Available online at www.sciencedirect.com

SciVerse ScienceDirect

Int. J. Human-Computer Studies 70 (2012) 156–168

International Journal of Human-Computer Studies

www.elsevier.com/locate/ijhcs

Juggling on a high wire: Multitasking effects on performance

Rachel F. Adler[a], Raquel Benbunan-Fich[b,*]

[a]The Graduate Center, City University of New York, United States
[b]Baruch College, City University of New York, United States

Received 24 January 2011; received in revised form 15 August 2011; accepted 21 October 2011
Communicated by S. Wiedenbeck
Available online 29 October 2011

Abstract

In this study, we develop a theoretical model that predicts an inverted-U relationship between multitasking and performance. The model is tested with a controlled experiment using a custom-developed application. Participants were randomly assigned to either a control condition, where they had to perform tasks in sequence, or an experimental condition, where they could discretionarily switch tasks by clicking on tabs. Our results show an inverted-U pattern for performance efficiency (productivity) and a decreasing line for performance effectiveness (accuracy). The results of this study indicate that the nature of the relation between multitasking and performance depends upon the metric used. If performance is measured with productivity, different multitasking levels are associated with an inverted-U curve where medium multitaskers perform significantly better than both high and low multitaskers. However, if performance is measured with accuracy of results, the relation is a downward sloping line, in which increased levels of multitasking lead to a significant loss in accuracy. Metaphorically speaking, juggling multiple tasks is much more difficult while balancing on a high wire, where performance mishaps can have serious consequences.

Keywords: Memory for goals; Multitasking; Task switching; Performance

1. Introduction

Is it really possible to juggle while walking on a high wire? Human multitasking has reached new heights these days and its effects on performance are not yet clear. Some researchers underscore the negative impacts of multitasking on task performance, particularly when multiple tasks interfere with each other and this juggling produces obstructions or distractions (Aral et al., 2006). Alternatively, others argue that multitasking can be conducive to better outcomes by promoting a more productive use of time, allowing ideas to mature or encouraging healthy breaks from complex tasks (Madjar and Shalley, 2008). Despite this debate, juggling multiple tasks with and without technological devices is a common practice at home, at school, at work and even during meetings (Benbunan-Fich and Truman, 2009; Gonzalez and Mark, 2004; Hembrooke and Gay, 2003; Mark et al., 2005; Wasson, 2004).

To a great extent, modern technology platforms enable and amplify multitasking. For example, contemporary operating systems in personal computers are designed to accommodate multitasking by allowing users to work on more than one task concurrently. Even current web browser interfaces provide tabbing capabilities to facilitate the performance of simultaneous web-based activities. Given the pivotal role of Information Technology in multitasking, our research is focused on *computer-based* multitasking behavior. We are particularly interested in the consequences of performing several unrelated computer-based tasks with a single technological device and in a specific period of time. By concentrating exclusively on computer-based tasks, and excluding all other concurrent tasks, such as for example eating or watching TV while using the computer, we seek to shed some light on the patterns and consequences of multitasking behavior.

Investigating multitasking behavior is particularly important for Human–Computer Interaction (HCI) researchers.

*Corresponding author. Fax: +1 646 312 3351.
E-mail addresses: RachelFAdler@gmail.com (R.F. Adler), rbfich@baruch.cuny.edu (R. Benbunan-Fich).

1071-5819/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2011.10.003

Although this is not a new area of research in HCI, it has not been the focus of intensive study (McCrickard et al., 2003b). This gap is more noticeable given the prevalence of multitasking behavior nowadays. Because of the popularity of multitasking, a deeper understanding of this behavior is poised to advance the HCI literature. Studying the performance consequences of computer-based multitasking can offer new insights to inform practitioners about optimal work arrangements and to improve systems that handle notification systems (McCrickard et al., 2003a) and support multitasking.

Our research question investigates the performance effects of different multitasking patterns. To this end, we develop a theoretical model predicting that multitasking behavior is beneficial until a point of diminishing returns in which too much multitasking begins to have a detrimental effect on performance. We test this prediction in a laboratory experiment using a custom-developed application with several tasks. Our results contribute to elucidate the complex relationship between multitasking and performance. To describe the study and its findings, the remainder of this paper proceeds as follows. The next section introduces the concept and possible strategies for multitasking. This conceptual introduction is followed by the theoretical model, whose foundation comes from the theory of memory-for-goals (Altmann and Trafton, 2002) and the Yerkes–Dodson law (Yerkes and Dodson, 1908). After the theoretical model, we develop the hypothesis and describe our research. The remaining sections address data analysis, results, discussion and limitations. The paper concludes by presenting the contributions of this study and outlining future research directions.

2. Theoretical background

Multitasking occurs when a user shifts attention to perform several independent but concurrent computer-based tasks. Benbunan-Fich et al. (2011) articulate two key principles to define multitasking, namely, task independence and performance concurrency. While the principle of *independence* suggests that ongoing tasks are self-contained, the notion of *concurrency* implies that these multiple tasks are carried out with some temporal overlap in a specific period of time. Depending on the amount of overlap, multiple tasks can be temporally organized following three different strategies, namely, sequential, parallel and interleaved. Each of these strategies implies a different degree of concurrency among ongoing tasks.

In a *sequential* strategy, each task starts after the completion of the previous task (Bluedorn et al., 1992). Since in this mode the level of concurrency is zero, it can be argued that this strategy is not consistent with the typical conceptualization of multitasking where individuals are juggling several ongoing activities. Although there is no task overlap, the sequential approach can be useful to establish a baseline case for comparing performance effects with other multitasking approaches.

In a *parallel* strategy, all concurrent tasks are attended to at the same time (Bluedorn et al., 1992). As such, there is a maximum degree of concurrency or overlap. In practice, however, true parallel performance is difficult to achieve because human attention cannot be simultaneously divided among many tasks, unless different types of attention are required as in the case of writing and listening to music at the same time (Salvucci and Taatgen, 2011). When attention shifting occurs at the cognitive level, it appears as if the two tasks are carried out in parallel. Given our interest in computer-based multitasking, we focus on situations where task switching is observable.

In an *interleaved* strategy, a task underway is voluntarily or involuntarily suspended to allocate attention to another task. Eventually, the original task is resumed but once again it is willingly abandoned or externally interrupted to attend to other tasks (Payne et al., 2007). The most common manifestation of this multitasking strategy is through task interleaving or task switching. This form of multitasking where different tasks are interleaved or interspersed is the most typical for computer-based activities.

Instead of categories, a more recent view of multitasking articulates the differences in terms of a continuum by noting the typical time spent on one task before switching to the other (Salvucci and Taatgen, 2011). At one extreme of the multitasking continuum there are tasks that involve frequent switching, perhaps every second or more often as in a normal conversation (i.e. talking while driving). This could be characterized as concurrent multitasking because the tasks are, in essence, performed at the same time. This characterization is similar to our description of parallel task performance presented above. At the other extreme, there are tasks that involve fairly long spans between switches. This could be characterized as sequential multitasking because a longer time (measured in minutes or even hours) might be spent on one task before switching to another. Unlike the non-multitasking-sequential case described above, in this situation, tasks have not been completed when there is a switch and therefore, this is an instance of multitasking. In this view, concurrent and sequential multitasking can be represented on the same continuum. Since these are the extremes of the continuum, our description of the task interleaving presented above, can be interpreted as representative of the situations in the middle of the multitasking continuum.

