RUNNING HEAD: EXPLORATION BEHAVIOR AND MENTAL FATIGUE

The Impact of Mental Fatigue on Exploration in a Complex Computer Task: Rigidity and Loss of Systematic Strategies

Dimitri van der Linden,

University of Amsterdam / Technical University of Eindhoven /University of Nijmegen

Roetersstraat 15, 1018 WB Amsterdam, The Netherlands

Den Dolech 2, 5600 MB Eindhoven, The Netherlands

Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

Michael Frese,

University of Giessen

Germany

Sabine Sonnentag,

Technical University of Braunschweig

Germany

Post-review version

Published in: Human Factors

Link: http://www.hfes.org/publications/productdetail.aspx?productid=1

Address for correspondence:

Dimitri van der Linden
University of Nijmegen
Section Work and Organizational Psychology
P. O. Box 9104
6500 HE Nijmegen
The Netherlands
Tel: +31 24 361 2743
Fax: +31 24 361 5937
E-mail: d.vanderlinden@psych.kun.nl

Abstract

This study investigates the impact of mental fatigue on exploration when performing a complex computer task. Exploration of participants who underwent a fatigue manipulation (N = 36) was compared with a control (non-fatigued) group (N = 32). A distinction was also made between participants with either a low or high level of general computer experience. Results showed fatigued participants to use significantly less systematic exploration and make more errors than non-fatigued participants. Fatigued participants with low experience also showed significantly more rigid behavior than other participants. No differences were found on the number of sub-tasks solved. Compared to low experienced participants, highly experienced participants showed significantly more systematic exploration, less unsystematic trial-and-error, solved more sub-tasks, and made fewer errors (marginally significant $p = .056$). Findings were interpreted as the result of reduced task engagement under fatigue and reduced involvement of executive control on behavior.

The Impact of Mental Fatigue on Exploration in a Complex Computer Task: Rigidity and Loss of Systematic Strategies

It is common in modern work places that people have to deal with increasingly complex problems such as managing a big company, finding the fault in a power plant, or learning a new computer program without specific instructions. Although in practice many of these tasks can be executed with routines or through the use of standardized procedures, there are also many complex tasks in which routines or procedures are not directly available or which need ad-hoc actions that go beyond formal procedures. Dealing with such non-routine tasks often requires exploration (Funke, 1991) to gain insight into the task and to find out which actions can accomplish task goals. Exploration can be done in a systematic or unsystematic way (Dörner 1980; Funke, 1991; van der Linden, Sonnentag, Frese, & van Dyck, 2001). Exploring systematically means that people behave in a goal-directed way and reflect on action feedback and on their own behavior (Trudel & Payne, 1995; van der Linden et al., 2001). In contrast, when exploring unsystematically, people often behave in an unstructured way and do not seem to follow a coherent path towards goal attainment (Dörner, 1980; Hollnagel, 1993). Instead, actions are executed impulsively or are guided by external stimuli that tend to capture attention (cf. Hollnagel, 1993).

An important factor involved in whether people explore in a systematic versus unsystematic way is their level of task engagement (Dörner, 1980). Task engagement refers to the level of cognitive resources such as attention, allocated to task relevant processes (Dörner, 1980; Trudel & Payne, 1995) and that involve problem solving steps such as goal-setting, hypotheses formation, planning, and feedback evaluation (Dörner, 1980). When many resources are allocated to problem solving, exploration is likely to be goal-directed and systematic. In contrast, when task engagement is low and fewer resources are invested, this tends to manifest itself in the use of unsystematic exploration strategies, which may generally be ineffective or inefficient (Hollnagel, 1993; Green & Gilooy, 1990). Although, many factors can influence the level of task engagement, in the current study we focus on one of these factors, namely mental fatigue. Specifically, we measure exploration behavior and assess how systematic versus unsystematic behavior changes under fatigue. As far as we know there is no other study that explicitly investigated exploration under fatigue. Nevertheless, knowing how exploratory behavior may change under fatigue is important because exploration is a substantial part of problem solving in complex tasks (Dörner, 1980; Hollnagel, 1993; Shrager & Klahr, 1986).

Mental Fatigue and Exploration Behavior

Mental fatigue can be defined as a psychophysiological state resulting from sustained performance on cognitively demanding tasks and coinciding with changes in motivation, information processing, and mood (e.g. Meijman, 2000). One of the main characteristics of mental fatigue is an increased resistance against further effort and a tendency to reduce task engagement (Holding, 1984; Meijman, 2000; Sanders, 1998). Thus, when possible, fatigued people will stop working on effortful tasks and postpone the work until they are no longer fatigued. However, even in situations where one cannot stop working, fatigued people still tend to reduce task engagement (often unintentionally)(Meijman, 2000; Sanders, 1998). Such reduced task engagement will not manifest itself as a complete withdrawal from the task or as a complete break-down of performance. More likely, periods of adequate performance will more frequently be alternated with lapses of task-engagement under fatigue (Sanders, 1998). During such lapses, behavior may not be directed by clear task goals but by more automatic cognitive processes (cf. Monsell & Driver, 2000). With regard to exploration behavior it can be expected that in those lapses, people will not show thoughtful, systematic exploration. Before being to able to study exploration under fatigue it is necessary to establish what are the behavioral manifestations of systematic and unsystematic exploration. Therefore, in the following sections we describe three major types of exploration behavior we assess in the current study. With these three types of exploration we do not intent to exhaustively cover all possible forms of exploration behavior but want to capture broad patterns of behavior that people show when working on complex, non-routine tasks (Hollnagel, 1993; Norman, 1991; Trudel & Payne, 1995). We labeled these types of exploration behavior <u>systematic exploration</u>, <u>unsystematic trial-and-error</u>, and <u>rigid behavior</u>.

