than those with high artefact complexity. On the other hand, people with high self-efficacy as a trait using simple Web site experienced higher perceptions of expressive aesthetics than those using complex Web site. In contradiction with Hypothesis 6b, the effect of artefact complexity was greater for those with higher self-efficacy as a trait bigger than for those with lower self-efficacy.

4.2 Qualitative results

Think-aloud protocols were conducted to obtain qualitative data. These qualitative findings from think-aloud sessions were as follows:

Participant 1 using version 2 reported positive aspects of the complex Web site as the availability of an index, menu on the left-hand side and use of colours, while she was specifying too much information as a negative aspect of the Web site. She found some tasks inconsistent with the links on the page and had some terminology difficulties.

Participant 2 using version 1 was not successful at performing tasks. She found simple tasks in the simple site difficult to do since she was not good at both understanding English and using a Web site. She likely does not have so much experience at using a Web site. She said that there were so many words she did not understand within the Web site. In summary, participant 2 found design of the Web site, use of colours and graphics positive. While she had experienced many terminology and comprehension difficulties with the Web site, she reported there was too much text in the Web site as a negative aspect of the Web site.

Participant 3 using version 3 found the pictures on the complex Web site pretty and thought that there was a lot of information on the Web site. Participant 3 thought that some pages were completely different from the general appearance of the Web site. In summary, participant 3 using version 3 reported too much information, the availability of an index and search engine as a good aspect of the Web site, while he found task and relevant link inconsistent. He also had terminology problems.

Participant 4 using version 3 had a tendency of using the three links at the top of the homepage, namely residents-businesses-visitors. This likely stemmed from their bright colours that are differed from the rest of the homepage. In addition, he preferred to use search engine rather than using the main menu at the left-hand side of the homepage. He explained the reason that there was too many links and it would take too much time to search and click within the site. He did not want to do this. In summary, while participant 4 using version 3 liked use of colours on the complex Web site, he found navigation on the Web site including search engine and other links too poor.

Participant 5 using version 1 reported positive aspects of the simple Web site as the availability of an index and site map, while she found task inconsistent with the link in one of the questions. Participant 6 using version 2 reported use of colours with user-friendly graphics and easy to read text as the positive aspects of the complex Web site. She reported negative aspect of the Web site as too much information and 'residents-businesses-visitors' tab. Participant 7 using version 3 found too many links and use of colours as favourable aspects, while he found events calendar and use of colours confusing. While participant 8 using version 3 found some tasks inconsistent with the link, she specified the positive aspects of the complex Web site as navigation, too much information, search engine and the availability of an index. Participant 9 using version 2 found search engine and layout of the complex Web site positive, while he reported negative aspects of the Web site as menu on the left-hand side and placement of the graphics.

The qualitative results from non-think-aloud data were presented inTable 6 andTable 7. They showed some usability problems for complex Web site: too many links, too much information, too many advertisements, poor usability, use of colours (poor diversity in colours for different services), poor table of contents (site map), too big letters, poor graphical design, poor navigation, unclear service name and content, unclear menu at left-hand side and poor search engine. Participants defined positive aspects of the complex Web site as good search engine, good graphic design, easy to use, consistent service name and content, use of colours (diversity in colours for different services), clear layout, too much information, clear Web site, easy to read and easy to navigate. One of the participants noted that screen shots at the beginning of the tasks were very useful to comprehend the complex Web site. The qualitative results from non-think-aloud data also showed some usability problems for simple Web site: too much information, poor navigation, too much text, too many links, poor search engine, lack of a dropdown menu, unclear service name and content and poor usability. Participants reported positive aspects of the simple Web site as the availability of an index, less text, less colours, easy to read, use of colours (diversity in colours for different services), clear layout, good search engine, nice graphic design, clear table of contents (site map) and easy to navigate. In general, participants found the simple Web site very clear and comprehensible.

Table 6 Best Aspects of the Complex and Simple Web Sites from Non-think-aloud data

Best Aspects of the Web Site			
for complex Web site		for simple Web site	
Ability of search engine	*18	Availability of an index	5
Graphic design	2	Less text	1
Easy to use	4	Less colours	1
Consistent service name and content	3	Easy to read	1
Use of colours	1	Use of colours	3
Layout	8	Layout	3
Too much information	2	Easy to navigate	2
Easy to navigate	8	Ability of search engine	1
Easy to read	1	Clear Web site	1
Clear Web site	2	Graphic design	1
Screenshots	1	Table of contents (site map)	1

* Frequency

Table 7 Worst Aspects of the Complex and Simple Web sites from Non-Think-Aloud Data

Worst Aspects of the Web Site			
for complex Web site		for simple Web site	
Ability of search engine	*1	Too much information	1
Poor graphical design and visuality	4	Poor navigation	3
Poor usability	1	Too much text	1
Unclear service name and content	1	Too many links	1
Use of colours	1	Poor search engine	1
Too much information	9	Lack of drop-down menu	1
Too many links	7	Poor usability	1
Poor navigation	3	Unclear Web site	1
Poor table of contents (site map)	1	Unclear service name and content	1
Too big letters	1		
Unclear menu on the left-hand side	1		
Too many advertisements	1		

*Frequency

The qualitative results from think-aloud data were presented inTable 8 andTable 9. They showed some usability problems for complex Web site: unclear service name and content, too much information, too many links, poor navigation and unclear 'residents-businesses-visitors' tab. Participants defined positive aspects of the complex Web site as ability of search engine, availability of an index, clear graphical design, use of colours (diversity in colours for different services), clear layout, too much information, easy to navigate, easy to read and clear menu on the left-hand side. The qualitative results from think-aloud data also showed some usability problems for simple Web site: too much information, too much text and unclear service name and content. Participants reported positive aspects of the simple Web site as availability of an index, use of colours (diversity in colours for different services), use of pictures, clear graphical design and clear table of contents (site map).

Table 8 Best Aspects of the Complex and Simple Web sites from Think-Aloud Data

Best Aspects of the Web Site			
for complex Web site		for simple Web site	
Ability of search engine	*3	Availability of an index	1
Availability of an index	3	Use of colours	1
Graphic design	1	Use of pictures	1
Use of colours	4	Graphic design	1
Clear layout	1	Table of contents (site map)	1
Too much information	1		
Easy to navigate	1		
Easy to read	1		
Clear menu on the left-hand side	2		

*Frequency

Table 9 Worst Aspects of the Complex and Simple Web sites from Think-Aloud Data

Worst Aspects of the Web Site			
for complex Web site		for simple Web site	
Unclear service name and content	2	Too much information	1
Too much information	1	Unclear service name and content	1
Too many links	1	Too much text	1
Poor navigation	2		
Unclear 'residents-businesses-visitors' tab	1		

*Frequency

Analyses of the both non-think-aloud and think-aloud data demonstrated an interesting result that too much information was found as both positive and negative aspects of the Web sites by participants. In other words, while some participants reported too much information as an unfavourable aspect of the Web site, some of them reported it as a favourable aspect of the Web site.

4.3 Conclusion

In this chapter, the results of the data analyses were demonstrated. Many statistical data analyses methods were conducted in order to test the hypotheses and to acquire quantitative results. Simple effect tests were conducted just for non-think-aloud data because in overall repeated analyses did not make too much contribution to the research. As mentioned before, because there were some low Cronbach's alpha scores in two questionnaires, analyses were conducted twice: once for non-think-aloud data with all items of questionnaires and once for non-think-aloud data without two items of questionnaires. Even though these repeated non-think-aloud data analyses yielded hardly different scores, the results of them were presented in this section. However, while the findings from non-think-aloud data analyses will be discussed in the next chapter, those from repeated non-think-aloud analyses will not be mentioned in the discussion chapter. The findings from qualitative data also will be discussed in the discussion chapter.

