A cognitive-experiential approach to modelling web navigation

Paul van Schaik\textsuperscript{a,},*, Jonathan Ling\textsuperscript{b}

\textsuperscript{a}Teesside University, School of Social Sciences and Law, Psychology subject group, Borough Road, Middlesbrough, Cleveland TS1 3BA, United Kingdom
\textsuperscript{b}University of Sunderland, Sunderland, United Kingdom

Received 17 August 2011; received in revised form 28 March 2012; accepted 31 March 2012
Available online 9 April 2012

Abstract

Flow experience, the degree to which a person feels involved in a particular activity, is an important influence on human–computer interaction. Building on Guo and Poole’s (2009) model of flow experience in Web navigation, and van Schaik and Ling’s (in press) cognitive-experiential approach to modelling interaction experience, this research demonstrates the crucial role of the preconditions of flow experience in human–computer interaction. In an experiment, the preconditions of flow experience – but not flow experience proper – mediated the effects of artefact complexity, task complexity and intrinsic motivation (as a situation-specific trait) on both flow and task outcome. However, preconditions did not predict overall artefact evaluation. Within a staged model of flow experience, the broader implications of this work for human–computer interaction are explored.

© 2012 Elsevier Ltd. All rights reserved.

Keywords: User-experience; Model; Flow experience; Cognition; Partial least squares

1. Introduction

1.1. Experience and task performance

Recently, there has been an apparent shift in human–computer interaction from a focus on cognitive-task performance to interaction experience (‘user-experience’), both in academic research (e.g., Law and van Schaik, 2010) and in engagement with practitioners (e.g., User Experience White Paper, http://www.allaboutux.org/uxwhitepaper). Attention to users’ interaction experience is useful in broadening both the scientific understanding of people’s interaction with computers and the concerns addressed by the process of designing interactive computer systems. Cognitive-task performance remains an active area of research (e.g., Stone and Dennis, 2011). However, task performance and experience are usually studied separately, without a concern for how they may be related. Indeed, it is important to note that the major published models of interaction experience are silent about this relationship (Desmet and Hekkert’s (2007) framework of product experience, Hartmann et al.’s (2008) model of users’ decision-making process for user-interface quality assessment, Hassenzahl’s (2003) user-experience model, Porat and Tractinsky’s (In press) environmental-psychology model and Thüring and Mahlke’s (2007) components-of-user-experience model). Recent work has, nevertheless, demonstrated that experience and task performance are not independent. For instance, a design with enhanced aesthetics can improve cognitive-task performance with a website under conditions of poor usability (Moshagen et al., 2009; see also Sonderegger and Sauer, 2010). Therefore, the current study set out to investigate in detail how one particular type of experience – flow – can influence task performance. In order to appreciate its role in human–computer interaction it is important to review some of the ideas underlying the concept of flow.

1.2. Flow experience and dimensions

Csikszentmihalyi (1988) distinguishes three functional subsystems of consciousness: attention (taking note of external or internal information), awareness (interpreting information) and memory (storing information). In this view, the content of consciousness is experience, conceptualised as “the sum of all the information that enters it,
and its interpretation by awareness” (Csikszentmihalyi, 1988, p. 17). Three of the most important processes in awareness are thought/cognition (recognising pieces of information and relating these to each other), feeling/emotion (valence towards the information that is being processed, which can be positive [like] or negative [dislike]), and conation/volition (keeping attention focused on a particular range of stimuli rather than being diverted to others). It is notable that, because all these processes need attention in order to work, “they are also ‘information’ in consciousness” (Csikszentmihalyi, 1988, p. 19). Therefore, they are subject to the same processing limitations that apply to information that originates outside of awareness (such as a maximum of seven chunks of information per unit of time; Miller, 1956). These processes compete for a relatively small number of stimuli at any one time. A crucial distinction is made between psychic entropy, where there is a conflict between the contents of consciousness and goals developed by the self, and negentropy/optimal experience/flow, where there is harmony among the contents of consciousness and with goals. During psychic entropy, there is “noise” in the information-processing system, accompanied by reduced efficiency — because attention is diverted from other tasks to attend to conflicting information — and there are negative experiences, depending on the particular information and the goals with which it conflicts (e.g., fear or boredom). Conversely, characteristics accompanying flow are increased efficiency and positive experience (e.g., pleasure, happiness, satisfaction and/or enjoyment).

Flow experience has been studied in many domains such as sport (Stavrou et al., 2007), education (Engeser and Rheinberg, 2008), computer gaming (Murphy et al., 2008) and human–computer interaction (e.g., P. van Schaik, J. Ling). Nine dimensions of flow (see Table 1) have been distinguished (Csikszentmihalyi, 1988) and measurement instruments for these dimensions have been developed and validated (e.g., Jackson and Marsh, 1996; Jackson and Eklund, 2002).

Flow is not a matter of ‘all or nothing’, as people can experience a degree of flow on each of the dimensions. Guo and Poole (2009) and van Schaik and Ling (in press) both measured the nine dimensions of flow experience in human–computer interaction. By conceptualising and measuring flow comprehensively, these studies represent an advance on previous work which employed incomplete models of flow — making it hard to generalise findings, elucidate the mechanisms that influence flow in human–computer interaction, or clarify how flow influences outcomes of this interaction (Guo and Poole, 2009).

### 1.3. Staged model of flow experience

Little research in human–computer interaction has considered the distinction between the preconditions of flow (the dimensions of balance of challenge and skill, clarity of goals and feedback) and its consequent: flow proper (the dimensions of concentration, control, mergence of action and awareness, transcendence of self, transformation of time and autotelic experience). However, Guo and Poole (2009) developed a staged model of flow experience, where the preconditions of flow influence flow proper, in order to develop a better understanding of the effects of other variables on flow experience. According to the theory of flow (Csikszentmihalyi and Csikszentmihalyi, 1988), the effect of situational factors on flow experience proper occurs as a consequence of their effect on the preconditions of flow. Guo and Poole (2009) examined the effect of website complexity on flow, mediated by its preconditions. They found that the negative effect of website complexity was partially mediated by preconditions of flow, but the study suffered from some limitations. First, complexity was not experimentally controlled. Rather, websites were selected that were less complex or more complex. Thus, numerous artefact characteristics may have been confounded with the effect of complexity, posing potential threats to internal validity. Second, only the effect of perceived complexity was analysed, but not the effect of actual site complexity on the preconditions of flow and flow proper.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance of challenge and skill</td>
<td>“The person perceives a balance between the challenges of a situation and one’s skills, with both operating at a personally high level.” (p. 18)</td>
</tr>
<tr>
<td>Mergence of action and awareness</td>
<td>“The flow activity is so deep that it becomes spontaneous or automatic.” (p. 18)</td>
</tr>
<tr>
<td>Goal clarity</td>
<td>“Goals in the activity are clearly defined (…), giving the person in flow a strong sense of what he or she is going to do.” (p. 19)</td>
</tr>
<tr>
<td>Feedback</td>
<td>“Immediate and clear feedback is received, usually from the activity itself, allowing the person to know he or she is succeeding in the set goal.” (p. 19)</td>
</tr>
<tr>
<td>Concentration</td>
<td>“Total concentration on the task at hand occurs when in flow” (p. 19)</td>
</tr>
<tr>
<td>Control</td>
<td>“A sense of exercising control is experienced, without the person actively trying to exert control.” (p. 19)</td>
</tr>
<tr>
<td>Loss of self-consciousness</td>
<td>“A sense of exercising control is experienced, without the person actively trying to exert control.” (p. 19)</td>
</tr>
<tr>
<td>Transformation of time</td>
<td>“Time alters perceptibly, either slowing down or speeding up” (p. 19)</td>
</tr>
<tr>
<td>Autotelic experience</td>
<td>“Intrinsically rewarding experience. An activity is autotelic if it is done for its own sake, with no expectation of some future reward or benefit.” (p. 20)</td>
</tr>
</tbody>
</table>
Therefore, it remains unknown whether actual complexity had an effect on perceived complexity, the preconditions of flow and flow itself. Third, Guo and Poole (2009) focused only on antecedents of flow and not on its consequents.

### 1.4. The need for 'anchoring' flow in objective outcomes

Significant gains in the study of experience (e.g., flow) would be made both theoretically and in practical terms if it could be demonstrated that changes in experience are reliably associated with objective 'anchors' such as health outcomes or the outcomes of task performance as a result of this experience (Sheldon et al., 2001). Apart from studying flow experience for its own sake, research in human–computer interaction has used non-experimental designs – with consequent threats to internal validity – to study flow experience mainly in relation to its influence on social-cognition variables such as evaluative judgments in the form of attitudes (Agarwal and Karahanna, 2000) and behavioural outcomes (Koufaris, 2002). However, although it has been demonstrated that flow experience is an independent positive predictor of task outcome (after controlling for other [cognitive] variables) in the domains of sport (Stavrou et al., 2007), education (Engeser and Rheinberg, 2008; Vollmeyer and Imhof, 2007), and computer-game playing (Engeser and Rheinberg; Vollmeyer and Rheinberg, 2003; Murphy et al., 2008), there is a lack of research in human–computer interaction studying flow in relation to cognitive-task outcomes.

### 1.5. Flow as a mediator of person-, artefact- and task characteristics on task performance

Finneran and Zhang’s (2003) person–artefact–task framework (see Fig. 1) was originally proposed for modelling flow experience in human–computer interaction. Building on the finding that flow is an independent predictor of task performance and using this framework (see Fig. 1), van Schaik and Ling (in press) studied the effect of experimentally manipulated artefact complexity and task complexity on flow experience and task performance in web navigation.