<u>Systematic exploration</u> implies that a person explores a system in a goal-directed, coherent way (Green & Gilooy, 1990; Trudel & Payne, 1995). This means that hypotheses about where to search for and about possible outcomes of actions are generated and that behavior is guided by these hypotheses (Shrager & Klahr, 1986). Moreover, exploring systematically implies to reflect on action feedback and on working methods. Several studies support the importance of goal-directed, reflective behavior for successful exploration (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Shrager & Klahr, 1986; Trudel & Payne, 1995). For example, Trudel and Payne (1995) analyzed verbalizations of people who explored a digital stopwatch and found that 'good' explorers (the ones who learned most about the watch) tended to verbalize discoveries they made earlier, assessed frequently what has been learned so far, and often confirmed or disconfirmed feedback, derived from testing ideas. In general, using systematic exploration involves a thoughtful, reflective approach to the task (Trudel & Payne, 1995, p. 325). As such, it can be argued that systematic exploration involves a relatively high level of engagement. However, as mental fatigue coincides with a reduction in task engagement (e.g. increased lapses of task engagement), it can be expected that the use of systematic exploration will decrease under fatigue (<u>Hypothesis 1</u>).

<u>Unsystematic trial-and-error</u> refers to exploration that is unstructured and does not seem to be guided by clear hypotheses, nor is it accompanied with signs of reflection (Hollnagel, 1993; Trudel & Payne, 1995). During unsystematic trial-and-error, people often

shift from one sub-goal to another, while none or only a few of these sub-goals are well-considered. In the literature on problem solving and human-computer interaction (HCI) there are many reports of such behavior even though different labels have been assigned to it, for example 'vagabonding' (Dörner, 1980), 'scrambled mode' (Hollnagel, 1993), or 'unsystematic trial-and-error' (Trudel & Payne, 1995). In general, unsystematic trial-and-error may coincide with a withdrawal of cognitive resources from hypotheses formation, planning, and reflection. Moreover, as the tendency to reduce task engagement increases under fatigue we can expect that the use of unsystematic trial-and-error also increases under fatigue (Hypothesis 2).

Rigid behavior is characterized by decreased cognitive flexibility and increased tendency to perseverate. During periods of rigid behavior, actions or ideas are often initially guided by habits or by salient cues that capture attention. Based on such habits or cues people relatively quickly adopt certain action patterns in which they persist even though feedback clearly indicates that this is no longer useful (Dörner, 1980; van der Linden, et al., 2001). Rigid behavior is another specific type of unsystematic behavior often reported in the problem solving or HCI literature. For example, in a study on learning a statistical program through exploration, Green and Gilhooly (1990) found poor learners to show a tendency to repeat methods, to pay less attention to feedback, and to fail to act appropriately on evaluation feedback. A finding that has been replicated in several other studies (e.g. Somsen, van der Molen, Jensen, & van Beek, 2000; Trudel & Payne, 1995; van der Linden, et al. 2001). We expect that a reduction in reflection and in the allocation of attention to action, will lead to an increase in rigid behavior. Thus, we expect an increase of rigid behavior under fatigue (Hypothesis 3).

Performance. In order to study exploration, we used a complex computer task (with Microsoft Excel 4.0) in which participants could freely explore. In accordance with the literature on exploration, we expected the use of systematic, reflective exploration to lead to better learning of options and procedures (Funke, 1991; Green & Gilooy, 1990; Hollnagel, 1993; Trudel & Payne, 1995; van der Linden, et. al, 2001). Moreover, if more procedures of the program are learned then more sub-tasks can be achieved. Thus, we expected a positive relationship between systematic exploration and performance, in terms of the number of sub-goals solved (Hypothesis 4a). Moreover, because systematic exploration involves thoughtful actions we hypothesized this type of exploration to be negatively related to the number of errors (Hypothesis 4b). In contrast, during periods of unsystematic exploration, task engagement is low and so is planning and reflection. This implies that many errors will be made and that relatively little is learned about new procedures or options. Consequently we expected a negative relationship between unsystematic trial-and-error and rigid behavior on the one hand and number of sub-tasks solved on the other hand (Hypothesis 5a). As unsystematic exploration strategies tend to lead to an increase of errors (e.g., rigid behavior leading to ineffective actions) we postulate Hypothesis 5b as a construct validity hypothesis, which states that the use of unsystematic exploration strategies are positively related to the number of errors.

With regard to mental fatigue we expected a negative relationship with solved sub-tasks (Hypothesis 6a) and a positive relationship with errors (Hypothesis 6b). Although such hypotheses make sense, we have to note that there is a surprising number of findings in the literature that do not show clear-cut relationships between fatigue and performance (Hockey, 1997; Holding, 1984; Sanders, 1998). The main reason for this is that people can reallocate resources, thereby forcing themselves to stay engaged in the task despite their fatigue (Hockey, 1997). Nevertheless, we hypothesized a relationship of fatigue with low number of subtasks solved and with a high number of errors, although, we are fully aware that these hypotheses are not easily supported in empirical studies on fatigue (Hockey, 1997; Sanders, 1998).