2.1. Memory for goals

Cognitive psychologists have proposed several theories to explain the mental processes that account for decreased task performance in multitasking situations (Meyer and Kieras, 1997). This body of literature explains the relationship between task concurrency and performance, primarily in the context of a dual task paradigm. Although the psychology literature has devoted comparatively less attention to user-defined tasks and goals and voluntary task switching, there are theories particularly well-suited to explain the relation between multitasking and performance. One of these is "memory-for-goals" (Altmann and Trafton, 2002). In this theory, a goal is defined as "a mental representation of an intention to accomplish a task, achieve some specific state of the world, or take some mental or physical action" (p. 39). The memory-for-goals theory posits that in order for people to initiate a new task, its goal must be strengthened in memory to the point where its activation rises above other goals (Salvucci et al., 2009). This theory explains hierarchical problem solving by allowing a high-level goal to be decomposed into subgoals. It also explains multiple task performance through the goal-activation process. Activation is the process whereby goals move up and become the focus of attention. The newest (or most recently activated) goal is the one that directs behavior, while the old goals are postponed (Altmann and Trafton, 2002).

Two different conditions cause active goals to be suspended or set aside temporarily in favor of new goals (Salvucci and Taatgen, 2011). The first possibility is an *external interruption* that requires immediate attention and produces a displacement of the current goal. This displacement results in a reorganization of goals in memory as people formulate the intention to resume the interrupted task later. The second possibility is a *voluntary decision to stop the current task (or break)* due to an obstacle that prevents its completion. This blockage causes the active goal to be suspended and allows for another goal to become the focus of attention until conditions change and the abandoned task can be resumed. From these two goal-displacement conditions (external interruptions and self-imposed interruptions), we are primarily interested in *self*-interruptions, where the empirical literature is sparse.

Due to the discretionary nature of self-imposed interruptions, the effects of voluntary task switching on performance are less clear than those documented in the interruptions literature, where there is an extensive body of work (Bailey and Konstan, 2006; Gillie and Broadbent, 1989; McFarlane, 2002; Speier et al., 2003). Some empirical studies have begun to shed light into the drivers of discretionary task interleaving, a necessary first step toward studying their relation with performance. For example, Payne et al. (2007) found that voluntary task switching is motivated by either the propensity to temporarily abandon a task that is no longer rewarding or by the tendency to switch to an unrelated task when a subtask is completed. In addition, Madjar and Shalley (2008) report that individuals had the highest creativity when they had the discretion to switch tasks, and each task had a specific goal. In the discretionary switching condition, the goals served to focus the participants' attention on working hard on the tasks, and the ability to switch at will afforded them the ability to take a break if needed and work on another task to refresh or clear their heads.

Memory-for-goals theory is useful to articulate two different conditions for goal displacement. In an internal self-initiated interruption, the user decides at her discretion to shift goals and perform a different task. The newly activated goal and associated task are entirely at the discretion of the user. However, in an external interruption the electronic notification indicates the new task and goal that require attention. The discretionary nature of self-interruptions poses an intriguing dilemma for analyzing the effects on performance. On the one hand, it is possible that some amount of task switching is better than no switching at all, as people get additional stimulation or arousal that leads to improved performance. On the other hand, it is also possible that swift task changes create an environment of confusion where an individual is unable to perform each task adequately due to overload or excessive arousal (Monsell, 2003; Palladino, 2007).

2.2. Yerkes–Dodson law

Research in psychology has found evidence of a curvilinear relationship between arousal and performance, which is typically known as the *Yerkes–Dodson law* (Yerkes and Dodson, 1908). According to this law, there is an optimal amount of arousal that leads to the best performance, while very low or extremely high levels of arousal lead to suboptimal performance. The law is attributed to Yerkes and Dodson because they were the first to document this inverted-U relationship. In their pioneer study in 1908, they found that when mice were given a difficult discrimination task their performance improved linearly with low and moderate levels of arousal. However, at the highest levels of arousal their performance was impaired forming an overall non-linear or inverted-U shaped relationship between arousal and performance. While there are numerous criticisms to Yerkes and Dodson (1908) – including the mice-to-man extension of their findings, their failure to systematically measure and control arousal in mice and the lack of statistical analyses – this work has been replicated with other living organisms and with people in different contexts (Miller, 1978; Staal, 2004).

In its initial formulation, the Yerkes–Dodson law intended to explain the relation between stimulus strength and habit-formation for varying degrees of task difficulty (Teigen, 1994). Over the last century, this law has been used to explain the effects of reward, motivation, arousal, and stress on learning, performance, problem-solving or memory (Bäumler, 1994). Teigen (1994) presents a discussion of the origins, evolution and applications of the Yerkes–Dodson law. Some researchers have even questioned the validity of the original law. In a critical review of the findings initially reported by Yerkes and Dodson, Bäumler (1994) reexamined the original data and concluded that this law perhaps never existed. Despite its detractors, the versatility of the law and the intuitive appeal of a curvilinear relationship have made it attractive to explain relations in various contexts and elevated its status to become a fundamental law in psychology (Teigen, 1994).

Empirical support for the inverted-U relation has been mixed depending upon the context in which it is applied.

For example, Muse et al. (2003) note that the relation has received scant support in the stress literature but ample support in the arousal/activation literature. Arousal refers to the psychological reactions that cause a person to be alert and attentive (Muse et al. 2003). At low levels, performance is impaired by a lack of alertness, while at high levels performance is characterized by a total disorganization of responses.

People's level of arousal fluctuates with different work demands. Increasing load produces a rise in arousal in order to mobilize cognitive resources (Kahneman, 1973). It can be argued that, at one extreme, under minimal amounts of arousal, people lack motivation to stay focused on the task at hand and their performance is likely to suffer. It is possible that a second task induces an increase in arousal to improve performance of the first task. However, at the other extreme, under high levels of arousal, humans might be unable to cope with the demands of the situation and their performance decreases. These extremes can be labeled as under-load and overload, respectively (Wiener et al., 1984). While under-load leads to a reduction in alertness and lowered attention, overload has an opposite effect, leading to distraction and interference, thus, the inverted-U relationship between arousal and performance. Performance is best in the middle of the curve, where people are in a relaxed-alert state (Palladino, 2007).

These differential effects of arousal on performance have been further explained in terms of information processing. The awareness of a pending task uses some resources because of "invisible" processing of irrelevant cues (Navon and Gopher, 1979). While it may not be detrimental at first, increased invisible processing of irrelevant cues, associated with pending tasks, is likely to produce performance degradation (Norman and Bobrow, 1975). With increased arousal, fewer cues are processed, excluding task-irrelevant cues, which help performance. However, at the highest levels of arousal, even task relevant cues are discarded, which results in impaired performance. Arousal increases attention and persistence but reduces the efficiency of information processing to a point where higher levels of arousal produce information processing deficiencies that result in suboptimal performance (Muse et al., 2003).

2.3. Hypothesis development

Multitasking situations present a resource-allocation problem to humans because they must decide how to assign limited resources (such as time or attention) across multiple competing tasks to achieve specific performance goals. In a completely voluntary switching environment, users exercise discretion on how and when to switch and whether to switch at all (Yeung, 2010). Therefore, they may choose to perform tasks sequentially, instead of interleaving ongoing tasks. From the perspective of memory-for-goals, sequential task execution implies that one and only one task goal exists at any given time. In contrast, interleaving situations require the activation of multiple goals and the simultaneous management of multiple problem representations. When people work on several tasks at the same time, their working memory helps them switch tasks by storing information related to the abandoned task and redirecting attention to the new task (König and Oberacher, 2010).