Van Schaik and Ling employed the framework to model people’s navigation of websites where a computer user navigates a website by visiting web pages in order to complete a task. In this framework, characteristics of a particular artefact (e.g., the complexity of a website), the task performed with the artefact (e.g., the complexity of a task) and the person (end-user) performing the task (e.g., a user’s spatial ability) all influence people’s (flow) experience (of navigating) and thereby the outcome of their task performance. Finneran and Zhang note that the effects of person, artefact and task on the process are not necessarily independent. Therefore, it is important to consider not only the main effects of these three entities separately, but also possible interaction effects, where this is theoretically justified and among artefact- and task characteristics that are experimentally manipulated. In van Schaik and Ling’s study, the main effects of the manipulations of artefact complexity and task complexity on flow were confirmed. They also demonstrated that flow was a mediator of the effect of experimental manipulations on task performance, and task performance was a mediator of the effect of flow on task outcome. They concluded that their results demonstrated the need for taking an integrated cognitive-experiential approach in the modelling of human–computer interaction.

However, a closer analysis of van Schaik and Ling (in press) work shows the following limitations. First, the modelling of flow experience was undifferentiated: the nine dimensions of flow were used as indicators of one higher-order construct of flow and a differentiation between the preconditions of flow and flow proper was not made. Consequently, the effect of experimental manipulations, through preconditions, on flow was not considered, let alone tested. Second, the measurement of flow experience was borrowed from the domain of sport (Flow State Scale, Jackson and Marsh, 1996). Although, given the wording of the items, the instrument appears to be applicable to the measurement of flow in different domains, the use of a comprehensive instrument (accounting for all nine dimensions of flow) that is more specific for measuring flow in human–computer interaction would be more appropriate. Third, the measurement of flow experience as a higher-order construct was ad hoc in that a subset of flow dimensions was selected based on statistical rather than theoretical grounds. Fourth, a single-variable measure (correctly completed tasks) was used for task outcome. Therefore, reliability, and convergent and divergent validity of task outcome could not be assessed. The use of multiple indicators would therefore strengthen the measurement.

### 1.6. Current study

Despite lacking an experimentally controlled manipulation of website complexity and not accounting for the effect of flow on further outcomes of human–computer interaction, Guo and Poole (2009) nevertheless provided empirical evidence for a theoretically justified staged model

![Fig. 1. Person-artefact-task model (adapted from Finneran and Zhang, 2003).](image-url)
of flow in human–computer interaction. Although they did not account for the preconditions of flow, and issues existed in their measurement of flow and task performance, van Schaik and Ling (in press) demonstrated that in Finneran and Zhang’s (2003) person-artefact-task framework flow is a mediator of the effect of artefact- and task complexity on task outcome. The aim of the current study was therefore to advance knowledge of the relationship between experience and task performance in human–computer interaction, while addressing the limitations of these studies. This was done by constructing and testing a cognitive-experiential model of flow experience within the person-artefact-task framework, with a staged conceptualisation of flow experience, measurement of flow that is specific to human–computer interaction and multiple measurements of task outcome. Specifically, the current study investigated the effects of the artefact characteristic of website complexity, the task characteristic of task difficulty and the person characteristic of intrinsic motivation (as a situation-specific trait) on the preconditions of flow, flow proper and task outcome of the navigation of an information-oriented website.

1.7. Development of hypotheses

1.7.1. Artefact complexity

One source of artefact complexity is the complexity of the navigation structure of a website, which can be reflected in the number of options that are available for selection on each page.\(^1\) Empirical support from recent research for the effect of artefact complexity shows that page complexity in terms of the number of navigation choices on a web page (Gwizdka and Spence, 2006) and structural complexity (Guo and Poole, 2009) can increase task difficulty. Moreover, Guo and Poole confirmed that perceived artefact complexity decreases flow experience and its preconditions. In further empirical evidence for the negative effect of complexity, Blackmon et al. (2002) found that the greater the number of links per page the lower the success rate (in terms of percentage correct links clicked on first click).

In this context, theoretical support for the effect of artefact complexity comes from Pierce et al.’s (1992) dual-criterion model, based on semantic memory models. This model can account for a decrease in the accuracy of option (i.e., link) selection with an increasing number of options (links). The model uses two criteria: high (H) and low (L). Comparisons between a target and criterion producing a value below L will immediately be rejected and those producing a value above H will immediately be selected, leading to a self-terminating search. Comparisons with values between L and H will produce candidates that are considered in case the search was not self-terminating. If the comparison process produces one candidate then the corresponding option will be chosen. If there are more candidates then (a subset of) these will be re-examined and a partially redundant search occurs. In the model, the mechanism through which an increasing number of options produces a larger set of candidates is a decrease in the value of H. This larger set of competing candidates increases the chance of incorrect selection. Applied to a hierarchical menu or a website, the chance of incorrect selection would be increased at each level of the hierarchy. Furthermore, as artefact complexity increases, the task of finding information will become more difficult, negatively affecting the balance of challenge and skill, and – because of additional extraneous presented information – making it more difficult to form and focus on goals and making it harder to perceive feedback and the consequences of acting on feedback; finally, consequently flow proper will decrease (Guo and Poole, 2009). Therefore, H1a: artefact complexity (page complexity) has a negative effect on task outcome.

H1b: artefact complexity (page complexity) has a negative effect on the preconditions of flow (balance of challenge and skill, clarity of goals and feedback).

H1c: artefact complexity (page complexity) has a negative effect on flow experience.

1.7.2. Task complexity

One source of task complexity follows from the navigation of a series of pages in a website. Specifically, this complexity is situated in the length of the path from the starting point (the home page of a website) to the destination (the page containing the required information to complete the task being performed). This has been identified as a factor that can decrease the quality of human–computer interaction (Gwizdka and Spence, 2006). Indeed, van Oostendorp et al.’s (2009) findings show that task complexity (path length, defined as the number of steps involved in finding the information) has a negative effect on task performance. Following Gwizdka and Spence (2006), two mechanisms for this effect can be distinguished. First, for a particular task goal and a given probability of selecting the correct link on each page (by identifying a semantic match between task goal and link content/text) on the path to the page containing the target information, a longer path to the target would result in a smaller probability of following a path consisting of only correct links. In other words, the longer the task path, the greater the chance of making at least one selection error and following an incorrect path. Second, a similar argument goes for the probability of making correct relevance judgments about the information presented in each of a sequence of web pages along the path. A longer path would lead to a greater chance of making an incorrect
relevance judgment, either by judging a relevant piece of information to be irrelevant for the task goal that is being pursued or vice versa. Therefore, task performance will decrease with path length. In addition, as task complexity increases the balance of challenge and skill, performance will be adversely affected, and – because the semantic match of page content with the task goal is inversely proportional to the serial position of a page on the path to the target page, – forming and focusing on goals and perceiving feedback and the consequences of acting on it become harder; finally, consequently flow experience will decrease (Guo and Poole, 2009). Thus,

**H2a:** task complexity (path length) has a negative effect on task outcome.

**H2b:** task complexity (path length) has a negative effect on the preconditions of flow (balance of challenge and skill, clarity of goals and feedback).

**H2c:** task complexity (path length) has a negative effect on flow experience.

### 1.7.3. Intrinsic motivation

Various individual-difference variables in persons (end-users) can have an effect on their navigation of web-based systems. For instance, Juvina and van Oostendorp (2006) demonstrated that spatial ability and domain expertise are positive predictors of task outcome (in terms of effectiveness defined as a combination of correctness and completeness). In the current study, we examine intrinsic motivation (“the inherent tendency to seek out novelty and challenges, to extend and exercise one’s capacities, to explore, and to learn” — Ryan and Deci, 2000, p. 70) as a person-characteristic (situation-specific trait) in web navigation. Intrinsic motivation in this sense is a positive predictor of task outcome in academic learning (Hirschfeld et al., 2008; Vansteenkiste et al., 2008) and flow experience in athletics (Stavrou et al., 2007). Theoretically, there are several reasons for this positive effect of intrinsic motivation on task performance (Zapata-Phelan et al., 2009). First, activity, concentration, initiative, resilience and flexibility can increase task performance. Second, intrinsic motivation has a stronger effect than external motivation on the persistence of effort. This, in turn, has a strong positive effect on the performance of tasks with artefacts. Third, internal motivation is expected to have a positive effect on task performance in the domain of employment. Fourth, those with higher intrinsic motivation are expected to focus their attention to a larger degree on a particular activity that they are performing, thereby enhancing the preconditions of flow (the attention-enhancing cognitive component of flow; see Section 1.7.4) a higher level of flow experience than others (Asakawa, 2004). Therefore,

**H3a:** intrinsic motivation has a positive effect on task outcome.

**H3b:** intrinsic motivation has a positive effect on the preconditions of flow experience.

**H3c:** intrinsic motivation has a positive effect on flow experience.

### 1.7.4. Flow experience and task outcome

Two pathways for the positive effect of flow on task outcome can be distinguished. First, flow is considered to be a ‘highly functional state’ (Engeser and Rheinberg, 2008, p. 158); consequently, it should facilitate task performance. In Csikszentmihalyi and Nakamura’s (2010) conceptualisation, the basis for this pathway could be seen as the function of the preconditions of flow — facilitating effortless attention, and thereby cognition and task performance. In particular, if the precondition of balance of challenge and skill is met then “every additional investment of attention can have the most immediate effect” (p. 187). When there is clarity of goals, each subsequent action in the action sequence presents a clear goal for the next step until the goal is reached, and thereby facilitates effortless attention. Immediate feedback facilitates effortless attention by sustaining attention (rather than attention having to be diverted/expended to seek feedback). Therefore, preconditions of flow can be considered as the attention-enhancing cognitive component of flow experience, which therefore should facilitate cognitive-task performance as an ‘enabler’.