Mental Fatigue and General Experience

If the allocation of cognitive resources to task relevant processing plays an important role in exploration, we can expect that other factors that influence the availability of resources and the ability to work systematically on a task, moderate relationships between fatigue and exploration. We argue that experience is such a moderator. In complex tasks, demands on cognitive resources during the initial learning phase are high because novices have to guide every step in the problem solving process consciously (Anderson, 1982). However, with growing experience, people develop action procedures that can be executed in a fairly automatic way and that do not require a high level of cognitive resources, such as attention (Anderson1982; Norman & Shallice, 1986). As a result, experienced people work more efficiently and are better able to plan their behavior and to interpret feedback. Thus, compared to low experienced people, experienced people will use more systematic exploration (Hypothesis 7a), and less unsystematic exploration such as trial-and-error and rigid behavior (Hypothesis 7b). Moreover, compared to low experienced people, experienced people will achieve more sub-goals (Hypothesis 8a) and make fewer errors (Hypothesis 8b). Such main effects of experience on exploration can be expected to be even bigger than the effects of mental fatigue on exploration as the effects of fatigue on behavior are often quite subtle (Broadbent, 1979; Hockey, 1997).

As experienced people can work on a task more efficiently and without excessive demands on cognitive resources, their behavior and performance may be less susceptible to the influence of sub-optimal states such as mental fatigue (Bainbridge, 1978). Stated differently, lapses in task engagement under fatigue may be less disruptive for experienced people as they efficiently can execute and plan their behavior even when task engagement is relatively low. Hence, we expect interactions between the effects of fatigue and the effects of experience on type of exploration behavior. Specifically, the changes in exploration behavior under fatigue will be less strong for experienced people compared to novices (Hypothesis 9).

Method

Participants

Sixty-eight psychology students participated in this study for additional study credits. Participants were randomly assigned to a fatigue or control group. None of the participants had experience with the computer program Excel, which we used in the experimental task.

Measures

Fatigue. Subjective fatigue was measured with the general activation subscale of the Activation-Deactivation Checklist (AD-ACL, Thayer, 1989, Cronbach's alpha = .81) and with the Rating Scale Mental Effort (RSME, Zijlstra, 1993, Cronbach's alpha = .86). Although, the RSME is often used as a single measure of fatigue, it consists of seven 150-point answer scales about several fatigue aspects. Two items relate to mental fatigue, two items to physical fatigue, the remaining items measure resistance against further effort, boredom, and visual fatigue.

General Computer Experience. We assessed general computer experience with five questions in a 5-point Likert Scale format (Cronbach's alpha = .83). The questions concerned the frequency of computer use during the last year and experience with a range of computer applications (e.g. Word, Windows, MacOS). In the analyses, participants were assigned to a high or low general experience group, based on a median split procedure.

Fatigue manipulation. We used a so-called scheduling task (Taatgen, 1999) on the computer as fatigue manipulation. In this task, participants assigned work time to fictional employees. Adequate planning in this task required strong task engagement as previous planning steps had to be kept in mind –taking notes was not allowed-, while participants simultaneous had to think on further planning steps.

Computer Tasks. Participants worked on two different tasks on a MacIntosch computer, namely a task with the spreadsheet program Excel (to test our hypotheses) and a task with the graphical program ClarisDraw (for practice purposes).

The Clarisdraw Task was introduced to make participants familiar with the exploration method, to allow them to practice thinking aloud (see Procedure section), and to instruct them on how to adequately think aloud. In the Clarisdraw task participants had to reproduce (draw) an example figure that was presented on the screen.

The Excel Task was a task with the software program Excel (version 4.0 for Mac). The Excel task was given directly after the manipulation and was used to study exploration. The overall goal in the task consisted of changing the format of a table on the screen according to an example, which was presented in printed form. The task consisted of eight sub-tasks such as moving or adding text, changing text alignment, and adding and coloring table rows. These sub-tasks were not explicitly mentioned in the instructions nor was instructed how to approach the task. Task instruction only mentioned that the table on the screen should look like the example table. For each experimental session, appearance and settings of Excel were standardized so that each participant would start out in the same environment in which only the worksheet, the standard toolbar and the formatting toolbar were visible. Participants could freely explore the program. An exception was the Help-function, which we had disabled to reduce behavioral freedom, hence simplifying the coding of behavior. Although disabling Help makes the experiment somewhat artificial compared to real life settings this had no substantial consequences for our study (which looks at exploration of a system without step-by-step instructions).

Thinking Aloud

We used a thinking aloud procedure (Ericsson & Simon, 1993). Participants had to verbalize their thoughts while working on the tasks. Although, verbalizations do not cover all cognitive processes during the tasks (e.g. sometimes people omit information and some processes cannot be verbalized) they have shown to provide useful indications about goals and intentions underlying behavior (Ericsson & Simon, 1993).

Procedure

Participants were tested individually in sessions that lasted approximately three hours. At the beginning of the session participants filled in questionnaires on general computer experience and level of fatigue (see Measures section). Thereafter they worked on the ClarisDraw task for 15 minutes. Participants had to think aloud and when necessary the experimenter clarified and corrected the thinking aloud procedure during this task. When participants stayed quiet for several seconds the experimenter asked them to “keep on thinking aloud please” (Ericsson & Simon, 1993). After the ClarisDraw task, participants in the fatigue condition continuously worked on the scheduling task for two hours. Participants in the control condition were told to wait for two hours in which they could read magazines or watch videos. After the manipulation participants again filled in fatigue questionnaires. All participants then worked on the Excel task for 15 minutes in which they had to think aloud again. The computer screen of the participants was directly connected to a video recorder that recorded all their actions and verbalizations.