5. Discussion and conclusion

The current study aimed to study the influence of individual-difference variables (need for cognition as a trait, self-efficacy as a trait, self-efficacy as a state and intrinsic motivation) on user experience and technology acceptance, moderated by artefact and task characteristics. Both a larger-scale quantitative experiment and a smaller-scale qualitative study using the same experiment with a qualitative procedure added have been employed to achieve this aim.

In summary, while some of the hypotheses proposed could be confirmed in the current study, a few of them were rejected by the findings. In overall, a significant effect of individual-difference variables on perceived enjoyment was found. In accordance with Hypothesis 1, people with a high need for cognition experienced higher perceived enjoyment than those with low need for cognition. However, in contradiction with Hypothesis 1, people with a high need for cognition experienced lower perceptions of expressive aesthetics than those with low need for cognition.

There are two possible explanations to this inconsistency. First, expressive aesthetics may have caused an ambiguity by breaking traditional design conventions and may have decreased the cognitive motivation of individuals with a high need for cognition for information acquisition and processing. Second, as supported in the paper of van Schaik and Ling (2008b), expressive aesthetics may have led to lower perceptions of aesthetics than classic aesthetics does.

The findings did not support Hypothesis 1 for other outcome measures: people with a high need for cognition did not experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceptions of classic aesthetics, higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low need for cognition.

In accordance with Hypothesis 2, people with high self-efficacy as a state experienced higher perceived enjoyment than those with low self-efficacy as a state. The findings did not support Hypothesis 2 for other outcome measures: people with a high self-efficacy (trait and state) did not experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceptions of aesthetics (classic and expressive), higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low self-efficacy (trait and state).

In accordance with Hypothesis 3, people with a high intrinsic motivation experienced higher average time-on-task per answered question than those with low need for cognition. No strong evidence was found to verify all the aspects of Hypothesis 3: people with a high intrinsic motivation did not experience less mental effort, lower perceived disorientation, higher perceived usefulness, higher perceived ease of use, higher perceived enjoyment, higher perceptions of aesthetics (classic and expressive), higher task performance and navigation behaviour, and higher perceived behavioural intention than those with low intrinsic motivation.

Because some items of two questionnaires (item 10 in the subscale of ‘identified motivation’ of SIMS and item 8 in the subscale of ‘transformation of time’ of Flow State Scale) diminished the reliability of the scales used in the current study, data analysis was repeated by removing these two items in order to both increase reliability of scales and compare the results with the non-think-aloud data analyses. This re-conducted data analysis was called ‘repeated non-think-aloud data analyses’ in the current study. However, the findings from repeated non-think-aloud data analyses demonstrated the same effects of individual-difference variables on outcome measures, as found in the non-think-aloud data.

In accordance with Hypothesis 4, high task complexity led to more mental effort and lower perceived usefulness of participants. However, high task complexity did not demonstrate an effect leading to higher perceived disorientation, lower perceived ease of use, lower perceived enjoyment, lower perceptions of aesthetics (classic and expressive), lower task performance and navigation behaviour, and lower perceived behavioural intention than low task complexity. In addition, high task complexity led to less number of correct main answers, more average time-on-task for incorrectly answered main questions, less average number of correct answers per answered question, more average number of visited pages per answered question, more average number of times of visiting the homepage per answered question, more average number of times of visiting a page with search results per answered question and more average time-on-task per answered question. The findings from repeated non-think-aloud data analyses have revealed the same effects of individual-difference variables on outcome measures, as found in the non-think-aloud data. Therefore, these findings are not presented here again.

In contradiction with Hypothesis 5, high artefact complexity led to less mental effort of participants than low artefact complexity. Therefore, people using a complex Web site experienced less mental effort than those using a simple Web site. A possible explanation to this is that the complex e-government Web site employed in the experiment may be more comprehensible and clearer than the simple one because it provided much more information to users than the simple Web site did. Nonetheless, high artefact complexity did not lead to higher perceived disorientation, lower perceived usefulness, lower perceived ease of use, lower perceived enjoyment, lower perceptions of aesthetics (classic and expressive), lower task performance and navigation behaviour, and lower perceived behavioural intention than low artefact complexity. On the other hand, high artefact complexity led to more average number of correct answers per answered question, less average number of visited pages per answered question, less average number of revisited pages per answered question, less average number of times of visiting the homepage per answered question and less average time-on-task per answered question. The findings from repeated non-think-aloud data analyses have revealed the same effect of task complexity on outcome measures, as found in the non-think-aloud data.

Comparing results of Hypothesis 4 and Hypothesis 5 indicated that while test subjects performing in complex Web site experienced lower mental effort, test subjects performing complex tasks experienced higher mental effort. This interesting finding obviously points out both the negative effect of task complexity and the positive influence of Web site complexity on mental effort of individuals.

In support of Hypothesis 6, a statistically significant moderated effect of combined individual-difference variables by task complexity on mental effort was found. Simple effect tests indicated that, when task complexity was high, the effect of individual-difference variables on mental effort was higher than when task complexity was low. On the other hand, people with higher individual-difference variables performing simple tasks experienced higher mental effort than those with lower individual-difference variables performing simple tasks. Therefore, in contradiction with Hypothesis 6a, moderation of individual-difference variables by low task complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was bigger compared to those with lower individual-difference variables. Besides, in support of Hypothesis 6, a statistically significant moderated effect of combined individual-difference variables by artefact complexity on both mental effort and perceptions of expressive aesthetics was found. Simple effect tests indicated that the effect of individual-difference variables by high artefact complexity on mental effort was equally low than those with low artefact complexity. On the other hand, people with higher individual-difference variables using complex Web site experienced higher mental effort than those with lower individual-difference variables using complex Web site. Therefore, in contradiction with Hypothesis 6b, moderation of individual-difference variables by high artefact complexity for individuals with higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation was bigger compared to those with lower individual-difference variables. Simple effect tests also indicated that the effect of self-efficacy as a trait by low artefact complexity on perceptions of expressive aesthetics was higher than those with high artefact complexity. On the other hand, test-users with high self-efficacy as a trait using a simple Web site experienced higher perceptions of expressive aesthetics than those using a complex Web site. However, test-users with low self-efficacy as a trait using a complex Web site experienced similar perceptions of expressive aesthetics to those using a simple Web site. Therefore, in contradiction with Hypothesis 6b, the effect of artefact complexity was greater for those with higher self-efficacy as a trait bigger than for those with lower self-efficacy. The findings from repeated non-think-aloud data analyses revealed the same moderated effect of task and artefact complexity.

In non-think-aloud data positive correlations were found between self-efficacy as a state and enjoyment and between intrinsic motivation and average time-on-task per answered question. While high self-efficacy as a state heightened perceived enjoyment of individuals, high intrinsic motivation heightened average time-on-task per answered question. The same correlations with non-think-aloud data were found in repeated non-think-aloud data. In think-aloud data, correlations were found between need for cognition and enjoyment, between self-efficacy as a state and average time-on-task per answered question, and between intrinsic motivation and disorientation, average pages loaded for correct answers, average time-on-task for correctly answered main questions, average number of visited pages per answered question and average number of revisited pages per answered question. While high need for cognition increased perceived enjoyment of individuals, high self-efficacy as a state decreased average time-on-task per answered question. High intrinsic motivation heightened disorientation of individuals, which was inconsistent with the proposition of Hypothesis 3.

In addition, the relation between the concepts of need for cognition and intrinsic motivation was investigated because some authors did not make a sharp distinction between need for cognition and intrinsic motivation. Although it was not statistically significant, there was an interesting negative correlation between need for cognition and intrinsic motivation. While people had a high need for cognition, those had a low intrinsic motivation. This finding demonstrates an unexpected relation between the concepts of need for cognition and intrinsic motivation, explaining that these two represent a contrary association. Therefore, these two must be elaborated as separate concepts.