Second, flow is a driver of motivation for continued activity, supporting processes of volition/conation; this leads people to select higher challenges in order to experience flow again. Therefore, flow proper can be considered as the motivational component of flow experience, which therefore should facilitate cognitive-task performance as a (motivational) ‘driver’. A similar argument, regarding the role of motivation during people’s interaction with an artefact, with supporting empirical evidence, is made by David et al. (2007). In their ‘information seek cycle’, as a result of the level of self-efficacy (rather than flow) from previous information- seek cycles, more challenging goals are formulated in subsequent cycles. Although Engeser and Rheinberg (2008) do not explicitly distinguish between flow proper and its preconditions, it appears that the first pathway from flow to task outcome applies to the preconditions of flow and the second to flow proper. This conclusion is a further justification for the conceptual distinction (also made by Guo and Poole, 2009) between the two higher-order constructs: flow proper and its preconditions. Thus,

**H4:** the preconditions of flow experience have a positive effect on task outcome, while person-, artefact- and task
factors (artefact complexity, task complexity and intrinsic motivation) are held constant.

H5: the preconditions of flow experience have a positive effect on flow experience, while person-, artefact- and task factors (artefact complexity, task complexity and intrinsic motivation) are held constant.

H6: flow experience has a positive effect on task outcome, while person-, artefact- and task factors (artefact complexity, task complexity and intrinsic motivation) and the preconditions of flow experience are held constant.

1.8. Overview of experiment

The current study tests the hypotheses using a computer-controlled experiment, in which artefact complexity and task complexity were manipulated; test-users’ intrinsic motivation was measured as an individual-difference variable. Furthermore, given that previous research has focused particularly on users’ evaluative judgments (of an artefact or the interaction with an artefact) rather than task outcome (which is the focus of attention here) as a consequence of flow experience, the results for task outcome were contrasted with those for an overall evaluative measure of artefact quality (‘goodness’; Hassenzahl, 2004) to explore how preconditions and flow differ in their effects on task outcome and evaluative judgment. In the experiment, users employed an information-oriented realistic mock intranet site to perform a series of information retrieval tasks.

2. Method

2.1. Design

A 2 × 2 between-subjects experimental design was used with the independent variables of artefact complexity (high and low) and task complexity (high and low). Low artefact complexity was defined as five links and high complexity as ten links on a web page. In a low-complexity task the answer to an information-retrieval task was available on a page two links from the homepage and in a high-complexity task the answer was available on a page four links from the homepage. In addition, the situation-specific trait of intrinsic motivation was measured by questionnaire (see Section 2.3) to establish its effect on flow experience and task outcome.

Dependent variables were the preconditions of flow, flow and task outcome as well as perceived task challenge, perceived artefact complexity and goodness. All variables, except task outcome, were measured by questionnaire (see Section 2.3). A reflective measure of task outcome was used with the following indicators: percentage of completed tasks (an indicator of speed), percentage of correct answers relative to the total number of tasks (completed or not) and percentage of correctly completed tasks (both indicators of efficiency, where the first is a more conservative indicator than the second).

2.2. Participants

One hundred and twenty-seven undergraduate psychology students (102 females and 25 males), with a mean age of 22.87 years (SD = 7.40) took part in the experiment as a course requirement. All participants had used the Web, but none had used the mock intranet site. Mean experience using the Web was 10.61 years (SD = 3.11), mean time per week spent using the Web was 17.22 h (SD = 13.74) and mean frequency of Web use was 15.37 times per week (SD = 10.02).

2.3. Materials and equipment

Questionnaires. Participants gave responses to several questionnaires, using 7-point scales (see Appendix). Guo and Poole’s (2009) 30-item flow experience scale (FES) measured the preconditions of flow experience (with dimensions of balance of challenge and skill, clarity of goals and feedback) and flow experience proper (with dimensions of concentration, perceived control, merger of action and awareness, transformation of time, transcendence of self and autotelic experience). ‘Shopping’ in Item 9 and ‘surfing’ in Item 10 were changed to ‘using the website’ because the task was not shopping and in order to avoid jargon, respectively. Four items from Guay et al. (2000) 16-item Situational Motivation Scale (SIMS) measured intrinsic motivation. Perceived task challenge was measured using a single item (PTC). Perceived artefact complexity was measured using Guo and Poole’s (2009) adaptation of Nadkarni and Gupta’s (2007) perceived website complexity scale (PCS).

Website. Each participant used one of two versions of a website that was modelled as a typical psychology site for university students, and specially designed for the experiment. In addition to the homepage, the main pages of the high-complexity version (see Fig. 2) of the site were Teaching, Research, Fees and Funding, Hall of Fame, Library, Staff, Sports and Leisure, Careers and About, with 9990 further Web pages. In addition to the homepage, the main pages of the low-complexity version (see Fig. 2) of the site were Teaching, Research, Fees and Funding, and Hall of Fame, with 620 further Web pages. All links and content of the low-complexity version were also included in the high-complexity version. The complex website had more pages than the simple website, but both sites had an equal number of four levels of depth (from the homepage) and therefore both site versions allowed simple and complex tasks to be completed.

Equipment. The experiment was programmed using Visual Basic 6.0 and ran on personal computers (Intel Pentium, 1.86 GHz, 2 GB RAM, Microsoft Windows XP operating system, 17-inch monitors). The screen dimensions...
were 1280 × 1024. Contrast (50%) and brightness (75%) were set to optimal levels.

2.4. Procedure

The experiment ran in a computer laboratory with groups of 15–20 participants who worked independently. The study consisted of three phases. In Phase 1, participants completed the SIMS. In Phase 2, an information retrieval task followed which included typical tasks that users perform with educational intranet sites. In each trial, a question appeared at the top of the screen, for instance ‘How long is the exam for the module Introduction to Personality and Social Psychology?’ Once participants had

Fig. 2. Website versions: (A) complex and (B) simple.
read the question, they had to click on a button labelled 'Show website'. The home page of the site then appeared on the screen and they had to find the answer to the question (which remained visible) using the site. Participants were told to take the most direct route possible to locate the answer. Having found this, they clicked on a button labelled 'Your answer', which opened a dialogue box at the bottom of the screen. Participants typed their answers into the box, clicked on 'OK' and moved on to the next question. After three practice questions, the main set of information retrieval tasks followed, with a duration of 20 min — in which a maximum of 37 further questions were presented. In Phase 3, participants completed the FES, the PTC and the PCS. Finally, participants answered questions requesting demographic details. The experiment took about 35 min to complete.

3. Results and discussion

First, the findings regarding psychometric properties of the measurement instruments are presented to establish the quality of measurement. Then, descriptive statistics and effect sizes are presented to show the effects of the experimental manipulations on the dependent variables and inferential statistics are presented to test these effects and the proposed hypotheses.\(^2\)

\(^2\)Essential in these analyses is evidence for the relationships in the model that support the hypotheses rather than the level of flow that is experienced. Indeed, variance in flow experience is necessary as a prerequisite for results that demonstrate this relationship. This would be impossible in the case of a ceiling effect (an extremely high level of flow, with insufficient or no variance) or a floor effect (an extremely low level of flow).

3.1. Psychometric properties of measurement Instruments

In testing the measurement model, reliability was analysed (see Table 2), and convergent and discriminant validity was assessed (see Table 3). The reliability of each individual reflective item is assessed by its loading on the construct of which it is an indicator, which should be 0.7 or higher (Henseler et al., 2009). More useful, in terms of the interpretation of results, is Chin’s (2010) recommendation to examine the squared loading because it represents the proportion of overlap in variance between an item and any construct.

Using a bootstrapping procedure, the loadings of all items were found to be statistically significant and there were no cross-loadings on any item across all scales. Although four items (SIMS-Item 1 and PCS-Items 1, 5 and 8) had loadings below the cut-off point, they were retained, given that (a) there were no cross-loadings, (b) the loadings on these four items were significant, (c) the scales have been previously validated and (d) in PLS including weaker items helps “to extract what useful information is available in the indicator to create a better construct score” (Barroso et al., 2010, p. 433), where weaker items contribute to construct scores with a lower weight. At the construct level, reliability was analysed using the composite reliability co-efficient, which needs to be 0.7 or higher. All the co-efficients exceeded this cut-off point.

Convergent validity – the extent of consistency among the items measuring a particular construct – was analysed using the average variance extracted (AVE) by a construct from its indicators, which should be 0.7 or higher (Henseler et al., 2009). All values exceeded this cut-off point with two exceptions: the SIMS (AVE = 0.62) and the PCS (AVE = 0.55); however, their AVE values exceeded 0.50, so – on average – more variability in the items of these scales was accounted for by their factors than was not. Discriminant validity (the extent to which a measure of a particular construct differs from measures of other constructs) was assessed by analysing the AVE by each construct from its indicators, which – according to the Fornell–Larcker-criterion – should be greater than its squared correlation with the remaining constructs. All values met this condition, with one exception: the squared correlation between clarity of goals and its higher-order construct preconditions of flow (0.82) exceeded AVE for clarity of goals (0.80). In conclusion, the reliability, and the convergent and discriminant validity of the multi-item constructs were confirmed. Per participant, a composite score was created for each of the factors, using the PLS weighted-average algorithm.

3.2. The effects of artefact complexity and task complexity on outcome measures

Descriptives. Descriptive statistics of the latent variables (see Table 4) show a negative effect of task complexity on
Table 2
Coefficients of reliability and convergent validity.