Coding of the Data

Exploration behavior in the Excel task was coded from the videotapes which contained the participants’ behavior on the computer screen and their verbalizations. Coders were blind to experimental condition and to the participants’ score on the general experience questionnaires. We used three behavioral categories that represented systematic or unsystematic exploration (see Introduction), namely, systematic exploration, unsystematic trial-and-error, and rigid behavior (description of categories below). In addition, we had one category for coding non-exploratory behavior. In complex tasks, such as our exploratory computer task, there are often no easy observable begin- or end points, therefore, we decided to use fixed time intervals of 20 seconds as coding units. First, coders decided whether a 20-second interval contained exploration behavior. If that was the case, the coders assessed whether exploration behavior fell into one of the three exploration categories.

From the sixty-eight videotapes analyzed, we randomly assigned twenty videos to each of two independent coders. We used intra-class correlations (ICC) as reported by Shrouth and Fleiss (1979) to assess interrater reliability. ICCs for the categories are reported below (60 < ICC < 75 = good interrater agreement, ICC > .75 = excellent interrater agreement, Ciccetti & Sparrow, 1981).

Coding Categories

Systematic exploration. Exploration was coded as systematic when the participants tried out functions or ideas in a structured and coherent way. This was the case when participants' actions either followed from explicit plans (e.g. "I will now try to use this same function for changing this part here..") or naturally followed from the previous actions (e.g. changing the settings of an option and then try that option again). Moreover, if coders found participants to evaluate what happened then such behavior was coded as systematic exploration too (e.g. "okay..this function serves to.."). ICC for systematic exploration was .75.

Unsystematic trial-and-error. Behavior was coded unsystematic trial-and-error if the participants did not show signs (either in terms of actions or of verbalization) of reflection or of feedback evaluation. For example, when the participants quickly 'jumped' from one option to another, showing no signs that options were well considered before going to the next one. ICC for this category was .83

Rigid behavior. Behavior was coded as rigid if the participants repeated (more than twice; Trudel & Payne, 1995) the same action sequence that already turned out to be unsuccessful in earlier attempts. Alternatively, behavior was coded rigid when participants continued to come back to the same options despite accumulating evidence that that option did not work. These criteria for coding rigid behavior in the computer task largely resembled perseverative behavior as assessed in traditional psychological tests that are used to diagnose deficits in the regulation of attention (e.g. Heaton, 1985). ICC for this category was .76.

Although coding of exploration behavior was done with fixed time intervals, coders took information from long-term goals and task context into account (e.g. actions belonging to a specific hypothesis could be executed over several consecutive 20-second periods). In this way we combined information about goals, intentions, and task context, with a fairly fine-grained analysis of behavior.

Non-exploratory behavior. In the Excel task, participants obviously also displayed non-exploratory behavior. For example, from their experience or as a result of exploration, participants discovered procedures in Excel, which they then applied to fulfill the task (e.g. changing the border around several parts of the table). Such application behavior often covered several 20-seconds coding periods. Application behavior was coded under a separate category (ICC = HAVE TO CHECK). In addition, there also was a proportion of behavior that was not exploration, nor application. This behavior was placed under a Residue-category (see Table 2).

Performance.

In task analysis of the Excel assignment we determined eight sub-tasks that had to be accomplished (see task description). An important performance variable was the number of sub-tasks solved within the time given. We also counted number of errors. Errors were defined as actions with negative consequences or actions that had no effect at all.

Results

Due to the relatively strong statistical power required to detect multivariate effects, we adopted an α of .10 for multivariate tests. For all univariate and post-hoc tests an overall α of .05 was used. Eta-squares (η^2) are reported as effect size.

Manipulation Check

Analyses of the RSME-scores confirmed that our manipulation was successful as there was a significant interaction effect of time of measurement (pre- versus post manipulation) by condition (fatigue versus not, see Table 1). Moreover, post-hoc tests showed the fatigue group to report significantly higher levels of fatigue after, but not before the manipulation (Table 1). A similar pattern of results was found for the general activation measure. To investigate in more detail, which aspects of fatigue were affected by our manipulation we analyzed the different fatigue aspects of the RSME separately. This analysis showed that our manipulation successfully induced mental fatigue and an increased resistance against further effort in the fatigue group. Compared to the control group, the fatigue group also reported significantly higher levels of boredom. Groups did not significantly differ in physical fatigue. The interaction of time of measurement by condition for visual fatigue was significant although non of the post-hoc comparisons reached significance (Table 1).