Awareness of multiple goals, as opposed to only one goal at a time, leads to increased levels of arousal. Working on several tasks at once requires people to maintain the temporary state information associated with each task (Borst et al., 2010). Goal displacement in memory as evidenced in task switching requires people to recall information associated with a previously suspended task to resume its performance (Payne et al., 2007). Constant switching among competing tasks distracts the user and degrades performance (Rubinstein et al., 2001). Switching tasks results in slower response times and increased error rates. The literature provides evidence of cognitive costs associated with switching between unrelated tasks (Payne et al., 2007). These costs are attributed to between-task interference, which typically arise from the residual tendency to keep thinking about the just-abandoned task that is now irrelevant (Yeung, 2010).

When people juggle many tasks simultaneously, they must remember what they left behind when they resume a previously abandoned task. Some task switching can be beneficial to increase levels of arousal and mobilize cognitive resources to handle increased levels of load. This increased arousal can result in improved performance (Navon and Gopher, 1979). However, under constant goal displacement and extreme high levels of arousal, any potential performance gain is lost due to the high costs of switching tasks and swapping associated problem state information in memory (Borst et al., 2010). At high levels, the failure to completely recall state information of the abandoned tasks impairs performance.

We posit that different multitasking situations have differential effects on performance. Thus, we distinguish different multitasking scenarios based on the amount of switching between ongoing tasks. In a sequential scenario, there is only one goal at a time and task interference is absent. In an interleaving task situation, some amount of task switching could be beneficial to induce an increased level of arousal that would lead to making a more productive use of time, or clearing one's head. However, a high amount of switching would be detrimental because of the difficulties associated with swapping problem representations in memory and the resulting task interference. Therefore, based on the *Yerkes–Dodson law*, an inverted-U relationship can be hypothesized where performance first improves with increases in multitasking levels, and then drops after reaching high levels of multitasking due to cognitive overloading and impairment resulting from task interference. More formally, we hypothesize

> Users who engage in multitasking will perform significantly better than those who are only slightly multitasking,

but there will be a point of diminishing returns where too much multitasking will significantly lower their performance.

Empirical evidence based on survey data has documented an inverted-U relationship between multitasking and output. For example, Aral et al. (2006) found that that the relationship between multitasking and output is non-linear. In their study, multitasking was defined at the project-level as "the act of taking on multiple simultaneous projects in parallel" and productivity was examined in terms of project completion rates and revenue. Their survey findings indicate that more multitasking is associated with greater revenue generation and project output to a point, after which there are diminishing marginal returns, then negative returns to increased multitasking. In other words, at lower levels of workload people who multitasked were more productive, but after a certain level, multitaskers' productivity declined.

While there is evidence of the curvilinear relation between productivity and multitasking based on survey data and on a definition of multitasking dependent upon projects and simultaneity, to our knowledge, there are no laboratory studies to date testing the inverted-U relationship with prescribed problem-solving tasks. To examine whether different levels of multitasking produce different effects on performance, we designed an experiment for individual users where tasks and times are controlled.

3. Research design

To study the effects of different multitasking strategies on performance, we developed a custom program consisting of six problem-solving tasks. Each task had an objectively correct answer. Our goal was to provide an experimental environment with multiple tasks of different duration and cognitive requirements where tasks interleaving is more likely to occur (Payne et al., 2007). Participants were given a fixed amount of time to work on six tasks of different duration that contributed independently to overall performance.

Using this application, we conducted a laboratory experiment with two conditions, using a between-groups design. In the control condition, the tasks were presented sequentially, and in the experimental condition, the tasks were presented all at once organized in six tabs. Participants in the control group carried out the tasks in sequence with *no multitasking*, while those in the experimental condition were able to *discretionarily multitask* by clicking on the corresponding tab at any moment. Fig. 1 shows the main screen of the experimental condition. In both conditions, the application managed total time on each task to control for the potential effects of time on performance. Maximum allotted time for the entire exercise was determined from pilot studies and was set at a limit intentionally shorter than the average amount of time users would need to complete these tasks. With such time restrictions, we sought to create an environment free of gaps (or idleness) due to early termination of tasks. In addition, to avoid the effects of task sequence, the tasks were presented in a random order in the non-multitasking condition, and the tabs were randomized in the discretionary multitasking condition. The main task, however, was always displayed first.

The main task was a Sudoku problem of medium difficulty. The goal of a Sudoku exercise is to fill in all the boxes in a 9×9 grid, so that each column, row, and 3×3 box have the numbers 1 through 9 without repetitions. The maximum

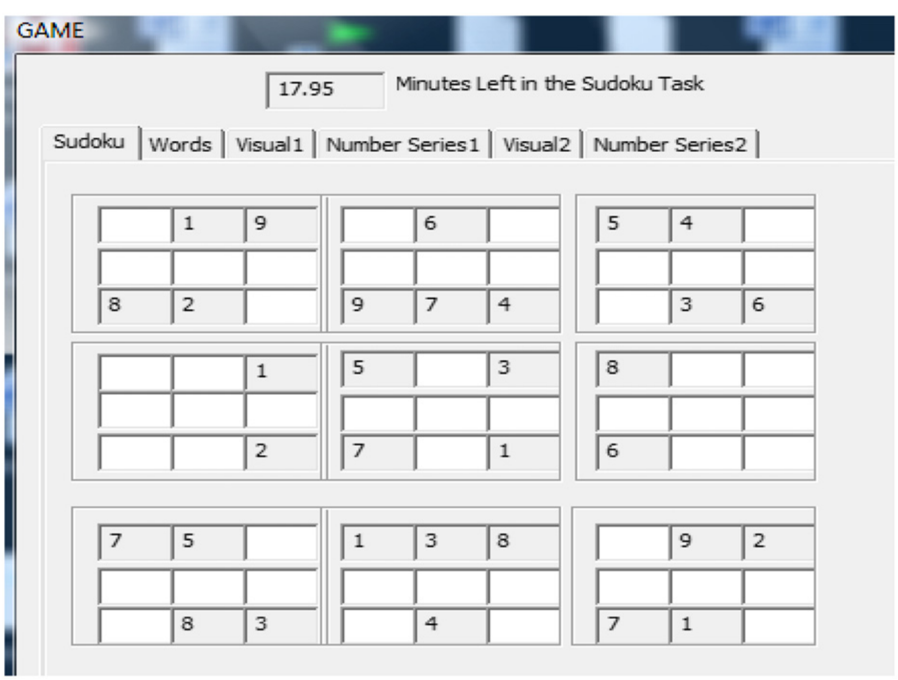

Fig. 1. Screen shot of experimental condition.

allotted time for this task was 18 min; a second Sudoku problem was available for those who finished earlier. This type of problem was chosen as the main task because it requires some mental concentration and a period of time for accurate completion and verification. Moreover, if this task is abandoned before its completion, participants need to remember their previous thought process in order to resume working.