<table>
<thead>
<tr>
<th>Construct/indicator</th>
<th>Average variance extracted</th>
<th>Composite reliability</th>
<th>Loading</th>
<th>Standard error</th>
<th>$t^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance of challenge and skill</td>
<td>0.93</td>
<td>0.96</td>
<td>0.96</td>
<td>0.01</td>
<td>77.23</td>
</tr>
<tr>
<td>- CS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarity of goals</td>
<td>0.80</td>
<td>0.94</td>
<td>0.88</td>
<td>0.03</td>
<td>31.30</td>
</tr>
<tr>
<td>- CG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CG2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CG3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CG4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
<td>0.03</td>
<td>24.94</td>
</tr>
<tr>
<td>- F1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- F2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- F3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- F4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration</td>
<td>0.79</td>
<td>0.94</td>
<td>0.86</td>
<td>0.03</td>
<td>32.30</td>
</tr>
<tr>
<td>- CN1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CN2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CN3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CN4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>0.88</td>
<td>0.97</td>
<td>0.91</td>
<td>0.02</td>
<td>39.39</td>
</tr>
<tr>
<td>- CT1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CT2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CT4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mergence of action and awareness</td>
<td>0.78</td>
<td>0.87</td>
<td>0.93</td>
<td>0.01</td>
<td>75.18</td>
</tr>
<tr>
<td>- M1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- M2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformation of time</td>
<td>0.77</td>
<td>0.91</td>
<td>0.93</td>
<td>0.02</td>
<td>53.94</td>
</tr>
<tr>
<td>- TT1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- TT2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- TT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcendence of self</td>
<td>0.79</td>
<td>0.92</td>
<td>0.95</td>
<td>0.02</td>
<td>49.43</td>
</tr>
<tr>
<td>- TS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- TS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- TS3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autotelic experience</td>
<td>0.82</td>
<td>0.95</td>
<td>0.85</td>
<td>0.03</td>
<td>28.59</td>
</tr>
<tr>
<td>- AE1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- AE2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- AE3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- AE4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preconditions of flow</td>
<td>0.80</td>
<td>0.94</td>
<td>0.80</td>
<td>0.04</td>
<td>21.87</td>
</tr>
<tr>
<td>- Balance of challenge and skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Feedback</td>
<td>0.77</td>
<td>0.96</td>
<td>0.83</td>
<td>0.06</td>
<td>14.77</td>
</tr>
<tr>
<td>- Clarity of goals</td>
<td>0.80</td>
<td>0.94</td>
<td>0.93</td>
<td>0.01</td>
<td>75.18</td>
</tr>
<tr>
<td>Flow</td>
<td>0.91</td>
<td>0.02</td>
<td>0.91</td>
<td>0.03</td>
<td>27.05</td>
</tr>
<tr>
<td>- Autotelic experience</td>
<td>0.69</td>
<td>0.96</td>
<td>0.80</td>
<td>0.04</td>
<td>12.20</td>
</tr>
<tr>
<td>- Concentration</td>
<td>0.80</td>
<td>0.94</td>
<td>0.87</td>
<td>0.02</td>
<td>18.35</td>
</tr>
<tr>
<td>- Control</td>
<td>0.64</td>
<td>0.96</td>
<td>0.47</td>
<td>0.10</td>
<td>4.80</td>
</tr>
<tr>
<td>- Mergence of action and awareness</td>
<td>0.64</td>
<td>0.96</td>
<td>0.47</td>
<td>0.10</td>
<td>4.80</td>
</tr>
<tr>
<td>- Transcendence of self</td>
<td>0.64</td>
<td>0.96</td>
<td>0.47</td>
<td>0.10</td>
<td>4.80</td>
</tr>
<tr>
<td>- Transformation of time</td>
<td>0.58</td>
<td>0.98</td>
<td>0.58</td>
<td>0.98</td>
<td>7.47</td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td>0.62</td>
<td>0.87</td>
<td>0.65</td>
<td>0.12</td>
<td>5.46</td>
</tr>
<tr>
<td>- IM1</td>
<td>0.86</td>
<td>0.05</td>
<td>15.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- IM2</td>
<td>0.81</td>
<td>0.07</td>
<td>11.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- IM3</td>
<td>0.81</td>
<td>0.07</td>
<td>11.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- IM4</td>
<td>0.81</td>
<td>0.07</td>
<td>11.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task outcome</td>
<td>0.81</td>
<td>0.93</td>
<td>0.84</td>
<td>0.03</td>
<td>30.15</td>
</tr>
<tr>
<td>- Tasks completed</td>
<td>0.89</td>
<td>0.03</td>
<td>31.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Tasks correct</td>
<td>0.97</td>
<td>0.01</td>
<td>159.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived task challenge</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- PTC1</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived artefact complexity</td>
<td>0.55</td>
<td>0.89</td>
<td>0.48</td>
<td>0.11</td>
<td>4.24</td>
</tr>
</tbody>
</table>
task outcome (large effect size$^3$) and the remaining measures (medium or medium to large effect sizes). Furthermore, there was a negative effect of artefact complexity on task outcome (medium effect size), perceived task challenge and balance of challenge and skill (small to medium effect size), but a negligible effect on the remaining measures. However, the effect of task complexity was stronger when artefact complexity was high and vice versa on most measures, thus indicating an interaction effect (small to medium effect size for balance of challenge and skill as well as for flow$^4$, and small effect sizes for feedback, preconditions, task outcome, perceived artefact complexity and goodness, but no effect on perceived task challenge and clarity of goal). These results provide preliminary evidence for the effectiveness of the manipulation of the independent variables.

$^3$Cohen’s (1988) conventions for effect size were used, where $r=0.10$ represents a small effect size, $r=0.30$, a medium effect size and $r=0.50$, a large effect size.

$^4$PLS allows the modelling of higher-order constructs. With reflective measurement, we used SmartPLS to create latent-variable scores on the higher-order constructs of preconditions and flow proper.

Manipulation check. Given these descriptive results, PLS analysis was used to conduct statistical inference tests of the effects of the experimental manipulations, while accounting for measurement error and maximising explained variance. First, as a manipulation check, the effects of artefact complexity and task complexity on perceived site complexity (PCS) and perceived task challenge (PTC), respectively, were tested. Task complexity should have a negative effect on PTC and artefact complexity should have a negative effect on PCS. The negative effects of task complexity, $t=3.75, p<0.001$, and artefact complexity, $t=2.69, p<0.01$, on PTC were significant, but the interaction effect was not, $|t|<1$. Furthermore, the negative effect of task complexity on PCS was significant, $t=3.75, p<0.001$, but the effect of artefact complexity, $|t|<1$, and the interaction effect, $t=1.18, p>0.05$, were not. Therefore, the measure of perceived task challenge was sensitive to the manipulation of both task complexity and artefact complexity, but the measure of perceived artefact complexity was only sensitive to the manipulation of task complexity. In addition, the effect of task complexity on goodness was significant, $t=4.51, p<0.001$, but the

---

### Table 2 (continued)

<table>
<thead>
<tr>
<th>Construct/indicator</th>
<th>Average variance extracted</th>
<th>Composite reliability</th>
<th>Loading</th>
<th>Standard error</th>
<th>$t^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS2</td>
<td>0.86</td>
<td>0.03</td>
<td>34.39</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>PCS3</td>
<td>0.85</td>
<td>0.03</td>
<td>25.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS4</td>
<td>0.88</td>
<td>0.02</td>
<td>38.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS5</td>
<td>0.69</td>
<td>0.07</td>
<td>10.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS6</td>
<td>0.77</td>
<td>0.07</td>
<td>11.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS8$^c$</td>
<td>0.62</td>
<td>0.07</td>
<td>9.11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Bootstrap, $N=5000$.

$^b$PCS-Item 7 was removed because of an extremely low loading (0.09).

$^c$Results for item after responses were reversed.

### Table 3

Coefficients of discriminant validity.

<table>
<thead>
<tr>
<th>IM</th>
<th>AE</th>
<th>CS</th>
<th>CN</th>
<th>CT</th>
<th>FB</th>
<th>CG</th>
<th>MA</th>
<th>TS</th>
<th>TT</th>
<th>CH</th>
<th>FL</th>
<th>PF</th>
<th>PCS</th>
<th>TO</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td>0.22</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>0.00</td>
<td>0.02</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>0.05</td>
<td>0.20</td>
<td>0.06</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.02</td>
<td>0.16</td>
<td>0.15</td>
<td>0.45</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB</td>
<td>0.06</td>
<td>0.08</td>
<td>0.16</td>
<td>0.12</td>
<td>0.20</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>0.02</td>
<td>0.06</td>
<td>0.47</td>
<td>0.07</td>
<td>0.26</td>
<td>0.25</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>0.02</td>
<td>0.07</td>
<td>0.18</td>
<td>0.08</td>
<td>0.43</td>
<td>0.13</td>
<td>0.32</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.09</td>
<td>0.13</td>
<td>0.03</td>
<td>0.06</td>
<td>0.08</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>0.10</td>
<td>0.26</td>
<td>0.00</td>
<td>0.11</td>
<td>0.09</td>
<td>0.04</td>
<td>0.02</td>
<td>0.09</td>
<td>0.05</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>0.02</td>
<td>0.01</td>
<td>0.24</td>
<td>0.01</td>
<td>0.05</td>
<td>0.08</td>
<td>0.21</td>
<td>0.19</td>
<td>0.02</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FL</td>
<td>0.12</td>
<td>0.47</td>
<td>0.12</td>
<td>0.64</td>
<td>0.77</td>
<td>0.21</td>
<td>0.22</td>
<td>0.41</td>
<td>0.23</td>
<td>0.33</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF</td>
<td>0.03</td>
<td>0.07</td>
<td>0.63</td>
<td>0.12</td>
<td>0.31</td>
<td>0.59</td>
<td>0.82</td>
<td>0.31</td>
<td>0.05</td>
<td>0.02</td>
<td>0.24</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS</td>
<td>0.01</td>
<td>0.04</td>
<td>0.10</td>
<td>0.09</td>
<td>0.21</td>
<td>0.14</td>
<td>0.20</td>
<td>0.28</td>
<td>0.06</td>
<td>0.02</td>
<td>0.14</td>
<td>0.20</td>
<td>0.22</td>
<td>0.56</td>
</tr>
<tr>
<td>TO</td>
<td>0.00</td>
<td>0.02</td>
<td>0.29</td>
<td>0.05</td>
<td>0.08</td>
<td>0.09</td>
<td>0.21</td>
<td>0.12</td>
<td>0.02</td>
<td>0.00</td>
<td>0.28</td>
<td>0.08</td>
<td>0.25</td>
<td>0.12</td>
</tr>
</tbody>
</table>

effect of artefact complexity, \( |t| < 1 \), and the interaction effect, \( t = 1.09, p > 0.05 \), were not. \(^5\)

Tests of \( H1a–H2c \). In order to test \( H1a–H2c \), further PLS analysis was conducted. There were significant negative main effects on task outcome of task complexity (in support of \( H2a \), \( t = 14.27, p < 0.001 \), and artefact complexity (in support of \( H1a \), \( t = 4.94, p < 0.001 \); however, the negative interaction effect was also significant, \( t = 2.46, p < 0.01 \), so the effect of task complexity was stronger when artefact complexity was high and vice versa. \(^5\) Significant negative effects on the balance of challenge and skill were those of task complexity (in support of \( H2b \),)

---

Table 4
Outcome measures as a function of site complexity and task complexity.