----insert Tabel 1 about here-----

Exploration Behavior

To test whether fatigue coincided with changes in exploration behavior and whether general computer experience moderated such effects, we submitted the three categories of exploration behavior to a Multivariate Analysis of Variance (MANOVA) with condition (fatigue vs. not) and experience (high vs. low) as independent variables. This revealed a significant multivariate main effect for condition ($\underline{F}$ (3, 62) = 4.22, $\underline{p}$ = .009) and a significant main effect for experience ($\underline{F}$(3, 62) = 4.50, $\underline{p}$ = .006). The multivariate interaction effect for condition by level of experience also reached significance at an α = .10 level ($\underline{F}$(3, 62) = 2.38, $\underline{p}$= .08). Univariate tests showed systematic exploration to decrease under fatigue (Hypothesis 1). We found a significant main effect for condition on systematic exploration ($\underline{F}$ (1, 64) = 3.97, MSE = 18.81, η^2 = .06, $\underline{p}$ = .050) with fatigued participants using significantly less systematic exploration ($\underline{M}$ = 7.23) compared to the control group ($\underline{M}$ = 9.06). There was also a main effect for general computer experience (Hypothesis 7a), with highly experienced participants using more systematic exploration ($\underline{F}$ (1, 64) = 10.39, MSE = 18.81, η^2 = .14, $\underline{p}$ <.01, $\underline{M}$ = 9.54 and $\underline{M}$= 6.21, respectively). There was no interaction effect for fatigue and experience on systematic exploration ($\underline{F}$ (1, 64) = 9.48, MSE = 18.81, η^2 = .01, $\underline{p}$ = .48, see Table 2 for means) indicating that fatigue had an effect on systematic, reflective behavior for both high and low general computer experience participants.

------Table 2 about here-----

Hypothesis 2 stating increase of unsystematic trial-and-error under fatigue was not confirmed as there was no significant main effect for condition ($\underline{F}$ (1, 64) = 1.52, MSE = 55.35, η^2 = .02, $\underline{p}$ = .22) and no significant interaction effect (Hypothesis 5a, $\underline{F}$ (1, 64) = .11, MSE = 55.35, η^2 = .002, $\underline{p}$ = .75). The main effect of general computer experience was significant, showing that, compared to low experienced participants, highly experienced participants used a lower level of unsystematic trial-and-error ($\underline{F}$ (1, 64) = 6.00, MSE = 55.35, η^2 = .09, $\underline{p}$ < .05, Hypothesis 7b, $\underline{M}$= 6.95 and $\underline{M}$= 11.31, respectively).

As expected, fatigue coincided with increased rigid behavior (Hypothesis 3). There was a significant main effect for condition on rigid behavior ($\underline{F}$ (1, 64) = 10.92, MSE = 7.33,

η^2 = .15, p < .01, M= 4.91 and M= 3.06, respectively). There was also a significant main effect for general computer experience (F (1, 64) = 4.71, MSE = 7.33, η^2 = .07, p < .05, Hypothesis 7b), with highly experienced participants using less rigid behavior (M = 3.44) than low experienced participants (M= 4.79). As the interaction effect for condition and level of general computer experience also was significant this confirmed the hypothesis that experience moderates the effects of fatigue on exploration (F (1, 64) = 7.33, MSE = 7.33, η^2 = .10, p < .01, Hypothesis 9). Post-hoc t-tests showed the mean of low experienced, fatigued participants to be significantly different from all other groups (see Table 2). This implies that fatigue leads to increased rigidity only when people are highly unfamiliar with the task at hand.

Analyses of non-exploratory behavior (the Application and Residue categories) showed that there were no significant main effects of condition or experience on either of these two categories nor were there any significant interaction effects of condition and experience (p values ranged from .19 to .79 for these analyses).

Check on number of verbalizations. To check for possible effects of fatigue on verbalizations, which might have effected the coding of exploration behavior, we analyzed the number of verbal statements of fatigued and non-fatigued participants. From a random sample of the participants (approximately 70 percent, 24 fatigued and 23 non-fatigued participants) we counted the number of statements within a 5-minute segment of the videotape. Analyses showed that fatigued and non-fatigued participants did not differ on the number of verbalizations during the Excel task (F(1, 47) = .01, η^2= .00, p = .91, Mean number of statements: non-fatigued versus fatigued participants were respectively M = 45.96 and 45.53). Thus, differences in exploration behavior could not be explained by differences in the number of verbalizations.

Performance

There was no significant main effect for condition (fatigue vs. not) on the number of sub task solved (Hypothesis 6a, F (1, 64) = 1.24, MSE = 6.00, η^2 = .02, p = .27). However, there was a significant effect for condition on number of errors (Hypothesis 6b); fatigued participants made significantly more errors (F(1, 64) = 4.77, MSE= 61.15, η^2 = .07, p < .05, means fatigue vs. not respectively, 23.11 and 19.57). Compared to low experienced participants, experienced participants solved significantly more sub-tasks (F(1, 64) = 7.79, MSE = 46.75, η^2 = .11, p = .007, Hypothesis 8a, M= 2.38 and M = 4.03, respectively) and made marginal significant fewer errors (F(1, 64) = 3.80, MSE = 232.45, η^2 = .06, p = .056, Hypothesis 8b, M = 23.45 and M =19.87, respectively). There were no significant interaction effects between condition and experience on number of sub-tasks solved (F(1, 64) = .11, MSE = 6.00, η^2 = .002, p = .74) and errors (F(1, 64) = 2.59, MSE = 61.15, η^2 = .04, p = . 11).