Regarding the quantitative findings presented, it can be said that this study confirmed the contribution of three interacting components of the PAT model (person, artefact and task) to user experience of public-sector Web sites. As stated in the PAT model, diverse interaction was found between person, artefact and task in a computer-based context. The findings demonstrated that the level of individual-difference variables had an effect on user experience outcomes as moderated by artefact characteristics (complexity of public-service Web sites) and task characteristics (complexity of task). Moreover, the current study contributed to the literature by developing a psychological model of users' experience of public-service Web sites (Figure 5). In overall, the final model of psychological model of users' experience of public-service Web sites enhances the understanding of user experience and moderated effect of user characteristics by Web site characteristics and task characteristics. Therefore, the final model suggests that moderation of individual-difference variables by low task complexity and high artefact complexity for individuals with a higher need for cognition, higher self-efficacy (trait and state) and higher intrinsic motivation the effect of task complexity and artefact complexity is bigger compared to those with lower need for cognition, lower self-efficacy (trait and state) and lower intrinsic motivation.

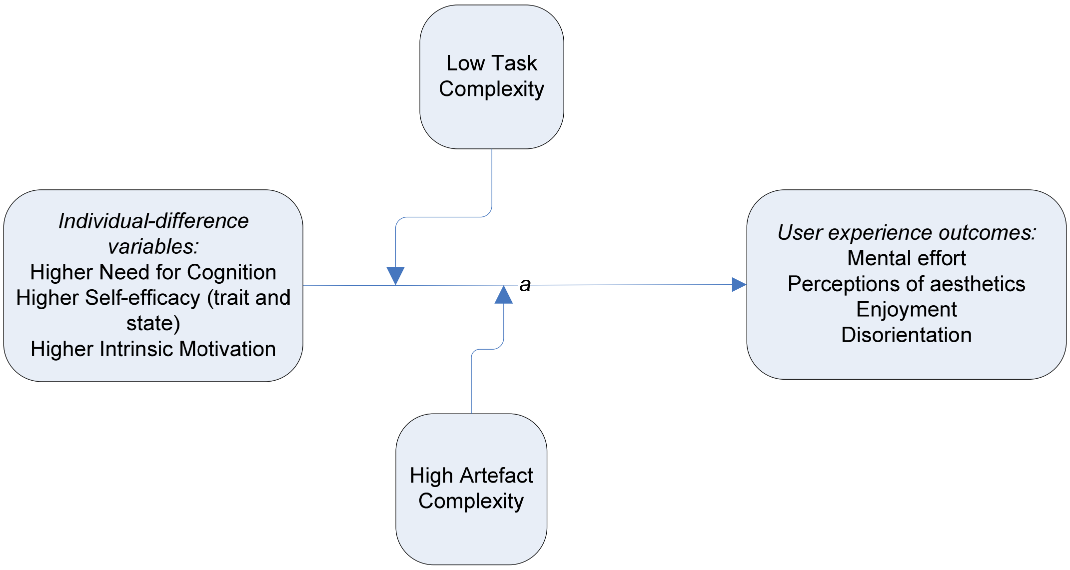

Figure 5 The final model of psychological model of users' experience of public-service Web sites (PAT-UX model)
[a] Higher effect of moderators

The findings of think-aloud sessions demonstrated that paradox of active user is a powerful principle in explaining the behaviour of users. Test-users tended to use the functions they are familiar with (e.g. using search engine) while performing tasks on the Web sites. They avoided using new methods, such as using the links on the homepage, to accomplish the tasks. If they would try some functions they do not use often, they would likely have found the answers rapidly and easily, especially to perform simple tasks. Most of the time participants preferred less efficient ways to perform the tasks. Therefore, the most efficient path to find the right differed from the path to the right answer taken by participants.

Even though the analysis of think-aloud data did not yield concrete quantitative results because of the low number of participants, the qualitative findings from this data were highly contributive in comprehending users' experience. The findings from non-think-aloud data demonstrated some positive aspects of the complex Web site: good search engine, good graphic design, easy to use, consistent service name and content, use of colours (diversity in colours for different services), clear layout, too much information, clear Web site, easy to read and easy to navigate. On the other hand, there were some unfavourable aspects of the complex Web site: too many links, too much information, too many advertisements, poor usability, use of colours (poor diversity in colours for different services), poor table of contents (site map), too big letters, poor graphical design, poor navigation, unclear service name and content, unclear menu at left-hand side and poor search engine. One of the participants noted that screen shots at the beginning of the experiment were very useful to comprehend the complex Web site.

The positive aspects of the simple Web site from non-think-aloud data were as follows: the availability of an index, less text, less colours, easy to read, use of colours (diversity in colours for different services), clear layout, good search engine, nice graphic design, clear table of contents (site map) and easy to navigate. In general, participants found the simple Web site very clear and comprehensible.

The qualitative results from think-aloud data showed some usability problems for complex Web site: unclear service name and content, too much information, too many links, poor navigation and unclear 'residents-businesses-visitors' tab. Participants defined positive aspects of the complex Web site as ability of search engine, availability of an index, clear graphical design, use of colours (diversity in colours for different services), clear layout, too much information, easy to navigate, easy to read and clear menu on the left-hand side. The qualitative results from think-aloud data also showed some usability problems for simple Web site: too much information, too much text and unclear service name and content. Participants reported positive aspects of the simple Web site as availability of an index, use of colours (diversity in colours for different services), use of pictures, clear graphical design and clear table of contents (site map).

Analyses of the both non-think-aloud and think-aloud data demonstrated an interesting result that too much information was found as both positive and negative aspects of the Web sites by participants. In other words, while some participants reported too much information as an unfavourable aspect of the Web site, some of them reported it as a favourable aspect of the Web site.

6. Limitations and future research

There were some limitations in the current study. The first limitation of this study was the lack of an automated tool for determining Web site complexity. Although recent research has mentioned the existence of an automated Web site evaluation tool (Ivory & Megraw, 2005; Ivory, 2003; Ivory et al., 2003), it was not possible to find and use one of these automated Web site measurement tools. Therefore, the complexity level of two public-service Web sites was measured by the researcher in terms of the objective Web site complexity metrics. This limitation should be considered by the future research. In addition, the current study used objective Web site complexity metrics and did not test participants' perception of artefact complexity. Their perceptions of artefact complexity could lead to different impact on user experience outcomes. Future research could measure participants' perception of artefact complexity and select complex and simple Web sites in terms of these complexity perceptions.

The second limitation of this study was the comprehension problems of test-users. Think-aloud sessions demonstrated that most of the participants had some comprehension difficulties because of the lack of understanding the English language. Participants in the quantitative experiment likely had these comprehension difficulties, too. Therefore, this limitation should be taken into account in evaluation of the findings and it is strongly recommended for future studies to use native test-users for both the quantitative experiment and think-aloud protocols.

The number of participants in the quantitative study was sufficiently high to achieve a relatively high power (according to a retrospective power analysis power was 0.72 for the effects of both artefact complexity and task complexity). The sample size of the qualitative study was deliberately small to allow detailed study of participants' behaviour and responses within the time available for the project. The qualitative data complemented the quantitative study. However, future research may use more participants for the quantitative and qualitative studies.

Even though recent research on persistence of individuals revealed a strong association between persistence and the constructs of need for cognition and self-efficacy, this was not tested by the current study. On the one hand, moderation of individual-difference variables by artefact complexity and task complexity is likely to have an influence on persistence level of individuals through use of public-service Web sites. On the other hand, persistence level of individuals likely has an effect on user experience outcomes. Therefore, future research should take into account persistence both as a user experience outcome and as an individual-difference variable.