<table>
<thead>
<tr>
<th>Site complexity</th>
<th>Task complexity</th>
<th>Clarity of goal</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>C</td>
<td>O</td>
<td>ES</td>
</tr>
<tr>
<td>Simple</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean</td>
<td>5.50</td>
<td>5.55</td>
<td>5.68</td>
</tr>
<tr>
<td>- SD</td>
<td>1.17</td>
<td>1.02</td>
<td>1.10</td>
</tr>
<tr>
<td>Complex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean</td>
<td>5.92</td>
<td>4.42</td>
<td>5.17</td>
</tr>
<tr>
<td>- SD</td>
<td>1.06</td>
<td>1.40</td>
<td>1.45</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean</td>
<td>5.86</td>
<td>4.98</td>
<td>5.42</td>
</tr>
<tr>
<td>- SD</td>
<td>1.11</td>
<td>1.34</td>
<td>1.30</td>
</tr>
<tr>
<td>Effect size</td>
<td>(^b)-0.34</td>
<td>(^c)-0.24</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preconditions of flow</th>
<th>Flow</th>
<th>Task outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site complexity</td>
<td>Task complexity</td>
<td>Task complexity</td>
</tr>
<tr>
<td>S</td>
<td>C</td>
<td>O</td>
</tr>
<tr>
<td>Simple</td>
<td>5.34</td>
<td>4.83</td>
</tr>
<tr>
<td>- Mean</td>
<td>1.13</td>
<td>0.98</td>
</tr>
<tr>
<td>Complex</td>
<td>5.55</td>
<td>4.43</td>
</tr>
<tr>
<td>- Mean</td>
<td>0.95</td>
<td>0.89</td>
</tr>
<tr>
<td>Overall</td>
<td>5.44</td>
<td>4.63</td>
</tr>
<tr>
<td>- Mean</td>
<td>1.04</td>
<td>0.95</td>
</tr>
<tr>
<td>Effect size</td>
<td>-0.38</td>
<td>-0.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perceived task challenge</th>
<th>Perceived artefact complexity</th>
<th>Goodness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site complexity</td>
<td>Task complexity</td>
<td>Task complexity</td>
</tr>
<tr>
<td>S</td>
<td>C</td>
<td>O</td>
</tr>
<tr>
<td>Simple</td>
<td>2.31</td>
<td>3.55</td>
</tr>
<tr>
<td>- Mean</td>
<td>1.33</td>
<td>1.57</td>
</tr>
<tr>
<td>Complex</td>
<td>2.97</td>
<td>4.25</td>
</tr>
<tr>
<td>- Mean</td>
<td>1.45</td>
<td>1.44</td>
</tr>
<tr>
<td>Overall</td>
<td>2.64</td>
<td>3.90</td>
</tr>
<tr>
<td>- Mean</td>
<td>1.42</td>
<td>1.53</td>
</tr>
<tr>
<td>Effect size</td>
<td>0.40</td>
<td>0.01</td>
</tr>
</tbody>
</table>

\( a \)Effect size for site complexity.

\( b \)Effect size for task complexity.

\( c \)Effect size for interaction effect (site complexity by task complexity).

---

\(^5\)Simple-effect tests showed that, with high artefact complexity, the effect of task complexity was significant, \( t = 17.67, p < 0.001 \), as well as with low task complexity, \( t = 7.92, p < 0.001 \). Furthermore, with high task complexity, the effect of artefact complexity was significant, \( t = 6.88, p < 0.001 \), but not with low task complexity, \( t = 1.66, p > 0.05 \).
$t = 4.47, p < 0.001$, artefact complexity (in support of $H_{1b}$), $t = 2.52, p < 0.01$, and the negative interaction between artefact- and task complexity, $t = 2.90, p < 0.01$.5 Task complexity had a significant negative effect on clarity of goals (in support of $H_{2b}$), $t = 3.96, p < 0.001$, but the main effect of artefact complexity (not in support of $H_{1b}$), $|t| < 1$, and the interaction effect, $t = 1.53, p > 0.05$, did not. The negative effect of task complexity on feedback was also significant (in support of $H_{2b}$), $t = 4.35, p < 0.001$, but the effect of artefact complexity (not in support of $H_{1b}$) and the interaction effect were not, both $|t| < 1$. Furthermore, the effect of task complexity on the higher-order construct of preconditions was significant (in further support of $H_{2c}$), $t = 4.84, p < 0.001$, but not significant were the effect of artefact complexity, $|t| < 1$, and the interaction effect, $t = 1.70, p > 0.05$. The effect of task complexity on flow was significant (in support of $H_{2c}$), $t = 4.84, p < 0.001$, but the effect of artefact complexity was not (not in support of $H_{1c}$), $|t| < 1$; however, the negative interaction effect was significant, $t = 2.74, p < 0.01$, so the effect of task complexity was stronger when artefact complexity was high.7 Altogether, the results presented in this subsection demonstrate that the manipulation of both artefact- and task complexity was effective.

3.3. The effects of intrinsic motivation and flow experience on outcome measures

Tests of $H_3$–$H_6$. In order to test $H_3$–$H_6$, PLS analyses were conducted with experimental manipulations (the manipulations of artefact complexity and task complexity combined as a single predictor variable) and intrinsic motivation as exogenous variables, and preconditions of flow experience, flow experience and task outcome as endogenous variables (see Fig. 3). The contribution of the experimental manipulations in explaining variance in the endogenous variables was analysed with PLS by using predicted scores on these variables from artefact complexity, task complexity and their interaction, simultaneously predicting each endogenous variable. This is because, in accordance with Hypotheses 3–6, the effects of the combined experimental manipulations (including their interaction) were of interest rather than separate elements of these. The combined predictors explained 57% of variance in task outcome (see Fig. 3(C)).

To test $H_{3a}$, intrinsic motivation was regressed onto task outcome (see Fig. 3(A)). However, the test result showed that intrinsic motivation was not a significant predictor of task outcome, not in support of $H_{3a}$. $H_{3b}$ was tested by regressing intrinsic motivation onto preconditions (see Fig. 3(A)). In support of $H_{3b}$, intrinsic motivation was a significant predictor of preconditions, but only with experimental manipulations not held constant. $H_{3c}$ was tested by regressing intrinsic motivation onto flow (see Fig. 3(B)). In support of $H_{3c}$, intrinsic motivation was a significant predictor of flow. To test $H_{4}$, preconditions was regressed onto task outcome (see Fig. 3(A)). In support of $H_{4}$, it was found that

![Fig. 3. The effects of experimental manipulations, flow and task performance on task outcome. Note. Experimental manipulations: the combinations of the manipulations of site complexity and task complexity. Figures in brackets show results of the total effect of antecedents on consequents. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.](image)
preconditions remained a significant predictor of task outcome when experimental manipulations and intrinsic motivation were held constant. In addition, experimental manipulations remained a significant predictor of task outcome and were a significant predictor of preconditions. Given the test results for H4 and the additional results, preconditions was a partial mediator of experimental manipulations on task outcome; this is because the effect of experimental manipulations remained significant with preconditions held constant. 

H5 was tested by regressing preconditions onto flow while experimental manipulations and intrinsic motivation were held constant (see Fig. 3(B)). In support of H5, preconditions remained a significant predictor of flow. In addition, experimental manipulations remained significant predictors of both flow and preconditions. Given the test results for H5 and the additional results, preconditions was a partial mediator of experimental manipulations on flow; this is because the effect of experimental manipulations remained significant with preconditions held constant.

Although this was not the main focus of the current study, to allow a more detailed comparison with Guo and Poole's (2009) results, a more fine-grained analysis was conducted where the three preconditions were tested separately as mediators of the effect of experimental manipulations on flow experience. Both clarity of goals, $\beta=0.31$, $t=2.24$, $p<0.05$, and feedback, $\beta=0.24$, $t=-2.69$, $p<0.01$, had a significant effect on flow, but balance of challenge and skill did not, $\beta=-0.08$, $|t|<1$, when each of these three preconditions was regressed onto flow. However, the effect of balance was significant when clarity of goals and feedback were not held constant, $\beta=0.23$, $t=2.41$, $p<0.01$.

To test H6, flow was regressed onto task outcome, while experimental manipulations, intrinsic motivation and preconditions were held constant (see Fig. 3(C)). Not in support of H6, flow was not a significant predictor of task outcome. However, flow was a significant predictor of task outcome when only intrinsic motivation was held constant. Nonetheless, preconditions remained significant while artefact complexity, task complexity, intrinsic motivation and now also flow experience were held constant (see Fig. 3(C)). Therefore, preconditions remained a partial mediator of the effect of experimental manipulations on task outcome. 