Fatigue, Exploration and Errors

To test relationships between exploration on the one hand, and number of sub-tasks solved and errors on the other hand (Hypotheses 4a and Hypothesis 4b), we first analyzed whether the correlations between these variables significantly differed between the different sub-groups (Fatigue vs. not and high vs. low general experience). To test this we used the Fisher r-to-Z procedure, which showed none of the correlations significantly differed between the sub-groups. Therefore, we report overall correlations, collapsed across all sub-groups. These correlations confirmed Hypothesis 4a as systematic exploration was positively correlated with number of sub-tasks solved (r (68) = .58, p < .001) and negatively with number of errors (r (68) = -.41, p < .001). Our hypotheses on the negative relationship between the use of unsystematic exploration and number of sub-tasks solved (Hypothesis 4a) and the positive relationship of unsystematic exploration with errors (Hypothesis 4b) were also confirmed. (trial-and -error/sub-tasks r (68) = -.43, p < .001, trial-and-error/errors r (68)=

.59, $p < .001$, rigidity/sub-tasks $r(68) = -.28$, $p < .05$, rigidity/errors $r(68) = .40$, $p = .001$). We also looked at the overall correlation between the number of errors and number of solved sub-tasks, which revealed that a high number of errors was related to a low number of sub-tasks solved ($r(68) = -.50$, $p < .001$.

In this study, we also looked at non-exploratory application behavior, which refers to the use of Excel procedures to fulfill the task. Such application often was the outcome of successful exploration in which the participants found out how to attain a sub-goal (e.g. “oh..so, this is the way to align text”). Thus, application behavior was related to performance. More specifically, application was correlated with sub-tasks solved ($r(68) = .34$, $p < .05$) and errors ($r(68) = -.66$, $p < .05$). To determine whether exploration stayed associated with performance measures beyond application, we also looked at partial exploration-performance correlations (controlling for application behavior). These analyses showed that the exploration-performance relationships stayed significant and did not notably change (exception was the rigid behavior, sub-tasks correlation which changed form $r = -.28$ to $-.18$, $p = .14$). Thus, for five of the six relationships tested, the relationship between exploration and performance was not accounted for by non-exploratory application behavior. Therefore, the overall results are consistent with the literature on systematic versus unsystematic problem solving behavior (e.g. Hollnagel, 1993, Trudel & Payne, 1995; van der Linden et al. 2001), and showed the different types of exploration behavior to be strongly related to performance on the Excel task. The amount of behavior in the residue category did not show any significant correlations with any other variable in this study.

Discussion

The main research question in our study was whether fatigues changes the way people explore a complex system. The present study showed this was the case; compared to non-fatigued participants, fatigued participants showed significantly fewer periods of systematic exploration. This finding suggests that fatigued participants were less thoughtful and reflective in their exploration behavior (Trudel & Payne, 1995). The results on unsystematic behavior were mixed; while fatigue did not coincided with significant changes in unsystematic trial-and-error, there was a significant effect on rigid behavior. Sub-group analyses showed that only participants with a low level of general computer experience displayed more rigid behavior under fatigue. Specifically, these low-experienced fatigued participants showed more inefficient perseveration and had the highest tendency to come back to the same, unsuccessful options. This finding was in accordance with our hypothesis that the effects of fatigue on exploration are stronger for people for whom the task has a high level of unfamiliarity.

Despite the fact that fatigue was related to exploration behavior, which in turn was related to number of sub-tasks solved, we did not find significant differences in number of sub-tasks solved between fatigued and non-fatigued participants. Although, admittedly somewhat speculative, this non-finding may be related to the relatively short duration of the Excel tasks (15 minutes). Within such a short time span, small differences in exploration behavior may not directly lead to a major (significant) breakdown of overall performance. This is a common finding in many fatigue studies which have shown that fatigue can lead to observable changes in behavior long before primary task performance starts to deteriorate (Holding, 1984; Hockey, 1997; Sanders, 1998). On the other hand, we found that, compared to non-fatigued participants, fatigued participants made more errors. These findings suggest that the frequency of making errors is more easily affected under fatigue than the primary task output (solving subtasks).

One important question of interpretation is what might have caused the changes in exploration behavior under fatigue. We assumed that under fatigue there is an increase of periods in which task engagement is temporary lowered. During such lapses of engagement, fatigued people do not invest a high level of resources into thoughtful reflection or planning. Hence, exploration during those periods will be less systematic. It is important to note that any reduction in task engagement under fatigue does not necessarily have to involve conscious decisions to do so (Meijman, 2000). Specifically, fatigued people might experience lapses in task engagement even when they, at an intentional level, want to perform well on the task.

Theoretically, we want argue that under fatigue, periods of reduced task engagement coincide with compromised executive control on behavior (cf. Monsell & Driver, 2000). Executive control refers to attentional processes that regulate perception and motor processes in order to ensure goal-directed behavior (Miller & Cohen, 2000, Norman & Shallice, 1986). Such executive control plays a major role in systematic behavior, based on planning and on cognitive flexibility (e.g. reacting to feedback). Moreover, one of the hallmarks of compromised executive control is reduced flexibility of behavior and the subsequent increased tendency to perseverate, particularly when a task is unfamiliar (cf. Miller & Cohen, 2000). In the current study, the overall reduction in goal-directed, systematic exploration and the increased tendency to perseverate for low-experienced participants indeed indicate such compromised executive control in fatigued participants. Moreover, these results are in accordance with other studies, which found fatigued participants to perform poorly on tasks that specifically tap executive control processes (van der Linden, Frese, & Meijman, in press). Future studies may want to directly combine measures of executive control with behavior in more applied tasks as in the current study.