7. Practical implications

The findings of this study provide some practical implications for the design of e-government Web sites. As suggested by the final model of this study, the strong moderator effects of high Web site complexity and low task complexity on mental effort, disorientation, perceptions of aesthetics and enjoyment should be taken into account while designing an e-government Web site for audiences with higher individual-difference variables. Accordingly, as governments intend to establish an online presence, they should consider this model. Furthermore, governments can consider this model while developing a segmentation strategy in order to focus unique needs of different segments. For instance, when segmenting citizens based upon individual-difference variables and an e-government Web site needs to address people with higher individual-difference variables, Web designer will know the fact that moderated effect of simple tasks in a complex Web site is strong. Ultimately, psychological model of users' experience of public-service Web sites flourishes the knowledge of governments on both the characteristics of their audiences and the effect of artefact and task complexity.

When considering the quantitative and qualitative results of this study, some design recommendations can be made for the future studies. Because people with high need for cognition experienced lower perceptions of expressive aesthetics, an e-government Web site should be designed according to the characteristics of classical aesthetics. In addition, providing diversity of colours to represent different services should be considered while designing a public-service Web site. The availability of an index, the ability of search engine, consistency of service name and content, availability of a clear menu structure, a comprehensive table of contents (site map), a successful graphical design and use of fewer advertisements are other dimensions of a usable e-government Web site design. Nevertheless, because some test-users found too much information as a positive aspect of an e-government Web site while some of them defined too much information as a negative aspect, it is overtly impossible to conclude the amount of information as a usable design recommendation.

References

Ajzen, I. (2002). Perceived behavioural control, self-efficacy, locus of control and the theory of planned behaviour. *Journal of Applied Social Psychology,* 32, 665-683.

Agarwal, R., Sambamurthy, V., & Stair, R. M., (2000). Research report: The evolving relationship between general and specific computer self-efficacy: An empirical assessment. *Information Systems Research*, 11, 418-430.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology use. *MIS Quarterly*, 24, 665-694.

Ahuja, J. S. & Webster, J. (2001). Perceived disorientation: An examination of e new measure to assess web design effectiveness. *Interacting with Computers*, 14, 15-29.

Alben, L. (1996). Quality of experience: Defining the criteria for effective interaction design. *Interactions*, 3, 11-15.

Arnold, A.G., 1999. Mental effort and evaluation of user interfaces: A questionnaire approach. In: Bullinger, H.-J., Ziegler, J. (Eds.), *Proceedings of the HCII '99 international conference on human– computer interaction, vol. 1.* Lawrence Erlbaum, Mahwah, NJ, pp. 1003–1007.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, 44, 1175-1184.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Education Psychologist*, 28, 117-148.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Bandura, A. (1977). *Social learning theory*. Prentice-Hall, Englewood Cliffs, NJ.

Banker, R. D., Davis, G. B., & Slaughter, S. A. (1998). Software development practices, software complexity, and software maintenance performance: A field study. *Management Science,* 44, 433-450.

Bélanger, F. & Hiller, J. S. (2006). A framework for e-government: Privacy implications**.** *Business Process Management*, 12, 48-60.

Bell, D. J. & Ruthven, I. (2004). Searcher's assessments of task complexity for Websearching. *Advances in Information Retrieval Proceedings,* 2997, 57-71.

Blackmon, M., Polson, P. G., Kitajima, M., & Lewis, C. (2002). Cognitive walkthrough for the Web. *CHI Letters*, 4, 463-470.

Byström, K. & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, 31, 191-213.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197-253.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48, 306-307.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review,* 13, 40-52.

Carroll, J. M. & Rosson, M. B. (1987). Paradox of the active user. In J. M. Carroll (Ed.), *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction* (pp. 80-111). Cambridge, MA: MIT Press.

Chen, H., Wigand, R. T., & Nilan, M. S. (1999). Optimal experience of Web activities**.** *Computers in Human Behaviour,* 15, 585-608.

Chi, E. H., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions on the Web. *CHI*, 3, 490-497.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly,* 19, 189-211.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.

Csikszentmihalyi, M. & Nakamura, J. (1989). The dynamics of intrinsic motivation: A study of adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 45-71). New York: Academic Press.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15-35). Cambridge, United Kingdom: Cambridge University Press.

Dalal, N. P., Quible, Z., & Wyatt, K. (2000). Cognitive design of home pages: An experimental study of comprehension on the World Wide Web. *Information Processing and Management*, 36, 607-621.

David, P., Song, M., Hayes, A., & Fredin, E. S. (2007). A cyclic model of information seeking in hyperlinked environments: The role of goals, self-efficacy, and intrinsic motivation. *International Journal of Human-Computer Studies*, 65, 170-182.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology**.** *MIS Quarterly**,*** 13, 318-340.

Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self- determination in human behaviour*. . New York: Plenum Press.

Dijk, J. A. G. M. van, Ebbers, W. E., Fennis, B. M., Geest, T. M. van der, Loorbach, N. R., Pieterson, W. J., Steehouder, M. F., Taal, E. & Vries, P. W. de (2005). Alter ego: State of the art on user profiling. *Telematica Instituut*.

Dijk, J. A. G. M. van, Peters, O., & Ebbers, W. (2007). Explaining the acceptance and use of government Internet services: A multivariate analysis of 2006 survey data in the Netherlands. *Government Information Quarterly,* 25, 379-399.

Doornbos, L. (2004). Vraaggestuurd werken en segmentatie. *Veranderend ondernemen in de non-profit sector.* Retrieved December 4 2007, via http://www.concern.nl/leesmap/ doornbos2004.pdf

Ebbers, W. E., Pieterson, W. J., & Noordman, H. N. (2008). Electronic government: Rethinking channel management strategies. *Government Information Quarterly*, 25, 181-201.

Finneran, C. M., & Zhang, P. (2003). A person-artefact-task (PAT) model of flow antecedents in computer-mediated environments. *International Journal of Human-Computer Studies*, 59, 475-496.

Finneran, C.M., Zhang, P., 2002. The challenges of studying flow within a computer-mediated environment. Paper presented at the Americas Conference on Information Systems, Dallas, TX.

Fishbein, M. & Ajzen, I. (1975). Belief, attitude, intention and behaviour: An introduction to theory and research. New York, NY: Addison-Wesley.

Frese, M. (1987). A theory of control and complexity: Implications for software design and integration of computer systems into the work place. In M. Frese, E.

Frey, B. S. (1997). *Not just for the money*. An economic theory of personal motivation. Brookfield: Edward Elgar Publishing.

Fu, W-T. & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science*, 28, 901-935.

Geissler, G. L., Zinkhan, G. M., & Watson, R. T. (2006). The influence of home page complexity on consumer attention, attitudes and purchase intent. *Journal of Advertising*, 35, 69-80.

Gotlieb, J. B., Grewal, D., & Brown, S. W. (1994). Consumer satisfaction and perceived quality: Complementary or divergent constructs? *Journal of Applied Psychology*, 79, 875-885.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion*, 24, 175-213.

Haak, M. J. van den., Jong, M. D. T. de., & Schellens, P. J. (2004). Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: A methodological comparison. *Interacting with Computers*, 16, 1153-1170.

Halstead, M. (1977). *Elements of software science*. North-Holland, New York: Elsevier.

Hassenzahl, M. & Ullrich, D. (2007). To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, 19, 429-437.

Hassenzahl, M., 2004. The interplay of beauty, goodness and usability in interactive products. *Human-Computer Interaction*, 19, 319-349.

Hassenzahl, M. & Tractinsky, N. (2006). User experience: A research agenda. *Behaviour & Information Technology*, 25, 91-97.

Haugtvedt, C. P., Petty, R. E., & Cacioppo, J. T. (1992). Need for cognition and advertising: Understanding the role of personality variables in consumer behaviour. *Journal of Consumer Psychology*, 1, 239-260.

Haugtvedt, C. P. & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology*, 63, 308-319.

Heijden, H. van der. (2003). Factors influencing the usage of Web sites: The case of a generic portal in the Netherlands. *Information & Management,* 40, 541-549.