In addition to the Sudoku problem, there were five additional tasks of shorter duration, one textual, two visual and two numeric series challenges. The duration of the smaller tasks was 1 min and a half for the textual task and about 1 min for the visual and numeric series. The textual task consisted of a word production challenge where users had to create 20 different words by unscrambling the letters of the word provided. The visual tasks consisted of "Odd One Out" problems, where users had to find the shape that did not fit the pattern presented in a series of four other shapes. There were two visual task sets with ten problems each. The numeric tasks consisted of number series problems where subjects had to fill in the missing number in the series. As in the case of the visual task sets, there were two numeric series task sets, each with 10 problems.

3.1. Experimental procedures

We recruited subjects from the undergraduate student population of a large urban college in the Northeast of the U.S. Participants were randomly assigned to each condition and received a ten-dollar incentive for their participation. The experimental sessions took place in a specially fitted lab with individual personal computers that were only running the custom-developed application. Participants were given a handout with the details of the exercise and instructions tailored to their condition. Those in the control non-multitasking condition were told that the system would present the tasks in succession, and those in the discretionary multitasking condition were told that they could complete the tasks in any order by clicking on the tabs. They were alerted to the time limits and the task requirements.

Before working on the tasks, the application presented each participant with a pre-test questionnaire to collect demographic information (gender, age, computer skills, and previous Sudoku experience). Then, the tasks were displayed depending upon the condition (one after the other for the non-multitasking group, or all at once in different tabs in the discretionary multitasking condition). Upon completion of the exercise, the application provided a post-test questionnaire to capture the participants' perceptions about the exercise. Two of the perceptual variables measured are perceived Sudoku ease and multitasking perception. Perceived Sudoku ease was measured in a 1–5 scale where 1 is easy and 5 is difficult. The variable was reversed to indicate ease. The scale used to measure multitasking perception was adapted from Bluedorn et al. (1992). The items were measured on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree):

- In this exercise, I switched between the assigned computer-based tasks.
- In this exercise, I tried to complete the assigned computer-based tasks at the same time.
- In this exercise, I worked on one computer-based task at a time (Reversed).
- In this exercise, I was carrying out several computer-based tasks at the same time.

3.2. Measures of variables

To measure multitasking activity we used *number of switches*. This variable was determined by counting the number of times a user clicked on the tabs during the session. Since each tab contained an independent task set, the number of switches is representative of task switching. Other studies (Czerwinski et al., 2004; Payne et al., 2007; Zhang et al., 2005) have operationalized multitasking with this type of measure. According to Benbunan-Fich et al. (2011), switches-related measures are rich measures of multitasking because they combine elements of task, user and technology. However, instead of the percentage-based metric for switches proposed in that study, we use an additive measure. In our case, the number of tasks and times on task are controlled by our custom-developed environment. Thus, a simpler count of switches offers a suitable multitasking metric that is comparable across subjects.

Performance is conceptualized in two dimensions: performance effectiveness, measured with *accuracy*, and performance efficiency, measured with *productivity*. Both of these performance measures were automatically calculated by the custom-developed application log data generated for each user.

The accuracy variable was determined with the user's scores in each task. Since each task had a correct answer, the scores were calculated with the number of correct answers as a percentage of the total answers required. For example, in the Sudoku task there were 49 empty spaces that needed to be filled out with the appropriate numbers. The score was the number of correct numbers entered divided by the total number of empty squares. For the word task, there were 20 acceptable words that could be generated from unscrambling the letters. The percent correct is the number of correct responses out of 20. The same method was applied to calculate the visual and number series tasks' scores. The total score was computed by averaging the accuracy scores of all six tasks.

The productivity measure was determined based on the amount of work completed per task, regardless of whether the answers were correct. For instance, the percentage of the 49 empty Sudoku spaces that were filled out, or the number of words out of 20 that were typed into the word solution boxes. The total productivity score was computed by combining the percentage completion scores across the six tasks.

4. Data analysis and results

We recruited 205 subjects (90 female and 115 male) and assigned them randomly to each condition: 102 subjects solved the tasks sequentially in the Non-Multitasking (NMT) condition, and 103 were allowed to switch tasks at will in the Discretionary Multitasking (DMT) condition. To ensure that randomization worked and to rule out alternative explanations, we first checked for pre-existing differences among participants assigned to both conditions. None of the continuous pre-test questionnaire variables (age, computer skills or Sudoku experience) showed a systematic variation. Computer Skills was measured with a 5-point scale (from 1=poor to 5=excellent). Sudoku Experience was measured with a 0–5 scale similar to computer skills, with a 0 for those who had never played Sudoku before. From the post-test variables, no differences across conditions are found in the perception of Sudoku ease. In contrast, there are significant differences in the perceptions of multitasking as explained in the next section. Table 1 presents these results.

Gender was coded with a dichotomous variable (men=0; women=1). A separate chi-square analysis for gender shows that male and female participants were equally distributed in each condition ($\chi^2=0.65$; $p=0.62$ ns).

4.1. Manipulation check

The multitasking perception scale was used to perform a manipulation check. A confirmatory factor analysis of these items showed a single factor with items loadings of 0.70 or higher, except for one item, which was dropped to create the index. The reduced scale had a high level of reliability (Cronbach's Alpha of 0.796). The multitasking perception means in the DMT condition are significantly higher than those in the NMT condition ($\text{Mean}_{\text{NMT}}=2.48$ vs. $\text{Mean}_{\text{DMT}}=2.84$; $F(1, 203)=8.69$, $p=0.004$). This test provides evidence of the integrity of the two conditions as implemented in our custom-developed environment. The results indicate that users in the experimental condition perceived the environment as involving more multitasking than those in the control condition.

4.2. Analysis of discretionary multitasking sub-sample

Descriptive statistics of the variables of interest for participants in the discretionary multitasking condition are presented in Table 2.

The number of switches for participants in the DMT condition ranges from 5 to 29. Five switches indicate that despite the flexibility afforded by this condition, some participants performed their tasks sequentially and chose not to multitask. Since there were six tasks/tabs, the minimum number of tab changes in this condition is five, which indicates that the participant never returned to a previously used tab. Participants were able to switch tasks at any point by clicking on the corresponding tab. However, if the time limit on a tab was reached that tab became disabled and the user was unable to return to that task. All tabs had to be attempted at least once as the multitasking environment did not terminate until every task/tab's time expired. Clicking on the current tab did not increment the number of switches.

To test our hypothesis, we used number of switches as the independent variable. We ran two quadratic models on the DMT sub-sample using either accuracy or productivity as the dependent variable. In both models, the number of switches was the main independent variable with two terms (linear and quadratic). The models also included control variables such as age, gender, computer skills and Sudoku experience. The general formula of the quadratic model is:

$$Y=\beta_0+\beta_1 X+\beta_2 X^2+\beta_3 Z_1+\beta_4 Z_2+\beta_5 Z_3+\beta_6 Z_4+\varepsilon$$

Table 2
Means of strategy and performance variables in experimental condition only.

[$N=103$]	Mean	Std. dev.	Minimum	Maximum
Number of switches	9.43	5.09	5	29
Accuracy	41.11	13.14	12.94	69.66
Productivity	61.14	12.90	31.89	93.33

Table 1
Descriptive statistics and comparison of conditions.