Limitations of the Study

Our study used a detailed analysis of exploration behavior, based on video recordings, which also contained participants' verbalizations. This method provided us with a rich source of data from which we could infer whether participants explored systematically or unsystematically. Nevertheless, in spite of their richness, our methods also had some limitations. One of these limitations refers to our focus on exploration behavior. Exploration plays a major role in dealing with complex tasks, yet, there are also many other types of behavior people might display. In the current study we captured much of these other types of behavior with a category labeled ‘application’. A more detailed analysis of such non-exploratory behavior and their interactions with exploration and performance may provide additional insight into the way people deal with (or learn) complex tasks. Nevertheless, because the fatigue and control group did not differ in overall amount of exploratory versus non-exploratory behavior and because exploration remained significantly related to performance also when controlling for non-exploratory behavior, this limitation does not compromise the study’s results or their interpretation.

Another limitation refers to the possible influence of fatigue on thinking aloud. However, two lines of arguments suggest that our results did not arise from mere verbalization effects. First, fatigued and non-fatigued participants did not differ in number of statements. This result is in accordance with findings of a pilot study in which we found very low correlations between pre-task measures of fatigue (RSME and general activation from the AD-ACL) and number of statements during a similar Excel task (r = -.02 and r = .05, respectively). Second, assessment of exploration behavior did not exclusively rely on verbal reports but also on participants’ actions (e.g. repetition of the same unsuccessful responses). Exploration in terms of systematic/unsystematic was inferred from the “…broader context of behavior” (Shrager & Klahr, 1986).

Practical implications

As fatigue to coincided with changes in exploration behavior, the results of this study have consequences for practice. Specifically, the results suggest that under fatigue, hypothesis generating and testing behavior (as in the Excel task) has a decreased likelihood of being systematic and coherent. Moreover, in an unfamiliar setting or task environment –as for the participants with little experience-, behavior under fatigue tends to be guided by salient cues in a rigid way. System designers should be aware of such tendencies because any not-well considered salient aspect of the human-computer interface may capture attention of uses under fatigue and subsequently may yield erroneous ideas or solutions. For example, in the current study, many participants were often misled by the label ‘Color’ in the option ‘Color Palette’. Because this label corresponded with a task goal (“change color of the table”), many participants initially thought this was the relevant option, when in fact it wasn’t. However, due to the effects of fatigue, low-experience participants seemed to have difficulties in detaching from such erroneous ideas and subsequently showed increased perseveration and rigidity. These findings show that system designers should construct the interface in such a way that labels are congruent with task goals, thereby reducing the need for users to engage in extensive reasoning. Moreover, obvious relationships between task goals and interface features can be expected to lead to less rigidity and errors.

As we argued that the effects of fatigue on exploration are related to general effects on information processing (e.g. executive control), the practical implications of the current study may also extend to different settings than the specific learning context we used. Specifically, it can be expected that similar manifestations of mental fatigue may be found in other situations in which people must deal with partial information, systematically have to generate and test hypotheses, and have to be flexible. A typical example of such situation is finding the fault underlying a system failure or incident. In such cases, not being systematic or using a rigid approach may have strong negative consequences. For example, in the Three-Mile-island nuclear plant incident, engineers who worked at a time when people generally feel fatigued, repeatedly failed to use an optimal exploration strategy to find the fault and repeatedly discarded useful information in a rigid way. Although the direct link between the behavior of those engineers and the results of this study is somewhat speculative, it illustrates the severity of consequences that rigidity and a lack of systematic behavior may have in real-life settings. This also suggests that it might be inefficient to rely on fatigued people to optimally generate and test hypotheses. When exploration is required and important (as in fault finding), it might be better to rely upon the exploratory abilities of well-rested people. In practice, this may be implemented by using back-up teams who are deployed when a situation requires systematic search of alternatives. Such teams may replace or support teams who are at the end of their shifts or have already worked for several hours during the night.

The findings of the current study also suggest another countermeasure against adverse manifestations of fatigue. Specifically, whereas low-experience participants showed more rigid behavior under fatigue, this was not the case for experienced participants. Hence, training and extended experience may to some limit reduce the detrimental effects of fatigue on exploration. Obviously, it is not possible to train people for every unforeseen situation that requires exploration. However, training in general heuristics and procedures of good exploration might reduce the need to think on good exploration strategies at times when such thinking is sub-optimal (as under fatigue). Specifically, such training should teach the steps of systematic exploration (generating specific goals and action plans to attain them, deliberate evaluation of action outcomes, Frese, 19..) and make people aware of the possible consequences of adverse conditions (e.g. fatigue) on exploration. Well-trained general principles of exploration and knowledge about the ‘dangers’ of sub-optimal energetic conditions may be tools that can help to prevent or overcome fatigue effects.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Bainbridge, L. (1978). Forgotten alternatives in skill and work-load. Ergonomics, 21, 169-185.

Broadbent, D. E. (1979). Is a fatigue test now possible? Ergonomics, 22, 1277-1290.

Chi, M. T. H., Bassok, M. Lewis, M. W., Reimann, P., and Glaser, R., 1989. Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing the interrater reliability of specific items in a given inventory: Applications to assessment of adaptive behavior. American Journal of Mental Deficiency, 86, 127-137.

Dörner, D. (1980). On the difficulties people have in dealing with complexity. Simulation & Games, 11, 87-106.