Tests of the model variables on evaluative judgment. The effects of experimental manipulations, intrinsic motivation, the preconditions of flow and flow proper on task outcome were contrasted with those on goodness to demonstrate how preconditions and flow differ in their effects on task outcome and evaluative judgment (see Fig. 4). In contrast to the results in the model for task outcome (where 57% of variance was explained), the combined predictors explained only 21% of variance in goodness (see Fig. 4(C)). In particular, while the effect of preconditions on task outcome was 6% ($f^2=0.15$, a medium effect size $^8$) in terms of additional variance explained, its effect on goodness was only 2% ($f^2=0.025$, a small effect size). The results also demonstrate that preconditions were not a mediator of the effect of experimental manipulations on goodness in the full model, but were a partial mediator in the model with flow removed. Furthermore, flow was not a mediator of the effect preconditions on goodness (but, with only intrinsic motivation held constant, small effect size).

---

$^8$Cohen’s (1988) conventions for effect size $R^2$, adapted for use in PLS path modelling (Henseler et al., 2009), were used, with values of 0.02 for a small effect size, 0.15 for a medium effect size and 0.35 for a large effect size.
flow was a significant predictor of goodness). In conclusion, the predictive value of the experimental manipulations, preconditions, flow and intrinsic motivation for goodness was markedly weaker than that for task outcome. In particular, the effect of preconditions was specific: although its effect on task outcome was significant, its effect on goodness was not with other variables held constant.

4. Discussion

4.1. Exploration of main findings

Given that the important relationship between experience and task performance in human–computer interaction remains unaddressed by the major published models of interaction experience, the aim of this study was to advance knowledge of this relationship, while addressing the limitations of previous work. Overall, the results have clarified the role of two higher-order factors of flow experience (flow proper and its preconditions) in relation to their antecedents (person-, artefact-, and task-characteristics) and consequents (task outcome and, as a comparison, goodness). Specifically, by constructing a staged model of flow experience, we were able to demonstrate the effect of experimentally manipulated task- and artefact factors (task complexity and artefact complexity) on task outcome, balance of challenge and skill, and flow experience proper (as per $H1a$–$H1c$ and $H2a$–$H2c$). In addition, the effect of intrinsic motivation (as a person characteristic) on flow was significant (as per $H3c$, even with experimental manipulations held constant), as was the effect on the preconditions of flow (as per $H3b$, but only without the experimental manipulations held constant), but not on task outcome (contrary to $H3a$). Furthermore, the effect of experimental manipulations on flow experience was mediated by the preconditions of flow (as per $H5$), and flow was simultaneously also influenced by the person characteristic of intrinsic motivation. Moreover, we demonstrated that the effect of experimentally manipulated task- and artefact factors on task outcome was mediated by preconditions (as per $H4$), but not by flow experience proper (contrary to $H6$). The effect of preconditions was specific in that it was a significant predictor of task outcome, but not of goodness (overall evaluation of artefact) with experimental manipulations, flow and intrinsic motivation held constant.

$H1$ and $H2$. The effect of artefact complexity (as per $H1a$) on task outcome is consistent with Gwizdka and Spence’s (2006) idea that task difficulty increases with the number of navigation choices and with Blackmon et al.’s (2002) finding of a lower success rate with increased number of links per page. Furthermore, the dual-criterion model of menu selection (Pierce et al., 1992) is consistent with our results. The effect of task complexity ($H2a$) on task outcome is consistent with van Oostendorp et al., (2009) findings and with the mechanisms derived from Gwizdka and Spence (2006). The effects of artefact- and task complexity on preconditions (balance of challenge and skill, clarity of goals and feedback) and flow proper (as per $H1b$, $H1c$, $H2b$ and $H2c$) is consistent with the idea – based on Guo and Poole (2009) – that as task complexity increases the preconditions are adversely affected and consequently flow experience decreases.

$H3a$–$H3c$. Previous human–computer-interaction research (van Schaik and Ling, in press) that investigated the effect of intrinsic motivation (as a person characteristic) on flow experience (as per $H3c$), with experimental manipulations of task- and artefact complexity held constant, failed to demonstrate this effect. It seems probable that this negative result can be accounted for by the use of a global flow measure that was based on statistical grounds. However, the current study was able to demonstrate the effect by using a theoretically justified conceptualisation of flow as a higher-order construct, separate from its preconditions. Specifically, intrinsic motivation was an independent positive predictor of flow, but not of preconditions or task outcome. The effect of intrinsic motivation on flow follows from the idea that, before starting a particular activity, those who are more intrinsically motivated will expend more effort in that activity (in this case web navigation) and will therefore have a higher level of involvement in the activity; in other words, they will experience a greater degree of flow. Although the total effect of intrinsic motivation on preconditions was significant, without experimental manipulations held constant, the effect was not significant. This result is not surprising, because, given the attention-enhancing cognitive nature of the preconditions of flow, they should be influenced by manipulations (e.g., those of task complexity and artefact complexity) that affect cognitive processes rather than by a motivational predisposition (e.g., intrinsic motivation). The effect of intrinsic motivation on task outcome was not significant, perhaps because the effect of intrinsic motivation would be mediated by flow (which was not an independent predictor of task outcome) and the information retrieval task was defined by the experiment (therefore, having an externally-defined goal) rather than chosen by the participants, whereas intrinsically-motivated individuals are those who engage in activities for the sake of these activities rather than to achieve some external goal.

$H4$–$H6$. First, it had to be established that the preconditions of flow were a mediator of the effect of the experimental manipulations of artefact complexity and task complexity on flow experience proper. The results of mediation analysis demonstrate that preconditions is indeed a (partial) mediator (as per $H5$). Second, it had to be established that flow experience is a mediator of the effect of preconditions on task outcome. The results of mediation analysis indicated that flow was not a mediator (contrary to $H6$), but preconditions remained a significant predictor of task outcome, with experimental manipulations, flow, intrinsic motivation held constant and were a mediator of the effect of experimental manipulations (as per $H4$). The results for $H4$ – preconditions of flow have a positive effect on task outcome – are consistent with the
results of previous research in other domains (Engeser and Rheinberg, 2008; Vollmeyer and Imhof, 2007). The positive effect of flow on task outcome was demonstrated (though previous research did not make a conceptual distinction between flow proper and its preconditions), supporting the first pathway from flow to task outcome, where a better functional state increases task performance. The results for \( H_6 \) – with experimental manipulations and preconditions held constant, flow has no positive effect on task outcome through a higher motivation to perform – do not fully support the second pathway; however, the effect was significant without these variables held constant. These results indicate that the influence of the cognitive flow component (preconditions as a higher-order factor) on task performance was stronger than the motivational component (flow proper as a higher-order factor). Furthermore, the results supporting \( H4 \) (preconditions of flow have a positive effect on task outcome) confirm the dual role of preconditions: as a mediator of the effect of task characteristics and artefact characteristics on both flow and task outcome. The different results for preconditions and flow proper can be accounted for through the work of Engeser and Rheinberg (2008), with reference again to the two pathways for the positive effect of flow on task outcome. As in Engeser and Rheinberg’s Study 2, the first pathway applies because the task could be seen as relatively low in importance (e.g., in terms of consequences for test-users’ performance on their programme of studies) and “the experimental situation was standardized and thus did not allow for additional practice” (p. 166). The second pathway would apply in situations where the task is seen as high in importance and in situations where people are learning as a non-compulsory activity (Engeser and Rheinberg’s Study 3).

The non-significant effect of artefact complexity on perceived artefact complexity indicates that artefact complexity is itself a complex construct and that the measure that was used (PCS) lacked sensitivity in this context. This is perhaps not surprising, as artefacts in general and websites in particular can vary on many dimensions of complexity. In this case, the aspect of artefact complexity that was manipulated was the navigation system (one of the three components of information architecture, alongside organisation system and labelling system, Morville and Rosenfeld, 2007) and operationalized as the number of links on a web page, and thereby potential navigation paths and overall website size. However, the items of the scale seemed to tap into other aspects of complexity, such as screen layout, which is not part of the information architecture of a website, but of its presentation design. Other measures of perceived complexity would be needed to capture the complexity created by the components of the information architecture. For example, Ahuja and Webster’s (2001) disorientation measure could be used as a measure of complexity associated with the navigation system or the organisation system.

4.2. Implications of the research model and future work

Within a broader context, the current study is part of the quest for a better understanding of the nomological network of flow experience. In particular, it is important to consider the antecedents of flow and its (cognitive-task-performance-enabling) preconditions, and the consequents of flow and its preconditions in order to appreciate the role of flow in psychological processes and human behaviour, and the consequences for human–computer interaction beyond web navigation.

4.2.1. Preconditions of flow as antecedents of flow proper

Guo and Poole (2009) call for the distinction between the preconditions of flow and flow proper from the original conceptualisation of flow experience (Csikszentmihalyi, 1988; Csikszentmihalyi and Nakamura 2010) to be consistently applied in human–computer interaction for conceptual clarity and as a basis for a better understanding of flow in relation to its antecedents and consequents. In addition to our findings, other empirical evidence confirms the preconditions of flow as a positive predictor of flow experience.

In experimental research, a balance of challenge and skill has been found to be a partial mediator of the experimentally manipulated balance on flow in computer-game playing (involvement/enjoyment; Keller and Bless, 2008). In addition, action orientation (persistence–volatility) was a moderator of the manipulated balance on flow experience: more persistent individuals experience more flow, but only when balance was optimal (Keller and Bless, 2008).

In non-experimental research, the preconditions of balance of challenge and skill, and feedback (but not clarity of goal) were partial mediators of the negative effect of perceived artefact complexity on flow in navigating e-commerce Internet sites (Guo and Poole, 2009). Furthermore, six preconditions of flow (balance of challenge and skill, feedback, clarity of goal, ‘playability’, ‘gamefulness’ and ‘frame story’) were correlated with flow experience in playing educational computer games (Kiili and Lainema, 2008). Moreover, the gap between challenge and skill was a predictor of flow experience in a virtual class environment (Shin, 2006).