Heijden, H. van der. (2004). User acceptance of hedonic information systems. *MIS Quarterly,* 28, 695-704.

Heyman, G. D. & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion,* 16, 231-247.

Hong, E. & O'Neil, H. F. Jr. (2001). Construct validation of a trait self-regulation model. *International Journal of Psychology*, 36, 186-194.

Horvath, M., Herleman, H. A., & McKie, R. L. (2006).Goal orientation, task difficulty, and task interest: A multilevel analysis. *Motivation and Emotion,* 30, 171-178.

Ivory, M. Y. (2003). Automated Web site evaluation: Researchers' and practitioners' perspectives. *Human-Computer Interaction Series*, vol. 4. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ivory, M. Y., Sinha, R. R., & Hearst, M. A. (2001). Empirically validated Web page design metrics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 53-60.

Ivory, M. Y. & Megraw, R. (2005). Evaluation of Web site design patterns. *ACM Transactions on Information Systems,* 23, 463-497.

Ivory, M. Y., Mankoff, J., & Le, A. (2003). Using automated tools to improve Web site usage by users with diverse abilities. *IT and Society,* 1, 195-236.

http://www.stanford.edu/group/siqss/itandsociety/v01i03/v01i03a11.pdf.

Jaccard, J. & Turrisi, R. (2003). *Interaction effects in multiple regression*. Sage, Thousand Oaks, CA.

Jackson, S. A. & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport & Exercise Psychology,* 18, 17-35.

Jacobs, B., Prentice-Dunn, S., & Rogers, R. W. (1984). Understanding persistence: An interface of control theory and self-efficacy theory. *Basic and Applied Social Psychology,* 5, 333-347.

Jiang, Z. (J.) & Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers’ product understanding. *MIS Quarterly,* 31, 475-500.

Jonassen, D. H. & Grabowski, B. L. (1993). *Handbook of individual differences, learning and instruction*, Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Juvina, I. & Oostendorp, H. van. (2006). Individual differences and behavioural metrics involved in modeling Web navigation. *Universal Access Information Society*, 4, 258-269.

Kieras, D. & Polson, P. G. (1999). An approach to the formal analysis of user complexity. *International Journal of Human-Computer Studies,* 51, 405-434.

King, W. R. & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43, 740-755.

Kommers, P. (1991). Hypertext and the Acquisition of Knowledge, Navigation and Meta-Cognition through Conceptual Representations. University of Twente.

Kuiken, F. & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing,* 17, 48-60.

Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of Web sites. *International Journal of Human-Computer Studies*, 60, 269-298.

Lazonder, A. W., Biemans, H. J. A., & Wopereis, I. G. J. H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51, 576-581.

Lindenberg, S. (2001). Intrinsic motivation in a new light. *KYKLOS,* 54, 317-342.

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention Web designers: You have 50 milliseconds to make a good first impression. *Behaviour & Information Technology,* 25, 115-126.

Lindsay, P. & Norman, D. (1977). *Human information processing: An introduction to psychology.* Academic Press, New York.

Marakas, G.M., Yi, M.Y., & Johnson, R.D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. *Information Systems Research*, 9, 126–163.

March, J. & Simon, H. (1958). *Organizations*. New York: Wiley.

McCabe, T. J. (1976). A complexity measure. *IEEE Transactions on Software Engineering*, 4, 308-320.

McCabe, T. J. & Butler, C. W. (1989). Design complexity measurement and testing. *Communications of the ACM*, 32, 1415-1425.

Miles, J. & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London: Sage Publications.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Moneta, G. & Csikszentmihalyi, M. (1996). The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality*, 64, 275-310.

Nadkarni, S. & Gupta, R. (2007). A task-based model of perceived Web site complexity. *MIS Quarterly*, 31, 501-524.

Pirolli, P. (2003). A theory of information scent. In J. Jacko & C. Stephanidis (Eds.), Human-Computer Interaction (Vol. 1, pp. 213-217). Mahwah, NJ: Lawrence Erlbaum.

Pirolli, P. (1997). Computational models of information scent following in a very large browsable text collection. In S. Pemberton (Ed.), *Proceedings of the conference on human factors in computing systems* (CHI' 97) (pp. 3–10). Atlanta, GA. New York: ACM Press.

Pitkow, J. & Kehoe, C. (1996). Emerging trends in the WWW user population. *Communications of the ACM*, 39, 106-108.

Rao, V. R. & Steckel, J. H. (1998). *Analysis for strategic marketing.* Addison Wesley Longman, Inc.

Rauterberg, M. (1996). How to measure cognitive complexity in Human-Computer Interaction. *Proceedings of the Thirteenth European Meeting on Cybernetics and Systems Research*, 2, 815-820.

Rosenfeld, L. & Morville, P. (2002). Information architecture for the World Wide Web. O'Reilly.

Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist*, 55, 68-78.

Schaik, P. van. & Ling, J. (2003). Using on-line surveys to measure three key constructs of the quality of Human-Computer Interaction in Web sites: Psychometric properties and implications. *International Journal of Human-Computer Studies*, 59, 545-567.

Schaik, P. van. & Ling, J. (2005). The psychometric evaluation of educational intranets. *Journal of Educational Computing Research*, 33, 81-100.

Schaik, P. van. & Ling, J. (2007). Design parameters of rating scales for Web sites. *ACM Transactions on Computer-Human Interaction*, 14, 1-35.

Schaik, P. van. & Ling, J. (2008a). Modelling user experience with Web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers*.

Schaik, P. van. & Ling, J. (2008b). Immediate, initial and subsequent perceptions of aesthetics. *International Journal of Human - Computer Studies*.

Schedler, K. & Summermatter, L. (2007). Customer orientation in electronic government: Motives and effects. *Government Information Quarterly*, 24, 291-311.

Silcock, R. (2001). What is e-government? *Parliamentary Affairs*, 54, 88-101.

Tarasewich, P. (2005). An investigation into Web site design complexity and usability metrics. *Quarterly Journal of Electronic Commerce*.

Thompson, E. P, Chaiken, S., & Hazlewood, J. D. (1993). Need for cognition and desire for control as moderators of extrinsic reward effects: A person x situation approach to the study of intrinsic motivation. *Journal of Personality and Social Psychology*, 64, 987-999.

Tractinsky, N., Katz, A.S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13, 127–145.

Vakkari, P. (1999). Task complexity, information types, search strategies and relevance: Integrating studies on information seeking and retrieval. *Information Processing and Management*, 35, 819-837.

Venkatesh, V. & Davis, F. D. (1996). A model of perceived ease of use: Development and test. *Decision Sciences*, 27, 451-481.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27, 425-478.

Wang, L., Bretschneider, S., & Gant, J (2005). Evaluating Web-based e-government services with a citizen-centric approach. *Proceedings of the 38th Hawaii International Conference on System Sciences,* 1-10.

Weinstein, A. (2004). *Handbook of market segmentation: Strategic targeting for business and technology firms*. The United States of America: The Haworth Press.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behaviour and Human Decision Processes,* 37, 60-82.

Yi, M. Y. & Hwang, Y. (2003). Predicting the use of Web-based information systems: Self-efficacy, enjoyment, learning goal orientation and the technology acceptance model. *International Journal of Human-Computer Studies,* 59, 431-449.

Zhang, Y. & Buda, R. (1999). Moderating effects of need for cognition on responses to positively versus negatively framed advertising messages. *Journal of Advertising*, 28, 1-15.

Zijlstra, R. (1993). *Efficiency in work behaviour: A design approach for modern tools*, Delft, the Netherlands: Delft University Press, 1993.