	Mean	S.D.	Min	Max	Comparison results
Pre-test variables					
Age[a]	21.56	3.26	18	35	$\text{Mean}_{\text{NMT}}=21.69$; $\text{Mean}_{\text{DMT}}=21.44$; $F(1, 202)=0.31$ (ns)
Computer skills	3.80	0.85	1	5	$\text{Mean}_{\text{NMT}}=3.80$; $\text{Mean}_{\text{DMT}}=3.80$; $F(1, 203)=0.00$ (ns)
Sudoku experience	1.65	1.40	0	5	$\text{Mean}_{\text{NMT}}=1.67$; $\text{Mean}_{\text{DMT}}=1.63$; $F(1, 203)=0.03$ (ns)
Post-test variables					
Perc'd Sudoku ease	2.85	1.49	1	5	$\text{Mean}_{\text{NMT}}=2.85$; $\text{Mean}_{\text{DMT}}=2.85$; $F(1, 203)=0.00$ (ns)
Multitasking perception	2.66	0.90	1	5	$\text{Mean}_{\text{NMT}}=2.48$; $\text{Mean}_{\text{DMT}}=2.84$; $F(1, 203)=8.69$***

[a]One subject typed an invalid age and was omitted from this analysis.
***Significance level: $p<0.001$.

where β_0 is the intercept, X is the number of switches, β_1 and β_2 are the coefficients of the linear and quadratic terms, β_3, β_4, β_5 and β_6 are the coefficients of the control variables (Z_1, Z_2, Z_3 and Z_4), respectively, and ε is the error term.

Table 3 presents the results of the models, the R^2, the estimated coefficients and the significance of each one.

Both models are significant at $p<0.001$. In the accuracy model, neither the quadratic term nor the linear term of total switches are significant. In contrast, in the productivity model, both the quadratic term and the linear term are significant (coefficients -0.08, $p<0.05$ and 2.38, $p<0.05$, respectively). Furthermore, the sign of the coefficient of the quadratic term in Model 2 is negative indicating an inverted-U relation. With respect to the control variables, the higher the level of Sudoku experience the better the performance, both in terms of accuracy and productivity. The effects of gender are only significant for accuracy. In the DMT sub-sample, 56 men have significantly higher accuracy than the 47 women (means 43.71 vs. 38.01). A separate analysis between gender and accuracy shows that this difference between men and women is significant (t-value$=$2.23; $p=0.028$). It should be noted that the results of the quadratic model without control variables are identical to those reported here but the R^2 is smaller.

Fig. 2 shows a simplified graph of the quadratic model for productivity. For visualization and ease of interpretation purposes, the plotted curve corresponds to the fitted values obtained from the equation: $Y=49.73–0.08X^2+2.38X$. The control variables are omitted for simplicity. The inverted-U pattern is clearly observed in this graph.

Since the coefficients of the single and square terms of number of switches were not significant in the quadratic model with accuracy as the dependent variable, we ran a linear model. The results are shown in Table 4. In this case, the linear model for accuracy is significant (Model $F(6, 96)=14.8$; $p<0.001$). The coefficient of the linear term of total switches is negative and significant indicating a decreasing line. In this model, both Sudoku experience and gender are significant explanatory variables. Consistent with the explanations above, higher levels of Sudoku experience are associated with better accuracy and men in this sample are more proficient than women. The variance explained by this model is 43%.

Taken together the results of these analyses indicate that in terms of accuracy, increased multitasking consistently produces a decreasing effect. In contrast, the effects for productivity do follow an inverted-U curve as predicted by our hypothesis.

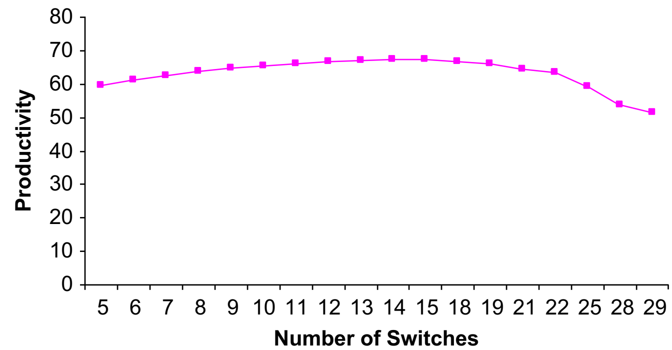

Fig. 2. Graph of quadratic model for productivity (Model 2).

Table 3
Quadratic models [$N=103$].

	Model 1: Accuracy	*Model 2:* Productivity
Model $F(6, 96)$	13.15***	2.32*
R^2 (%)	45.1	12.6
Parameters	*Estimate (t,p)*	*Estimate (t,p)*
Intercept	25.57 (2.29*)	49.73 (3.59***)
Total switches	1.00 (1.23 ns)	2.38 (2.39*)
Total switches squared	−0.05 (−1.79 ns)	−0.08 (−2.50*)
Sudoku experience	5.13 (7.04***)	1.98 (2.20*)
Gender	−6.41 (−3.15**)	−2.54(−1.01 ns)
Age	−0.00 (−0.00 ns)	−0.20 (−.47 ns)
Computer skills	1.70 (1.34 ns)	0.27 (0.17 ns)

*Significance level: $p<0.05$.
**Significance level: $p<0.01$.
***Significance level: $p<0.001$.

Table 4
Linear model [$N=103$].

	Accuracy
Model $F(6, 96)$	14.80***
R^2 (%)	43.3
Parameters	*Estimate (t,p)*
Intercept	33.36 (3.20**)
Total switches	−0.41 (−2.01*)
Sudoku experience	5.25 (7.16***)
Gender	−5.91(−2.90 **)
Age	−0.00 (−0.00 ns)
Computer skills	1.51 (1.18 ns)

*Significance level: $p<0.05$.
**Significance level: $p<0.01$.
***Significance level: $p<0.001$.

4.3. Comparative analysis of performance

An initial comparison of performance metrics (accuracy and productivity) for participants in both conditions was performed. The means of accuracy and productivity scores in the non-multitasking condition though slightly higher than those in the discretionary condition, are not significantly different (see Table 5).

For a more detailed analysis of performance within the experimental sub-sample, we divided number of switches into three different categories: Low (number of switches below 10), Medium (number of switches between 10 and 15) and High (15 or more switches). The cutoff points (10 and 15) were

Table 5
Comparison of performance control vs. experimental conditions.

	Mean	S.D.	Min	Max	Comparison results
Accuracy	41.50	12.87	10.75	75	$Mean_{NMT}=41.90$; $Mean_{DMT}=41.11$; $F(1, 203)=0.19$ (ns)
Productivity	60.99	13.22	26.99	93.33	$Mean_{NMT}=60.83$; $Mean_{DMT}=61.14$; $F(1, 203)=0.03$ (ns)

determined from the mean and standard deviation of this variable for the DMT sub-sample. These three categories represent the extent to which participants engaged in multitasking behavior, whether they were light multitaskers, medium multitaskers, or heavy multitaskers. According to the predictions of the inverted-U hypotheses, those in the medium multitasking category should have the highest productivity when compared to those with low or high levels of multitasking. The results of *t*-tests confirm this prediction. The most productive are those with medium levels of multitasking (71.20), while those at the extremes with low or high multitasking are least productive. Their average productivity levels are 59.17 and 56.66, respectively. Consistent with the findings of an inverse relation between multitasking, the results of additional *t*-tests confirm that high multitaskers had the lowest accuracy (30.89) compared with the low multitaskers (41.62) and medium multitaskers (47.05). These results are presented in Table 6.