In sum, evidence supports the idea of the preconditions of flow as antecedents and a mediator of the effect of challenge/skill balance on flow. In addition, those who have a stronger disposition of persistent action experience more flow when challenge and skill are in balance. Our finding of preconditions as an independent positive predictor of flow further supports the staged model of flow.

---

\(^{9}\)In non-experimental studies it is more difficult to infer causal effects, as this research does not test the effect of a manipulation that specifically targets each of the preconditions of flow separately and mediation of the effect on flow proper.
4.2.2. Antecedents of flow and its preconditions

According to Finneran and Zhang’s (2003) person–artefact–task model, the characteristics of three main factors – person, artefact and task – all influence flow experience. From the perspective of flow theory, person characteristics can contribute towards flow in the form of (a) in-born or developed capabilities that facilitate task performance by focusing attention and (b) motivational predispositions that produce flow-producing behaviour, with flow (in turn) as a motivation to repeat this behaviour. Another justification for the search for person characteristics as predictors of flow comes from the finding that between-individual variance accounts for 26% of the total variance in flow, while 74% of the variance was attributable to within-individual variation in architecture students (Fullagar and Kelloway, 2009). From the perspective of flow theory, artefacts and tasks can promote flow through (a) characteristics that facilitate focused attention and (b) characteristics that motivate people to engage in flow-producing behaviour, which acts as a motivation for repeat-behaviour. A wide range of empirical evidence supports person-, artefact- and task characteristics as antecedents of flow.\(^\text{10}\)

Person. Some research has studied person characteristics as antecedents of flow. Intrinsic motivation (as a personality trait) was a positive predictor and amotivation a negative predictor of flow, but extrinsic motivation was not a predictor in architecture students (Fullagar and Mills, 2008). Furthermore, need for autonomy positively moderated the relation between intrinsic motivation and flow, with a stronger influence of motivation in those with a greater need for autonomy. Given a balance of challenge and skill, those with a high implicit achievement motivation (hope of success) experienced a higher level of flow than those with a high achievement motive (fear of failure) in learning statistics (Engeser and Rheinberg, 2008) and studying psychology (Schüller, 2007). Moreover, Baumann and Scheffer (2010, 2011) proposed and provided evidence for an intrinsic component of the achievement motive (and measured by the Operant Motive Test) as a stable characteristic that predisposes people to experiencing flow. The achievement flow motive was offered as a way to operationalise Csikszentmihalyi’s (1988) concept of autotelic personality.

Furthermore, both self-concept and psychological skills were positive predictors of dispositional flow in elite athletes, and in turn, dispositional flow and psychological skills were independent predictors of flow state in elite athletes (Jackson et al., 2001). In addition, in online shopping, interest in product type on sale and possession of Web skills were independent positive predictors of concentration (as a flow dimension; Koufaris, 2002). Finally, (previous) meditation experience was a positive predictor of flow dimensions at customer-service work (Kuo and Ho, 2010).

Other research has studied autotelic personality (flow proneness) as an antecedent of other personality characteristics or behavioural outcomes. Asakawa (2010) investigated the concept of autotelic personality in Japanese college students. The findings indicate that frequency of flow experience was correlated with relatively stable psychological characteristics: positively with self-esteem and the use of active coping strategies, and negatively with anxiety and the use of passive coping strategies. Therefore, autotelic personality may contribute towards stable healthy psychological characteristics.

In sum, there is evidence demonstrating that a range of person variables are antecedents of flow, most notably variants of intrinsic motivation, but also dispositional flow, psychological skills, interest in task domain, and the possession of domain-specific and general skills. Our finding of intrinsic motivation as an independent positive predictor of flow provides further support for the role of person variables. There is also other evidence for autotelic personality as a potential antecedent of other personality characteristics or behavioural outcomes. When the influence of person characteristics is taken into account a more nuanced analysis of the impact of potential of effects artefact- and task variables or the conditions for their impact becomes possible.

Artefact. Experimental research has demonstrated that, with text-to-speech voice communication, voice-only communication resulted in a higher level of flow experience of live help in e-commerce than text-and-voice presentation (Qiu and Benbasat, 2005). However, presentation mode of e-learning was an antecedent of concentration (as a dimension of flow), with more concentration when presentation was audio-text-video rather than using in other combinations of two media including audio (Liu et al. 2009). These apparently conflicting results indicate that the effect of communication mode on flow may depend on the type of task: additional modes can be distracting in two-way communication by reducing focused attention, but can enhance flow in a learning task by redundant coding (Mayer, 2001).

Results from non-experimental research (Choi et al., 2007; Guo and Poole, 2009; Huang, 2003; Skadberg and Kimmel, 2004) indicate that interface/usability characteristics (through cognitive facilitation) are antecedents of flow. Further research findings (Choi et al., 2007; Ding et al., 2010; Huang, 2003; Skadberg and Kimmel, 2004) show that content/functionality characteristics (through motivational facilitation) are also antecedents of flow. In sum, evidence indicates that the effect of communication mode on flow may be moderated by task type. Furthermore, both usability-related and functionality-related artefact characteristics have an effect on flow. Our finding of the effect of experimentally manipulated artefact complexity on both

\(^{10}\)Perhaps understandably, given the different aims of the various research studies on flow, most work has tended to focus on only one of the main factors. As a consequence, it is often impossible to determine whether the effect of a particular variable is an independent predictor of flow experience.
flow and its preconditions further supports the influence of artefact characteristics on flow.

**Task.** Experimental research has found that flow was highest when task difficulty was optimal rather than extremely low or extremely high in playing a computer game (Rheinberg and Vollmeyer, 2003). However, when less extreme levels of difficulty were used, flow was higher when task difficulty was optimal or difficult rather than easy (Schiefele and Roussakis, 2006).

In non-experimental research, for tasks with low perceived importance, a balance of challenge and skill produced flow experience, but otherwise flow was experienced when skill exceeded challenge in statistics learning, computer-game playing and learning a foreign language (Engeser and Rheinberg, 2008). Independent positive predictors of flow experienced by managers in their work were the activity types of planning, problem solving, and evaluation (Nielsen and Cleal, 2010). In sum, evidence indicates that task difficulty has a curvilinear effect on flow, but perceived task importance moderates the effect of challenge/skill balance on flow and particular types of work activity increase flow. Our finding of experimentally manipulated task complexity on flow and its preconditions provides further evidence for the influence of task characteristics on flow.

### 4.2.3. Consequents of flow and of its preconditions

Research has reported three types of consequent of flow experience: objective, behavioural and subjective. Sheldon et al. (2001) have called for types of objective outcome to ‘firmly ground’ experience (here, flow experience), including task-performance outcomes and health outcomes. Our H4 and H6 and supporting evidence provide a basis for the effect of flow on task performance. From the perspective of flow theory, because flow experience leads to a desire to experience flow again, its behavioural effect is that it acts as a motivating force to repeat flow-producing behaviour (Csikszentmihalyi, 1988). According to flow theory, flow experience (by creating harmony between the content of consciousness and goals) will produce order in consciousness, which is experienced as enjoyable (Csikszentmihalyi, 1988). In this sense, the subjective consequent of enjoyment in the short term is (just) an epiphenomenon of flow. However, in the longer term, as a consequence of memory processes in consciousness, memories of positive emotion and volition in previous flow episodes could produce positive predispositions, such as satisfaction with life and intrinsic motivation.

**Objective outcomes.** In competitive sport, several flow dimensions were independent predictors of objectively measured athletic task performance (challenge/skill balance, goal clarity and autotelic experience), intersubjectively rated performance (challenge/skill balance, goal clarity, feedback and autotelic experience) and subjective self-rated performance (challenge/skill balance and autotelic experience; Jackson et al., 2001; Stavrou et al., 2007). However, flow can also have an indirect effect. In marathon running, a more complex effect of flow was found: flow experienced during a race was related to future running motivation, but not directly linked to race performance. Rather, pre-race training behaviour was a predictor of race performance and this behaviour was promoted by flow experienced during training (Schüler and Brunner, 2009). This is because, given the positive feelings accompanying flow, flow acts as a motivator for repeat-behaviour (Csikszentmihalyi, 1988).

In computer-game playing, flow experience is positively related to and a marginal independent positive predictor of task performance (Murphy et al., 2008; Engeser and Rheinberg, 2008). The latter result is qualified by a positive correlation between flow and task performance only when the level of task difficulty is optimal (Vollmeyer and Rheinberg, 2003). Furthermore, flow experience during an academic course was a positive predictor of exam task performance in undergraduate psychology (Schüler, 2007), statistics and a foreign language (Engeser and Rheinberg, 2008).

Flow was a full mediator of the relation between characteristics of academic work (autonomy, role clarity and feedback) and psychological well-being. In addition, psychological well-being was a full mediator of the relationship between flow and physical health (Steele and Fullagar, 2009). In sum, the reviewed evidence further supports flow as an antecedent of task performance in voluntary and non-voluntary settings, and provides evidence for the beneficial effect of flow on psychological and physical health. Our finding that preconditions of flow are an independent predictor of task performance, mediating the effects of experimentally manipulated artefact- and task complexity, further supports the staged model of flow and the influence of flow on objective outcomes.