Appendix

Appendix A: 11 Elements of Ivory et al. (2001)

Metric	**Description**
Word count	Total words on a page
Body text %	Percentage of words that are body vs. display text (i.e., headers)
Emphasized body text %	Portion of body text that is emphasized (e.g., bold, capitalized or near !'s)
Text positioning count	Changes in text position from flush left
Text cluster count	Text areas highlighted with color, bordered regions, rules or lists
Link count	Total links on a page
Page size	Total bytes for the page as well as elements graphics and style sheets
Graphic %	Percentage of page bytes that are for graphics
Graphics count	Total graphics on a page (not including graphics specified in scripts, applets

	and objects)
Colour count	Total colours employed

Appendix B: Information Retrieval Questions Used in the Experiment

Experiment Version 1: Low Artefact Complexity-Low Task Complexity

Ballymoney Borough Council

http://www.ballymoney.gov.uk/default.aspx

Practice tasks:

Task 1: What is the phone number of Ballymoney waste disposal service?

Answer: 028 2766 2408

(click 'w': http://www.ballymoney.gov.uk/w.aspx.)

Task 2: What is the total cost of purchasing a dog?

Answer: £16.75

(click 'environment': http://www.ballymoney.gov.uk/environment.aspx; click 'dog warden service': http://www.ballymoney.gov.uk/Dog_Warden_Service.aspx.)

Or:

(click 'your council': http://www.ballymoney.gov.uk/your_council.aspx; click 'council services': http://www.ballymoney.gov.uk/council_services.aspx; click 'dog warden service': http://www.ballymoney.gov.uk/Dog_Warden_Service.aspx.)

Task 3: What is the cost of marriage in a church?

Answer: £35.50

(click 'marriages': http://www.ballymoney.gov.uk/Marriages.aspx.)

Main tasks:

Task 1: What is the name of the forth company awarded for its maintenance service in the Annual Tenders list?

Answer: Painter Brian Lamont

(click 'tenders' at the home page: http://www.ballymoney.gov.uk/Tenders.aspx; click 'annual tenders': http://www.ballymoney.gov.uk/Annual_Tenders.aspx;.)

Task 2: How many Euros are there in the budget of the Northern Ireland Programme for Building Sustainable Prosperty?

Answer: 575 million Euros.

(click 'business': http://www.ballymoney.gov.uk/business.aspx; click 'local economic development programmes': http://www.ballymoney.gov.uk/Local_Economic_Development_Programmes.aspx.)

Task 3: What is the address of Bushvalley primary school?

Answer**:** 175 Ballinlea Road

(click 'community': http://www.ballymoney.gov.uk/community.aspx; click 'schools in Ballymoney': http://www.ballymoney.gov.uk/Schools.aspx.)

Or:

(click 'schools in Ballymoney' at home page: http://www.ballymoney.gov.uk/Schools.aspx.)

Task 4: How many times a year do the members of Ballymoney Borough Arts Committee get together?

Answer: 5

(click 'culture' at the home page: http://www.ballymoney.gov.uk/culture.aspx; click 'Ballymoney borough arts committee': http://www.ballymoney.gov.uk/ballymoney_borough_arts_committee.aspx.)

Task 5: On which day of July 2008 does the Health and Environmental Service Committee convene?

Answer: 22nd

(click 'timetable of meetings': http://www.ballymoney.gov.uk/Council_Meetings_Timetable.aspx.)

Task 6: How many members are there in the Audit Committee?

Answer: 8

(click 'committees' at home page: http://www.ballymoney.gov.uk/Committees.aspx; click 'audit committee': http://www.ballymoney.gov.uk/Audit_Committee.aspx.)

Task 7: How many priorities does the PEACE II programme implemented by Ballymoney Borough Local Strategy Partnership have?

Answer: 5

(click 'partnerships' at home page: http://www.ballymoney.gov.uk/Partnerships.aspx; click 'peace II': http://www.ballymoney.gov.uk/Peace_II.aspx.)

Task 8: How many fashion shops are there in Ballymoney?

Answer: 1

(click 'business directory' at home page: http://www.ballymoney.gov.uk/business-directory.aspx; click 'f': http://www.ballymoney.gov.uk/business_f.aspx.)

Task 9: How many semi-natural agriculture areas are there in Ballymoney?

Answer: 2

(click 'natural environment' at home page: http://www.ballymoney.gov.uk/natural_environment.aspx; click 'agriculture': http://www.ballymoney.gov.uk/Agriculture.aspx.)

Task 10: How many independent members are there in Ballymoney Council?

Answer: 1

(click 'your council': http://www.ballymoney.gov.uk/your_council.aspx; click 'council members': http://www.ballymoney.gov.uk/Council_Members.aspx.)

Experiment Version 2: High Artefact Complexity-Low Task Complexity

Manchester City Council

http://www.manchester.gov.uk/site/

Practice tasks:

Task 1: How many people live in Manchester?

Answer: Around 440,000

(click 'community and living': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100003.)

Task 2: How often is the newspaper Manchester People delivered to households across the city?

Answer: once every three months

(click 'your council': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100004; click 'Manchester People…': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100004&documentID=1754.)

Task 3: When did summer term start for schools in 2007/2008 in Manchester?

Answer: Monday 7 April 2008

(click 'term dates for Manchester schools' at home page: http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=1258.)

Main tasks:

Task 1: How many executive members are there among local councilors?

Answer: 6

(click 'your council': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100004; click 'your councillors': http://www.manchester.gov.uk/site/scripts/council_democracy_index.php.)

Task 2: What was the cost of Town Hall?

Answer: Around £1 million

(click 'your council': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100004; click 'town hall': http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=1986.)

Task 3: What was the top prize for Pride of Manchester in Community Awards 2007?

Answer: £750

(click 'community and living': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100003; click 'community awards 2007': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100003&documentID=3108.)

Task 4: What is the name of the cycle lights which are popular nowadays for their robustness and long battery life?

Answer: LED (Light emitting diode)

(click 'travel and roads': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100011; click 'safe cycling tips': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100011&documentID=3229.)

Task 5: What is the name of the service providing to find a home from a not-for-profit Manchester landlord?

Answer: Homeswap

(click 'housing': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100007; click 'rehousing: homefinder': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100007&documentID=294.)

Task 6: What is the phone number of Manchester environmental services centre?

Answer: 0161 954 9000

(click 'environment and planning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100006; click 'environment on call': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100006&documentID=526.)

Task 7: What is the phone number of the service for reporting abandoned vehicles?

Answer: 0161 872 5050

(click 'environment and planning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100006; click 'abandoned vehicles': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200001.)

Task 8: How many councillors are there in Manchester City Council?

Answer: 96

(click 'find your councillor': http://www.manchester.gov.uk/site/scripts/council_democracy_index.php.)

Task 9: What is the post code of contact address for Manchester City Council?

Answer: M60 2LA

(click 'contact us': http://www.manchester.gov.uk/contactus.)

Task 10: Is 'NVQ Level 2 Food Processing and Cooking' course available in Manchester School FEAST (Food Excellence and Skills Training) programme?

Answer: Yes

(click 'education and learning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100005; click 'Manchester FEAST':

http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100005&documentID=3580.)

Experiment Version 3: High Artefact Complexity-High Task Complexity

Practice tasks:

Task 1: What time does this year's Great Manchester Run take place?

Answer: 11:00-13:00

(click 'leisure, libraries and culture' link: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100009; click 'events': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100009&documentID=2552; click 'Manchester spring events': http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=3326; click 'mini and junior great Manchester run': http://www.manchester.gov.uk/site/scripts/events_info.php?eventID=890.)

Task 2: Does Withington Adult Learning Centre in South Manchester have any crèche service?

Answer: Yes

(click 'education and learning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100005; click 'Manchester adult education services': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100005&documentID=118; click 'find your nearest adult learning centre': http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=134; click 'South Manchester': http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=134&pageNumber=4; click 'Withington…': http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=442.)

Task 3: What is the name of the fort built by Roman soldiers in Castlefield location?