Given the performance differences within the DMT condition, depending on the multitasking category (low, medium or high) it is possible that the use of average performance is masking the existence of significant differences. To examine this possibility, we compared each segment of the discretionary multitasking sample (low, medium and high) with the non-multitasking condition in each of the two performance metrics. From these comparisons, only two were significant. First, *high multitaskers* had significantly lower accuracy scores than participants in the NMT (control) condition (30.89 vs. 41.90; t-value$=3.08$, $p<0.01$). Second, *medium multitaskers* obtained significantly better productivity scores than non-multitaskers (71.20 vs. 60.83; t-value$=-3.16$, $p<0.01$). These results are also presented in Table 6.

Consequently, two noteworthy findings emerge from the comparison of performance. First, high multitaskers performed significantly worse than everybody else (medium multitaskers, low multitaskers, and non-multitaskers) in terms of accuracy. Second, medium multitaskers obtained the highest productivity scores compared to the other multitaskers and non-multitaskers.

4.4. Analysis of switching patterns

Different multitasking strategies, applied by participants in the DMT condition, are captured by the number of switches. Patterns of switches are illustrated with timeline graphs. In these graphs, the line segments indicate the approximate amount of time (in seconds) that each task remained active based on its start and end time.

Table 6
Comparison of means.

Accuracy		Mean	*t*-Values
DMT-Low	DMT-Med	41.62 vs. 47.05	−1.79 ns
DMT-Low	DMT-High	41.62 vs. 30.89	3.06**
DMT-Med	DMT-High	47.05 vs. 30.89	3.45**
NMT	DMT-Low	41.90 vs. 41.62	0.15 ns
NMT	DMT-Med	41.90 vs. 47.05	−1.66 ns
NMT	DMT-High	41.90 vs. 30.89	3.08**
Productivity		**Means**	***t*-Values**
DMT-Low	DMT-Med	59.17 vs. 71.20	−4.24***
DMT-Low	DMT-High	59.17 vs. 56.66	0.57 ns
DMT-Med	DMT-High	71.20 vs. 56.66	2.99**
NMT	DMT-Low	60.83 vs. 59.17	0.89 ns
NMT	DMT-Med	60.83 vs. 71.20	−3.16**
NMT	DMT-High	60.83 vs. 56.66	1.08 ns

*Significance level: $p<0.05$.
**Significance level: $p<0.01$.
***Significance level: $p<0.001$.

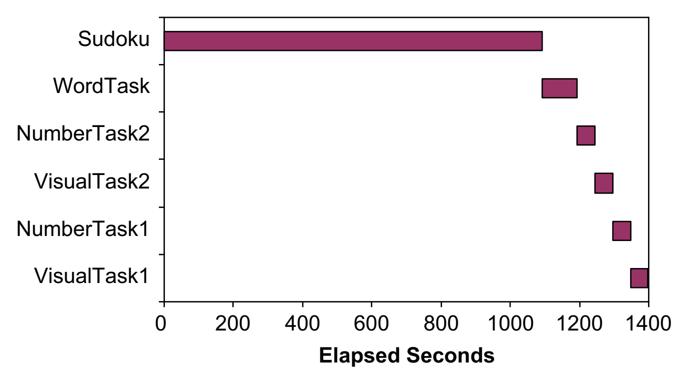

Fig. 3. Participant applying a non-multitasking strategy.

Fig. 3 shows the timeline graph of a subject who chose not to multitask and allocated his/her time in a completely sequential fashion, with only five switches. We note that sixteen subjects (assigned to the experimental condition) applied a strictly sequential working strategy. As illustrated in the graph, the user first completed the Sudoku task (SUD), followed by the Word task (WRD), then the Number Series task (NUMB2), the Visual task (VIS2), the other Number Series task (NUMB1), and finally the other Visual task (VIS1).

Fig. 4 shows the timeline of a subject who had 10 switches and multitasked more than the non-multitasking user. In this case, after about 2 min playing Sudoku,

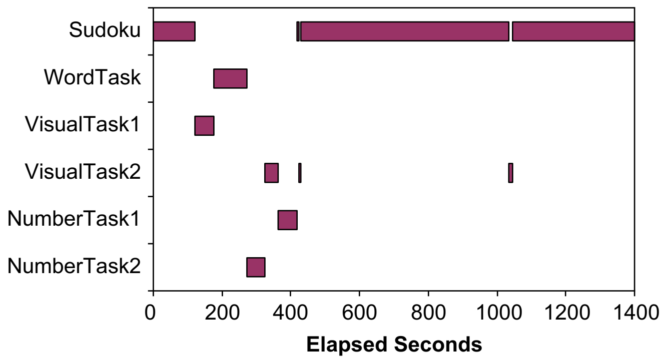

Fig. 4. Participant applying an interleaving strategy.

Table 7
Correlations with number of switches in experimental condition [$N=103$].

	Accuracy	Productivity	Prior Sudoku experience	Multitasking perception	Perceived Sudoku ease
Total switches	−0.22*	−0.03	−0.08	0.31**	0.21*

*Significance level: $p<0.05$.
**Significance level: $p<0.01$.

the subject switched to the first visual task (VIS1), completed it and went on to start and finish the word task (WRD) and the number series task (NUMB2). The subject then began the other visual task (VIS2), but before completing it switched to the first number series task (NUMB1) which s/he completed. This participant then went back and forth between playing Sudoku and switching to the Visual task (VIS2). When time ran out in the visual task, s/he spent the last minutes working on the Sudoku problem.

Measuring actual multitasking with number of switches shows the extent to which participants engaged in this behavior. However, two participants with the same number of switches may have different switching patterns throughout their experimental session. To investigate this possibility, we analyzed the end times of the secondary tasks and compared them to the end time of Sudoku. A total of 64 people (62%), compared to only 39 (38%) finished all the secondary tasks before completing the Sudoku problem and this percentage is significant (Chi-square$=6.07$; $p=0.014$). This finding is consistent with Yeung (2010) who reports that in situations where people choose to multitask they have a tendency to finish the easier or shorter tasks first.

The total number of switches was further analyzed with respect to the other variables collected in this experiment to gain insights into why some people switch more than others. Pairwise correlations are presented in Table 7.

These correlations confirm the findings reported earlier. Accuracy is negatively related to total switches, while productivity does not have a significant correlation due to the curvilinear association between these variables. It is noteworthy that the number of switches is significantly correlated with multitasking perceptions, thus showing that subjective perceptions of multitasking activity are consistent with the objective measure of switches. In addition, the significant correlation between perceived Sudoku ease and total switches indicates that those who perceived Sudoku to be easier tended to switch more than those who found this task to be difficult.

5. Discussion

In this paper, we propose a theoretical model based on memory-for-goals theory and the Yerkes–Dodson law to explain the relation between computer-based multitasking behavior and performance. Based on cognitive theories of goal-directed attention and different multitasking strategies, the model predicts an inverted-U relationship between multitasking and performance. This model provides the foundation to systematically compare the effects of alternative multitasking strategies. To test the model, we conducted a controlled experiment with a custom-developed application that featured several problem-solving tasks of different durations. Participants were randomly assigned to one of two conditions: non-multitasking condition (control group) or discretionary multitasking condition. Those in the control condition were presented with the tasks one at a time in sequence. Subjects in the discretionary multitasking condition were allowed to switch tasks at will. Consistent with our treatment, the discretionary multitasking condition reported significantly higher perceptions of multitasking than participants in the control condition, which provides evidence of the integrity of our experimental conditions.