**Behavioural outcomes.** Research has found evidence for both positive and negative behavioural effects of flow. Flow was a positive predictor of self-reported exploratory behaviour and learning in educational-game playing (Kiili and Lainema, 2008) and of exploratory behaviour in Internet-site use (Zhao et al., 2011). Flow in computer use in a workplace was a predictor of exploratory use behaviour and the effect of flow on the extent of computer use was mediated by exploratory behaviour (Ghani and Deshpande, 1994). Flow in e-learning was a predictor of self-reported level of achieving learning outcomes (Ho and Kuo, 2010). Asakawa (2010) found positive correlations of flow frequency with specific tasks in college students’ life: active commitments to college life, search for future career, and daily activities. However, flow was also an independent positive predictor of problematic Internet use (Thatcher et al., 2008) and of addictive behaviour in cyber-game playing (Chou and Ting, 2003). In sum, there is evidence that (the motivational component of) flow is a predictor of both unproblematic and problematic behaviour.

**Subjective outcomes.** Regarding short-term effects of flow, experimental research has demonstrated that flow experience in college students is a mediator of the effect of
flow induction as a task characteristic on positive affect (Rogatko, 2009). In non-experimental research, momentary flow in architecture students was predictive of momentary mood and not vice versa (Fullagar and Kelloway, 2009) and flow in students of undergraduate psychology was a positive predictor of positive affect (Schüler, 2007).

Regarding long-term effects of flow, Asakawa (2010) found positive correlations of flow frequency in college students with indicators of the quality of life more generally: Jujitsu-kan (a Japanese sense of fulfilment) and greater satisfaction with life. Therefore, a more frequent experience of flow may result in greater memory for positive emotions experience in flow state as a basis for positive thoughts about life (see also Csikszentmihalyi and Nakamura, 2010). In sum, research evidence supports the short-term positive effect of flow on positive affect and the potential long-term positive effect on quality of life.

4.2.4. Implications for human–computer interaction and challenges for research

A consequence of the conceptual distinction between flow and its preconditions is that usable design can promote flow by enhancing the challenge/skill balance, clarity of goals and feedback (see also van Schaik and Ling, in press) and thereby enhance task performance. Indeed, perhaps unexpectedly, Norman’s (1988) action cycle and his principles of good design are helpful in this respect. This is because the gulf of execution in the action cycle appears to map onto the challenge/skill balance and clarity of goals, and the gulf of evaluation onto feedback. Moreover, the principle of a good conceptual model should assist in creating a better match between the challenge posed by a user-interface and a user’s skills. Visibility and good mapping should help in making goals clear by making procedures and actions visible, and mapping actions onto their expected consequences. Finally, feedback about (for instance) completed actions and a user’s location within the interface structure should help in focusing the user’s attention.

A systematic consideration of person-, artefact- and task variables in relation to flow would allow designers to take into account their implications for creating better interfaces. In terms of person characteristics, an important implication of flow research is that these can affect the extent to which flow is experienced. Therefore, a research challenge and a practical task lie in the (online) measurement of flow-related person characteristics and this could be a basis for adapting human–computer interaction. An obvious example would be that a user’s skill level would need to be matched with the level of challenge. However, human–computer interaction design becomes more complex when intrinsic motivation as a person characteristic is taken into account. This is because the reviewed research implies that the best challenge/skill balance for achieving flow can depend on a user’s level of intrinsic motivation. For example, those with a high hope of success (such as expert computers users) would require a match between perceived challenge and skill in their interaction (e.g., Norman, 1981), but those with a high fear of failure would require a lower level of challenge than skill in order to achieve flow. As an alternative to adaptation, an implication would be that designers should consider the range of user-profiles in terms of person characteristics that influence flow or moderate the effect of other variables on flow, in order to help estimating the potential impact of design decisions on users’ flow. A further research challenge would be to create and validate tools to support designers to enhance flow in end-users.

Regarding artefact characteristics, the use of multiple presentation modes in interaction design could reduce users’ flow by distracting attention. However, judicious use of a combination of modes could enhance flow by redundant coding. Specific usability-related design characteristics could enhance the preconditions of flow (and thereby flow itself). Functionality, however, could promote flow/order in consciousness by creating a match between the contents of awareness (e.g., functionality that is offered) and a user’s goals (e.g., task goal).

In relation to task characteristics, the difficulty of task procedures will affect the challenge/skill balance and thereby flow. Task importance has implications for human-computer interaction design in terms of the best challenge-skill ratio: in important tasks (e.g., ‘mission-critical’ systems) the skills should (far) exceed challenge level. This is not only for safety reasons by reducing the chance of accidents, but also to increase flow and thereby enhance task performance. However, in less important tasks (e.g., in computer games), challenge and skills should be in balance in order to achieve flow.

Regarding objective outcomes of flow, task-performance outcomes could be enhanced by flow-promoting interactive computer systems through (a) effortless attention (Csikszentmihalyi and Nakamura, 2010) and motivation (b) enhanced training of skills and consequent enhanced task performance. Although an apparently more demanding research challenge, flow-promoting computers might enhance psychological well-being during each episode of interaction and thereby, ultimately, physical health. Even though research (Steele and Fullagar, 2009) highlights the potential of such research, the actual technical design and implementation of such a computer-based ‘intervention’ remain to be elucidated.

According to flow theory, by promoting flow, motivation towards repeat-behaviour at a more challenging level can be increased. This principle forms the basis of many computer-game designs, but might be applied more generally to human–computer interaction. However, designers should also be aware of the potential to create flow-inducing artefacts with negative behavioural outcomes (e.g., addictive behaviour).

In terms of subjective outcomes, in the short term, flow-promoting interactions might produce positive feelings, which could be a goal in its own right in terms of increasing users’ (momentary) well-being, but also as a potential
vehicle for marketing flow-producing artefacts. Non-experimen-
tal-research evidence (e.g., Asakawa, 2010) suggests
that, in the longer term, flow could create stable human
dispositions, such as satisfaction with life and flow prone-
ness (seeking out flow-producing activities). An interesting
research challenge is then how to facilitate flow in a
controlled fashion towards creating such dispositions, with
potential major benefits to users and ultimately to society.

In a broader context, taking the perspective of need
fulfilment, flow may be “understood as a variant of a
competence experience” (Hassenzahl et al., 2010, p. 361).
We are not suggesting that the needs that flow experience
can fulfil are the only ones that should be considered in
human–computer interaction at the expense of other basic
psychological needs, such as relatedness and autonomy
(Sheldon, 2011). Rather, our objective is to demonstrate
the important implications of flow as a powerful tool for
improving human–computer interaction.

The current study used an appropriate match of persons
(psychology students), artefact (psychology intranet), task
(information retrieval) and objective outcome measure
(information-retrieval task performance). However, in order
to establish the wider applicability of (a) the approach of
cognitive–experiential modelling of human–computer inter-
action with a staged model of flow and (b) the findings — in
particular the critical role of the preconditions of flow, future
research should investigate a range of appropriate combina-
tions of person, artefact, task and (objective, behavioural
and subjective) outcome measures.

5. Conclusion

This research makes novel contributions, both concep-
tually and experimentally, by demonstrating the crucial
role of the preconditions of flow experience in human–
computer interaction. Taking a wider perspective, the role
of flow in human–computer action is important as a
mediator of the effects of person-, artefact- and task
variables on task performance, behavioural outcomes
and (potentially) stable dispositions, such as life satisfac-
tion. Future work in human–computer interaction should
exploit the potential of computers to promote flow experi-
ence and ultimately thereby the quality of life.

Appendix. Questionnaire items

Intrinsic motivation (Guay et al., 2000)
Why are you currently engaged in this activity?
IM1 Because I think that this activity is interesting.
IM2 Because I think that this activity is pleasant.
IM3 Because this activity is fun.
IM4 Because I feel good when doing this activity.

Response format: 7-point Likert scale with endpoints ‘Corresponds not at all’ and ‘Corresponds exactly’.

Flow (Guo and Poole, 2009)
Balance of challenge and skill
CS1 My abilities matched the challenge of the situation.
CS2 I felt I was competent enough to meet the demands of the situation.

Clarity of goals
CG1 I knew clearly what I wanted to do.
CG2 I had a strong sense of what I wanted to do.
CG3 I knew what I wanted to achieve.
CG4 My goals were clearly defined.

Feedback
F1 It was really clear to me that I was doing well.
F2 I was aware of how well I was performing.
F3 When using the website, I had a good idea about how
well I was doing.
F4 I could tell by the way I was using the website how well
I was doing.

Concentration
CN1 My attention was focused entirely on what I was
doing.
CN2 It was no effort to keep my mind on what was
happening.
CN3 I had total concentration.
CN4 I was completely focused on the task at hand.

Control
CT1 I felt in total control of what I was doing.
CT2 I felt like I could control what I was doing.
CT3 I had a feeling of total control.
CT4 I felt in total control of my action.

Mergence of action and awareness
M1 I reacted to the website automatically.
M2 I did things spontaneously and automatically without
having to think.

Transformation of time
TT1 Time appeared to go by very quickly.
TT2 I lost track of time.
TT3 Time flew.

Loss of self-consciousness
TS1 I was not concerned with what others may have been
thinking of me.
TS2 I was not concerned with how I was presenting myself.
TS3 I was not worried about what others may have been thinking of me.

Autotelic experience

AE1 I really enjoyed the experience.
AE2 I loved the feeling experienced and I want to capture it again.
AE3 The experience left me feeling great.
AE4 I found the experience extremely rewarding.

Response format: 7-point Likert scale with endpoints ‘Not challenging at all’ and ‘Very challenging’.

Perceived task challenge

How challenging did you find the tasks that you have been working on?

Response format: 7-point Likert scale with endpoints ‘Strongly agree’ and ‘Strongly disagree’.

Perceived artefact complexity (Guo and Poole (2009) — adapted from Nadkarni and Gupta (2007))

I judge the Web pages to be

PC1 Open–Cluttered
PC2 Coherent–Incoherent
PC3 Logical–Ilogical
PC4 Organised–Non-organised
PC5 Uniform–Varied
PC6 Congruent–Incongruent
PC7 Sparse–Dense
PC8 Distracting–Non-distracting

Response format: 7-point semantic differential with given endpoints.

References