Ericsson, K. A., and Simon, H. A., 1993. Protocol analysis: Verbal reports as data (rev. ed). (The MIT Press, Massachusetts).

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), Complex problem solving: Principles and mechanisms. (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Green, A. J. K. & Gilhooly, K. J. (1990). Individual differences and effective learning procedures: The case of statistical computing. International Journal of Man-Machine Studies, 33, 97-117

Heaton, R. K. (1981). Wisconsin card sorting test manual. Odessa: University of Colorado.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. Biological Psychology, 45, 73-93.

Holding, D. (1983). Fatigue. In R. Hockey (Ed.), Stress and Fatigue in Human Performance (pp. 145-164). Durnham: John Wiley & Sons.

Hollnagel, E. (1993). Human reliability analysis: Context and control. London: Academic Press.

Meijman, T. F. (2000). The theory of the stop-emotion: On the functionality of fatigue. In D. Pogorski & W. Karwowski (Eds.), Ergonomics and safety for global business quality and production. (pp. 45-50). Warschaw: CIOP.

Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. Annual Review of Neuroscience, 24, 167-202.

Monsell, S., & Driver, J. E. (2000). Control of cognitive processes: Attention and performance XVIII. Cambridge: the MIT Press.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Swartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in theory and research. (Vol. 4, pp. 1-18). New York: Plenum Press.

Sanders, A. F. (1998). Elements of human performance. London: Lawrence Erlbaum Associates.

Shrager, J., & Klahr, D. (1986). Instructionless learning about a complex device: The paradigm and observations. International Journal of Man Machine Studies, 25, 153-189.

Schrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Somsen, R. J. M., van der Molen, M. W., Jennings, J. R., & van Beek, B. (2000). Wisconsin Card Sorting in Adolescents: Analysis of performance, response time, and heart rate. Acta Psychologica, 104, 227-257.

Taatgen, N. A. (1999). Learning without limits. From problem solving towards a unified theory of learning. Unpublished dissertation, University of Groningen, Groningen, The Netherlands.

Thayer, R. E. (1989). The biopsychology of mood and arousal. New York: Oxford University Press.

Trudel, C. I., & Payne, S. J. (1995). Reflection and goal management in exploratory learning. International Journal of Human Computer Studies, 42, 307-339.

Van der Linden, D., Sonnentag, S. Frese, M. & van Dyck, C. (2001). Exploration strategies, error consequences, and performance when learning a complex computer task. Behaviour and Information Technology, 20, 189-198.

Van der Linden, D., Frese, M. & Meijman, T. F. (in press). Mental fatigue and the control of cognitive processes: Effects on perseveration and task engagement. Acta Psychologica.

Zijlstra, F. R. H. (1993). Efficiency in work behavior: A design approach for modern tools. Delft: Delft University Press.

Table 1. Pre- and post manipulation measures of fatigue

		Pre-manipulation			Post-manipulation		
	Time X Condition F(1, 64)	Control (n= 32)	Fatigue (n = 36)	t	Control (n = 32)	Fatigue (n= 36)	t
RSME (total)	16.39**	24.07 (11.15)	23.40 (13.22)	.40	26.57 (16.88)	45.01 (24.94)	3.58**
General activation	7.16*	12.64 (2.16)	12.26 (2.59)	.54	11.21 (2.32)	9.47 (2.28)	3.24**
Mental fatigue	25. 64**	53.12 (25.12)	48.74 (32.37)	.62	48.12 (37.26)	97.88 (56.96)	4.26**
Physical fatigue	.96	59.21 (39.10)	60.97 (42.11)	-.1.20	73.46 (52.05)	88.47 (51.47)	-.09
Resistanc to effort	22.94**	25.18 (14.50)	25.03 (21.27)	.10	14.33 (16.66)	46.82 (38.04)	4.56**
Boredom	7.54*	18.46 (14.34)	12.77 (19.84)	-.68	14.76 (27.54)	21.50 (39.16)	-3.18*
Visual fatigue	5.88*	14.88 (15.62)	15.65 (11.19)	1.34	35.30 (15.87)	60.41 (26.95)	-1.64

Table2. Means (and Standard Deviations) of strategy-use and performance for the different groups.

	Non-Fatigued		Fatigued	
	Low experience (n = 14)	High Experience (n = 18)	Low experience (n = 15)	High experience (n = 21)
Exploration behavior measures				
Systematic exploration	7.60 (5.80)	10.28 (4.78)	4.71 (3.38)	8.09 (3.13)
Unsystematic trial-and-error	9.93 (9.08)	6.06 (4.68)	12.79 (10.0)	7.71 (5.11)
Rigid behavior	2.87 (2.36)	3.22 (3.02)	6.86 (2.74)	3.62 (2.62)
Performance measures				
Number of sub-tasks solved [a]	2.80 (1.97)	4.28 (3.27)	1.93 (1.54)	3.81 (2.42)
Number of errors	19.93 (7.01)	19.28 (6.51)	27.21 (7.78)	20.38 (9.28)
Non-exploratory behavior				
Application Behavior	21.60 (8.89)	22.83 (3.89)	18.42 (7.79)	20.33 (5.35)
Residue	3.00 (4.27)	2.61 (3.92)	2.21 (3.14)	4.42 (4.87)

Note. Means denote the frequency with which a strategy occurred during the task.

[a] number of sub-tasks solved (Range from 0 to 8)