To test the predictions of the Yerkes–Dodson law, we used the number of switches as an indicator of multitasking. Given the nature of our experimental setting, the number of switches is an appropriate metric to determine the extent to which each participant engaged in multitasking behavior. This metric is also consistent with the premise of memory-for-goals theory. Switches are indicative of self-interruptions and goal shifts. Regardless of the amount of effort put into the new task, by switching the subject is diverting attention from the previous task goal and focusing on a new goal. Thus, the higher the amount of switches, the higher the number of times a user shifts goals.

In our analysis of the discretionary multitasking subsample, we found that a quadratic model explains performance efficiency (productivity). This finding is consistent with our theoretical development which proposes that increased multitasking levels are beneficial to a point after which there are negative effects. The inverted-U curve can also be explained in terms of the costs associated with multitasking. At lower levels of multitasking, the cognitive switching costs associated with swapping problem state information in memory are offset by the efficiency gains achieved from increased arousal. This enhanced state of alertness mobilizes additional cognitive resources that result in better productivity. However, at high levels of multitasking, any potential performance improvement is lost to between-task interference.

In contrast to the inverted-U curve found for performance efficiency, a linear model explains performance effectiveness (accuracy). It is likely that the nature of the tasks chosen for this experiment along with the strict time controls account for the linear degradation of performance for those that discretionarily engaged in multitasking behavior. Based on the predictions of the theoretical development, it appears that accuracy gains are not experienced with increased levels of multitasking at any point. On the contrary, accuracy steadily decreases as multitasking behavior increases. This finding has to be carefully examined in other contexts with different tasks or more lenient time controls.

On average, performance metrics (accuracy and productivity) did not show any statistically significant differences between participants in the control and the experimental condition. To investigate whether performance averages were masking important differences, the discretionary multitasking sub-sample was segmented in three categories (low, medium and high), according to the number of switches. Comparisons between each sub-group and the control condition showed two significant differences. On the one hand, high multitaskers had the lowest accuracy scores than other multitaskers and non-multitaskers. On the other hand, medium multitaskers had the best productivity scores than anybody else (including other multitaskers and non-multitaskers).

The most intriguing question is why some participants switch more than others. To find possible explanations, we ran pair-wise correlations with number of switches. The correlation with prior Sudoku experience is not significant ($\rho = -0.08$, $p = 0.40$ ns). Although prior experience with Sudoku has a significant influence on performance effectiveness for subjects who worked in the discretionary multitasking condition (i.e. the coefficient for Sudoku experience is significant in all the models), there is no significant pair-wise correlation between prior level of experience and number of switches. In contrast, there is a significant correlation between number of switches and perceived Sudoku ease. This finding suggests that those who found the main task easy did switch more than those who found it difficult. In the context of our theoretical development, it appears that task difficulty can act as a deterrent to switching, while on the other hand the easier the task, the higher the likelihood of switching.

Overall, the results of this study suggest that the relation between multitasking and performance depends upon the metric used for assessing performance. If performance is measured with accuracy of results (i.e. as performance effectiveness), the relation approximates a downward slopping curve, in which increased levels of multitasking lead to a significant loss in accuracy. In contrast, when performance is measured with productivity (i.e. as performance efficiency), different multitasking levels are associated with an inverted-U curve. Those with medium levels of multitasking activity had better productivity than those in the lower or higher levels.

5.1. Limitations

Several limitations should be acknowledged in the interpretation of our results. The first limitation stems from our choice of research method. By conducting an experiment we attempted to maximize precision at the expense of realism and generalizability (Jung et al., 2010). As Dennis and Valacich (2001, p. 5) indicate, "no one method is better or worse than any other," some methods are better at some aspects and worse at others. The use of a laboratory experiment allowed us to precisely control the variables of interest (tasks and times) and test our hypothesis in the absence of extraneous influences and confounding factors.

Notwithstanding the freedom to switch tasks in the discretionary multitasking condition, the empirical findings we report may have been influenced by the parameters imposed by the custom-developed application. The nature of the tasks (problem solving under time constraints) could have affected the extent to which participants exhibited multitasking behavior. In particular, the knowledge that total time on a tab was limited may have affected how participants chose to allocate their time.

It should also be noted that participants for our study were drawn from the student population at a major urban college in the Northeast of the U.S. Furthermore, as is typical in most laboratory experiments, participants did not have a real stake in the outcome. Although they were compensated for their participation, the payment was not tied to their performance. Nevertheless, the quantitative data as well as results of informal interviews suggest that participants took the tasks seriously, particularly because they found the multitasking environment engaging, challenging and fun.

Overall, the limitations stemming from the choice of research methods (laboratory experiment), the design of the multitasking environment (problem solving tasks with time controls) and the nature of the subjects who participated in this research (college students), suggest caution when generalizing these results to other populations and settings. The inability to make generalizations was traded off for the precision and control afforded by laboratory research. Although our experimental environment may be deemed as artificial, it gave us the opportunity to control task and times, and measure performance with two different metrics (accuracy and productivity) and under different multitasking scenarios. As a result of these methodological choices, our research has provided evidence for the complex relation between multitasking behavior and performance. According to our results, when multitasking increases there is an inverted-U curve for productivity but a linearly decreasing function for accuracy.

5.2. Implications for theory, research and practice

At the theoretical level, this study explains why the Yerkes–Dodson law occurs in multitasking situations using memory-for-goals theory. Our model posits that performance increases at lower levels of multitasking when increased arousal from shifting goals is beneficial to keep the users alert and engaged with their tasks. However, performance decreases after a certain point when constant goal displacement has a negative effect on performance. With this theoretical contribution, we have shown the limits to multitasking from the perspective of memory-for-goals theory. Goal shifting is evidenced in task switching, which is related to arousal. Given the link between arousal and switching, a logical expansion to this framework would consist of theorizing about the drivers for task switches with situational and individual antecedents of multitasking, such as personality or cognitive styles, which would offer a more complete view of this type of behavior.

From a research perspective, our study provides the foundation to examine the relation between multitasking behavior and user performance. Exploring the boundaries of this relation and the extent to which it can be replicated with other tasks and in other settings offers a fertile ground for future research. One of the most intriguing questions is why some participants multitask more than others. Further research should investigate whether individual preferences (such as personality traits or multitasking propensity), or situational conditions (such as task complexity, interruptions or notifications) account for differential levels of multitasking activity. According to our findings, the degree of difficulty of the main task influences multitasking behavior.

This study can be extended in multiple directions. One possibility is to add another condition (mandatory interruptions) and compare the results on performance using a larger sample. Another option is to include additional tasks with different levels of difficulty and different priorities. Finally, another potentially fruitful extension consists of allowing participants to initiate their own tasks (such as checking email, browsing the web, etc.). Although the addition of participant-initiated tasks would make the environment more representative, overall performance would be more difficult to measure.

The practical implications of our results are noteworthy. While medium levels of multitasking tend to increase individual productivity, high levels of multitasking are detrimental to accuracy. Accordingly, information workers should be mindful of the implications of multitasking in their own performance and consider reverting to a strictly sequential approach, instead of increasing their own multitasking levels to a point where all productivity gains are lost and performance suffers. This insight can help design novel features to support and control multitasking behavior in computer-based environments.