Answer: Mamucium

(click 'environment and planning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100006; click

‘conservation’:
http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200023; click ‘conservation areas’:
http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=511; click ‘Castlefield conservation area’:
http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=511&documentID=972; click ‘history’:
http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=511&documentID=972&pageNumber=2.)

Main tasks:

Task 1: What was the fifth item to be discussed at the meeting of the Health and Well-being Committee Thursday 6th March 2008?

Answer: Teenage pregnancy

(click ‘your council’:
http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100004; click ‘meetings and minutes’:
http://www.manchester.gov.uk/site/scripts/meetings_index.php; click ‘health and well-being…’:
http://www.manchester.gov.uk/site/scripts/meetings_committees.php?headerID=30; click ‘Thursday 6th march…’:
http://www.manchester.gov.uk/site/scripts/meetings_info.php?meetingID=658; click ‘agenda’: http://www.manchester.gov.uk/downloads/March_08.pdf.)

Task 2: What is the percent of authorised absence of King David High School students in 2006?

Answer: 6.2

(click ‘education and learning’:
http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100005; click ‘statistics on education’:
http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100005&documentID=2176; click ‘authorised/unauthorised absence 2006’:
http://www.manchester.gov.uk/site/scripts/download_info.php?downloadID=410&fileID=700; download file:
http://www.manchester.gov.uk/downloads/D4_Absence06_1_.pdf.)

Task 3: What is the phone number for reporting noise from another home as a type of pollution?

Answer: 0161 953 2525

(click 'business': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100002; click 'pollution': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200075; click 'pollution control-noise': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=412; click 'pollution control-noise': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=412&documentID=2956; click 'noise from another home': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=412&documentID=2956&pageNumber=2.)

Task 4: What is the cost of a Street Trading Licence in Manchester?

Answer: £625 per year

(click 'business': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100002; click 'business and street trading licences': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200063; click 'street trading in Manchester': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200063&documentID=1530; click 'street trading licence costs': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200063&documentID=1530&pageNumber=3.)

Task 5: What is the minimum penalty for depositing litter in the summary page of environmental crimes in Manchester?

Answer: £50

(click 'environment and planning': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100006; click 'environmental enforcement': http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=500012; click 'environmental crimes and legislation': http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=500012&

documentID=1028; click ‘summary page’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=500012&documentID=1036.)

Task 6: What is the phone number of Chorlton High school?

Answer: 0161 882 1150

(click ‘education and learning’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100005; click ‘Manchester school finder’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100005&documentID=2860; click ‘view the school finder’: http://www.manchester.gov.uk/site/custom_scripts/school_finder.php; click ‘click secondary schools’: http://www.manchester.gov.uk/site/custom_scripts/school_finder.php?schoolType=Secondary&page=1&recordsPerPage=10.)

Task 7: When must your Kerbit Twin Bins be left out on your recycling collection day?

Answer: At 7.30am

(click ‘community and living’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100003; click ‘recycling, rubbish and waste’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200084; click ‘kerbside recycling in Manchester’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200084&documentID=622; click ‘kerbit twin bins’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200084&documentID=622&pageNumber=2.)

Task 8: What is the phone number to be called in a domestic violence emergency?

Answer: 999

(click ‘community and living’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100003; click ‘domestic violence’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200036; click ‘domestic violence’:

http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200036&documentID=550; click ‘what to do in an emergency’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=200036&documentID=550&pageNumber=3

Task 9: What is the name of the second street within the Loxford Court (Hulme) resident parking zone?

Answer: Chervil Close

(click ‘environment and planning’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100006; click ‘parking’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=200072; click ‘street parking-permits’: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=474; click ‘resident’s permit’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=474&documentID=596; click ‘list of eligible streets-Loxford’: http://www.manchester.gov.uk/site/scripts/documents_info.php?documentID=596&pageNumber=6.)

Task 10: What is the code for checking bus times using a mobile phone?

Answer: 84268

(click ‘travel and roads’ at the home page: http://www.manchester.gov.uk/site/scripts/documents.php?categoryID=100011; click ‘greater Manchester passenger transport executive information’: http://www.manchester.gov.uk/site/scripts/documents_info.php?categoryID=100011&documentID=3438; click ‘Bus service and Timetable information’: http://www.gmpte.com/content.cfm?subcategory_id=104202; click ‘SMS press release’: http://www.gmpte.com/content.cfm?subcategory_id=103073&news_id=2368604.)

Appendix C: Questionnaires Used in the Experiment

Short Form of the Need for Cognition Scale of Cacioppo and Petty (1984)

Instructions: For each of the statements below, please indicate to what extent the statement is characteristic of you. If the statement is extremely uncharacteristic of you (not at all like you) please write a "1" to the left of the question; if the statement is extremely characteristic of you (very much like you) please write a "7" next to the question. Of course, a statement may be neither extremely uncharacteristic nor extremely characteristic of you; if so, please use the number in the middle of the scale that describes the best fit. Please keep the following scale in mind as you rate each of the statements below: 1 = extremely uncharacteristic; 4 = uncertain; 7 = extremely characteristic.

1	I would prefer complex to simple problems.
2	I like to have responsibility of handling a situation that requires a lot of thinking.
3	Thinking is not my idea of fun.*
4	I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.*
5	I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something.*
6	I find satisfaction in deliberating hard and for long hours.
7	I only think as hard as I have to.*
8	I prefer to think about small, daily projects to long-term ones.*
9	I like tests that require little thought once I've learned them.*
10	The idea of relying on thought to make my way to the top appeals to me.
11	I really enjoy a task that involves coming up with new solutions to problems.
12	Learning new ways to think doesn't excite me very much.*
13	I prefer my life to be filled with puzzles that I must solve.
14	The notion of thinking abstractly is appealing to me.
15	I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

16	I feel relief rather than satisfaction after completing a task that required a lot of mental effort.*
17	It is enough for me that something gets the job done; I don't care how or why it works.*
18	I usually end up deliberating about issues even when they do not affect me personally.

* Reverse scoring is used on this item.

Combined Self-Efficacy Scale

(Combination of two scales, namely computer self-efficacy scale of Compeau and Higgins (1995) and generalised self-efficacy scale of Hong and O'Neil (2001).

PART 1 (TRAIT-personality)

Self-efficacy scale of Hong and O'Neil (2001):

Instruction: Please answer the following questions, using a scale from 1 to 7, where 1 indicates "Not at all confident," 4 indicates "Moderately confident," and 7 indicates "Totally confident."

1. I always manage to solve difficult problems if I try hard enough. 1 2 3 4 5 6 7
2. If someone opposes me, I can find means and ways to get what I want. 1 2 3 4 5 6 7
3. It is easy for me to stick to my aims and accomplish my goals. 1 2 3 4 5 6 7
4. I am confident that I could deal efficiently with unexpected events. 1 2 3 4 5 6 7
5. Thanks to my resourcefulness, I know how to handle unforeseen situations. 1 2 3 4 5 6 7
6. I can solve most problems if I invest the necessary effort. 1 2 3 4 5 6 7
7. I remain calm when facing difficulties because I can 1 2 3 4 5 6 7

rely on my coping abilities.

8. When I am confronted with a problem, I usually find several solutions. 1 2 3 4 5 6 7

9. If I am in a bind, I can usually think of something to do. 1 2 3 4 5 6 7

10. No matter what comes my way, I'm usually able to handle it. 1 2 3 4 5 6 7

NO REVERSED SCORES

PART 2 (STATE-experience)

Self-efficacy scale of Compeau and Higgins (1995):

Instruction: Often in our jobs we are told about software packages that are available to make work easier. For the following questions, imagine that you were given a new software package for some aspect of your work. It doesn't matter specifically what this software package does, only that it is intended to make your job easier and that you have never used it before.