6. Conclusion

Multitasking is a contemporary phenomenon resulting from a fast-paced, technologically-driven world. At home, at school or at work, people no longer focus their attention on one task at a time, but tend to juggle multiple tasks simultaneously. Although the upside of multitasking could be the illusion of productivity, the downside is their potential negative effects on performance. To investigate this phenomenon, we propose a theoretical model predicting an inverted-U-relation between multitasking and performance, and test it with a laboratory experiment using a custom-developed application. Our results indicate that some multitasking actually improves productivity, but too much multitasking has a negative effect. When accuracy is used to measure performance, the results are not encouraging. More multitasking has a deleterious effect on performance effectiveness. Metaphorically speaking, juggling multiple tasks is much more difficult while balancing on a high wire, where performance mishaps can have serious consequences.

References

Altmann, E.M., Trafton, J.G., 2002. Memory for goals: an activation-based model. Cognitive Science 26 (1), 39–83.

Aral, S., Brynjolfsson, E., Van Alstyne, M.W., 2006. Information, Technology and Information Worker Productivity: Task Level Evidence. MIT Working Paper. Retrieved 20 January 2010. From: 〈http://ssrn.com/paper=942310〉.

Bailey, B.P., Konstan, J.A., 2006. On the need for attention-aware systems: measuring effects of interruption on task performance, error rate and affective state. Journal of Computers in Human Behavior 22 (4), 658–708.

Bäumler, G., 1994. On the validity of the Yerkes–Dodson law. Studia Psychologica 36, 205–209.

Benbunan-Fich, R., Adler, R.F., Mavlanova, T., 2011. Measuring multitasking behavior with activity-based metrics. ACM Transactions on Computer–Human Interaction 18 (2), 1–22.

Benbunan-Fich, R., Truman, G.E., 2009. Multitasking with laptops during meetings. Communications of the ACM 52 (2), 139–141.

Bluedorn, A.C., Kaufman, C.F., Lane, P.M., 1992. How many things do you like to do at once? An introduction to monochronic and polychronic time. Academy of Management Executive 6 (4), 17–26.

Borst, J.P., Taatgen, N.A., Van Rijn, H., 2010. The problem state: a cognitive bottleneck in multitasking. Journal of Experimental Psychology: Learning, Memory, and Cognition 36 (2), 363–382.

Czerwinski, M., Horvitz, E., Wilhite, S., 2004. A diary study of task switching and interruptions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 175–182.

Dennis, A.R., Valacich, J.S., 2001. Conducting research in information systems. Communication of the Association for Information Systems 7 (5), 1–41.

Gillie, T., Broadbent, D., 1989. What makes interruptions disruptive? A study of length, similarity, and complexity. Psychological Research 50 (4), 243–250.

Gonzalez, V.M., Mark, G., 2004. 'Constant, constant, multi-tasking craziness': managing multiple working spheres. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 113–120.

Hembrooke, H., Gay, G., 2003. The laptop and the lecture: the effects of multitasking in learning environments. Journal of Computing in Higher Education 15 (1), 46–64.

Jung, J.H., Schneider, C., Valacich, A.J., 2010. Enhancing the motivational affordance of information systems: the effects of real-time performance feedback and goal setting in group collaboration environments. Management Science 56 (4), 724–742.

Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, NJ.

König, C.J., Oberacher, L., 2010. Personal and situational determinants of multitasking at work. Journal of Personnel Psychology 9 (2), 99–103.

Madjar, N., Shalley, C.E., 2008. Multiple tasks' and multiple goals' effect on creativity: forced incubation or just a distraction? Journal of Management 34 (4), 786–805.

Mark, G., Gonzalez, V.M., Harris, J., 2005. No task left behind?: Examining the nature of fragmented work. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 321–330.

McCrickard, D.S., Catrambone, R., Chewar, C.M., Stasko, J.T., 2003a. Establishing tradeoffs that leverage attention for utility: empirically evaluating information display in notification systems. International Journal of Human–Computer Studies 58 (5), 547–582.

McCrickard, D.S., Czerwinski, M., Bartram, L., 2003b. Introduction: design and evaluation of notification user interfaces. International Journal of Human–Computer Studies 58 (5), 509–514.

McFarlane, D., 2002. Comparison of four primary methods for coordinating the interruption of people in human–computer interaction. Human–Computer Interaction 17 (1), 63–139.

Meyer, D.E., Kieras, D.E., 1997. A computational theory of executive cognitive processes and multiple-task performance: part I, basic mechanisms. Psychological Review 104 (1), 3–65.

Miller, J.G., 1978. Living Systems. McGraw-Hill, New York.

Monsell, S., 2003. Task switching. Trends in Cognitive Sciences 7 (3), 134–140.

Muse, L.A., Harris, S.G., Field, H.S., 2003. Has the inverted-U theory of stress and job performance had a fair test? Human Performance 16 (4), 349–364.

Navon, D., Gopher, D., 1979. On the economy of the human-processing system. Psychological Review 86 (3), 214–255.

Norman, D.A., Bobrow, D.G., 1975. On data-limited and resource-limited processes. Cognitive Psychology 7 (1), 44–64.

Palladino, L.J., 2007. Find Your Focus Zone: An Effective New Plan to Defeat Distraction and Overload. Free Press, New York.

Payne, S.J., Duggan, G.B., Neth, H., 2007. Discretionary task interleaving: heuristics for time allocation in cognitive foraging. Journal of Experimental Psychology: General 136 (3), 370–388.

Rubinstein, J.S., Meyer, D.E., Evans, J.E., 2001. Executive control of cognitive processes in task switching. Journal of Experimental Psychology: Human Perception and Performance 27 (4), 763–797.

Salvucci, D.D., Taatgen, N.A., 2011. The Multitasking Mind. Oxford University Press, New York.

Salvucci, D.D., Taatgen, N.A., Borst, J.P., 2009. Toward a unified theory of the multitasking continuum: from concurrent performance to task switching, interruption, and resumption. In: Proceedings of the CHI'09 Human Factors in Computing Systems, pp. 1819–1828.

Speier, C., Vessey, I., Valacich, J.S., 2003. The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. Decision Sciences 34 (4), 771–797.

Staal, M.A., 2004. Stress, cognition, and human performance: a literature review and conceptual framework. NASA Technical Memorandum 2004-212824.

Teigen, K.H., 1994. Yerkes–Dodson: a law for all seasons. Theory and Psychology 4 (4), 525–547.

Wasson, C., 2004. Multitasking during virtual meetings. Human Resource Planning 27 (4), 47–61.

Wiener, E.L., Curry, R.E., Faustina, M.L., 1984. Vigilance and task load: in search of the inverted U. Human Factors: The Journal of the Human Factors and Ergonomics Society 26 (2), 215–222.

Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology 18 (5), 459–482.

Yeung, N., 2010. Bottom-up influences on voluntary task switching: the elusive homunculus escapes. Journal of Experimental Psychology: Learning, Memory, and Cognition 36 (2), 348–362.

Zhang, Y., Goonetilleke, R.S., Plocher, T., Liang, S.-F.M., 2005. Time-related behaviour in multitasking situations. International Journal of Human–Computer Studies 62 (4), 425–455.