The following questions ask you to indicate whether you could use this unfamiliar software package under a variety of conditions. For each of the conditions, please rate your confidence in your ability to complete the job using the software package, by choosing a number from 1 to 7, where 1 indicates "Not at all confident," 4 indicates "Moderately confident," and 7 indicates "Totally confident".

I CAN PERFORM A CERTAIN TASK WITHIN A WEBSITE…

11. …If there is no one around to tell me what to do as I go. 1 2 3 4 5 6 7

12. … If I have never used a website like it before. 1 2 3 4 5 6 7

13. … if I had only the software manuals for reference. 1 2 3 4 5 6 7

14. … if I have seen someone else using the website before trying it myself. 1 2 3 4 5 6 7

15. … if I can call someone for help when I get stuck. 1 2 3 4 5 6 7

16. … if someone else has helped me get started. 1 2 3 4 5 6 7

17. … if I have a lot of time to perform the task in the website. 1 2 3 4 5 6 7

18. … if I have just the built-in help facility for assistance. 1 2 3 4 5 6 7

19. … if someone shows me how to use it first. 1 2 3 4 5 6 7

20. … if I have used similar websites before this one to perform the same task. 1 2 3 4 5 6 7

NO REVERSED SCORES

The Situational Motivation Scale (SIMS) of Guay et al. (2000)

Directions: Read each item carefully. Using the scale below, please circle the number that best describes the reason why you are currently engaged in this activity. Answer each item according to the following scale: 1: *corresponds not all*; 2: *corresponds a very little*; 3: *corresponds a little*; 4:*corresponds moderately*;

5: *corresponds enough*; 6: *corresponds a lot*; 7: *corresponds exactly*.

Why are you currently engaged in this activity?

1. Because I think that this activity is interesting 1 2 3 4 5 6 7
2. Because I am doing it for my own good 1 2 3 4 5 6 7
3. Because I am supposed to do it 1 2 3 4 5 6 7
4. There may be good reasons to do this activity, but personally I don't see any 1 2 3 4 5 6 7
5. Because I think that this activity is pleasant 1 2 3 4 5 6 7
6. Because I think that this activity is good for me 1 2 3 4 5 6 7
7. Because it is something that I have to do 1 2 3 4 5 6 7
8. I do this activity but I am not sure if it is worth it 1 2 3 4 5 6 7
9. Because this activity is fun 1 2 3 4 5 6 7
10. By personal decision 1 2 3 4 5 6 7

11. Because I don’t have any choice 1 2 3 4 5 6 7

12. I don’t know; I don’t see what this activity brings me 1 2 3 4 5 6 7

13. Because I feel good when doing this activity 1 2 3 4 5 6 7

14. Because I believe that this activity is important for me 1 2 3 4 5 6 7

15. Because I feel that I have to do it 1 2 3 4 5 6 7

16. I do this activity, but I am not sure it is a good thing to pursue it 1 2 3 4 5 6 7

Codification key: Intrinsic motivation: Items 1, 5, 9, 13; Identified regulation: Items 2, 6, 10, 14; External regulation: Items 3,7, 11, 15; Amotivation: Items 4, 8, 12, 16.

NO REVERSED SCORES

Disorientation scale of Ahuja and Webster (2001)

I felt lost

I felt like I was going around in circles

It was difficult to find a page that I had previously viewed

Navigation between pages was a problem

I didn’t know how to get to my desired location

I felt disoriented

After browsing for a while I had no idea where to go next

NO REVERSED SCORES

Adapted perceived ease of use and usefulness scale of Davis (1989)

Perceived Ease of Use
1. Learning to use the site was easy.
2. Becoming skillful at using the site was easy.
3. The site was easy to navigate.

Perceived Usefulness
1. Using the site would improve my effectiveness in finding information about my local council.
2. Using the site would improve my productivity in finding information about my local council.
3. I would find the site useful for finding information about my local council.
4. Using the site would improve my performance in finding information about my local council.

NO REVERSED SCORES

The Flow State Scale of Jackson and Marsh (1996)

Please answer the following questions in relation to your experience in the event you have just completed. These questions relate to the thoughts and feelings you may have experienced during the event. There are no right or wrong answers. Think about how you felt during the event and answer the questions using the rating scale below. Circle the number that best matches your experience from the options to the right of each question.

Rating Scale: (From 1 to 7)

Strongly disagree	Neither agree nor disagree	Strongly Agree
1	4	7

1. I was challenged, but I believed my skills would allow me to meet the challenge.

2. I made the correct movements without thinking about trying to do so.

3. I knew clearly what I wanted to do.

4. It was really clear to me that I was doing well.

5. My attention was focused entirely on what I was doing.

6. I felt in total control of what I was doing.
7. I was not concerned with what others may have been thinking of me.
8. Time seemed to alter (either slowed down or speeded up).
9. I really enjoyed the experience.
10. My abilities matched the high challenge of the situation.
11. Things just seemed to be happening automatically.
12. I had a strong sense of what I wanted to do.
13. I was aware of how well I was performing.
14. It was no effort to keep my mind on what was happening.
15. I felt like I could control what I was doing.
16. I was not worried about my performance during the event.
17. The way time passed seemed to be different from normal.
18. I loved the feeling of that performance and want to capture it again.
19. I felt I was competent enough to meet the high demands of the situation.
20. I performed automatically.
21. I knew what I wanted to achieve.
22. I had a good idea while I was performing about how well I was doing.
23. I had total concentration.
24. I had a feeling of total control.
25. I was not concerned with how I was presenting myself.
26. It felt like time stopped while I was performing.
27. The experience left me feeling great.
28. The challenge and my skills were at an equally high level.
29. I did things spontaneously and automatically without having to think.
30. My goals were clearly defined.
31. I could tell by the way I was performing how well I was doing.

32. 1 was completely focused on the task at hand.

33. I felt in total control of my body.

34. I was not worried about what others may have been thinking of me.

35. At times, it almost seemed like things were happening in slow motion.

36. I found the experience extremely rewarding.

NO REVERSED SCORES

Sub-scales of Flow State Scale:

1. Challenge-skill balance
Q1

Q10

Q19

Q28

2. Action-awareness merging
Q2

Q11

Q20

Q29

3. Clear goals
Q3

Q12

Q21

Q30

4. Unambiguous feedback
Q4

Q13

Q22

Q31

5. Concentration on task at hand

Q5

Q14

Q23

Q32

6. Paradox of control
Q6

Q15

Q24

Q33

7. Loss of self-conscious
Q7

Q16

Q25

Q34

8. Transformation of time
Q8

Q17

Q26

Q35

9. Autotelic experience
Q9

Q18

Q27

Q36

Perceived Behavioural Intention Scale in Venkatesh et al. (2003)

Perceived behavioral intention to use the system
1. I would intend to use the site for finding

information about my local council.
2. I predict that I would use the site for finding information about my local council.
3. I plan to use the site for finding information about my local council.

NO REVERSED SCORES

Perceptions of aesthetics scale of Lavie and Tractinsky (2004)

1. Classic aesthetics:

10. Aesthetic

8. Pleasant

4. Clear

2. Clean

9. Symmetrical

2. Expressive aesthetics:

3. Sophisticated

5. Fascinating

6. Creative

7. Uses special effects

1. Original

NO REVERSED SCORES

Appendix D: Instructions to Think-Aloud Sessions

In this test, we are interested in what you think about when you perform some tasks with a web site that I am going to ask you to do. Please note that the site is being tested, not you. In order to do this, I am going to ask you to THINK ALOUD as you perform the tasks. What I mean by 'think aloud' is that I want you to tell me EVERYTHING you are thinking from the time you first see the question until you finished a task or until I tell you to stop working on the task. I don't want you to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. It is most important that you keep talking. If you are silent for any long period of time, I will ask you to talk. Please try to speak as clearly as possible, as I shall be recording you as you speak. Do you understand what I want you to do?