ABSTRACT

PANEPINTO, MARIE PATRICE. The Effects of Voluntary versus Forced Task Switching on Task Performance. (Under the direction of Douglas J. Gillan, PhD).

Research on task switching has focused on the relatively well known task switching cost, usually defined as an increase in RT on a trial directly following a switch. Two main issues with previous studies suggest that their results may not be applicable to real world scenarios; one, that they typically use short and arbitrary tasks in comparison to real work situations and two, that the vast majority force participants to switch rather than allowing them to do so voluntarily, as is common in the workplace. The current experiment utilized two longer lasting tasks (document proofreading and a Sudoku puzzle) to more closely resemble real world situations and four task switching groups. One group switched voluntarily, one was forced without warning, one was forced with a cue that a switch would be coming, and one served as a no switch control group. Performance, reaction time, and mental workload (NASA-TLX) were measured. Task switch group produced no differential effects on these variables, and no task switching cost was found. Though the hypothesis were not supported, these results lend support to the notion that previous lab studies may not adequately resemble real world scenarios and that micromanaging small tasks, and switching between comparatively longer lasting tasks may not be the same thing. More research on this area may help to produce a better understanding of why people task switch and what they experience cognitively when they do so.
The Effects of Voluntary versus Forced Task Switching on Task Performance

by
Marie Patrice Panepinto

A thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Master of Science

Psychology

Raleigh, North Carolina

October 28, 2009

APPROVED BY:

Christopher B. Mayhorn, PhD
Anne Collins McLaughlin, PhD

Douglas J. Gillan, PhD
Committee Chair
Dedication

For my parents
Biography

Marie Panepinto was born and raised in northern New Jersey. She attended college at SUNY, Fredonia on partial academic scholarship and was a member of the honors program. She graduated summa cum laude in May 2006 with a BA in psychology and a minor in theatre. In August 2006, she entered the Human Factors & Ergonomics graduate program in the psychology department at North Carolina State University. She has served as a teaching assistant, teaching courses in statistics and ergonomics. Her professional human factors experience includes working as a user experience intern at HumanCentric.
# Table of Contents

List of Tables ............................................................................................................................ v
List of Figures .......................................................................................................................... vi
The Effects of Voluntary versus Forced Task Switching on Task Performance ...................... 1
Introduction ........................................................................................................................... 1
Method .................................................................................................................................. 9
Participants ........................................................................................................................ 9
Apparatus/Materials .......................................................................................................... 9
  Computer ....................................................................................................................... 9
  Document .................................................................................................................... 10
  Sudoku ....................................................................................................................... 10
  NASA-TLX ............................................................................................................... 10
  Cognitive Tests .......................................................................................................... 11
  Demographic Survey ................................................................................................. 11
Procedure .......................................................................................................................... 11
  General ....................................................................................................................... 11
  Voluntary Switch Group ............................................................................................ 12
  Forced Switch Group .................................................................................................. 12
  Cued Switch Group ..................................................................................................... 12
  No Switch Control Group .......................................................................................... 13
Results ................................................................................................................................. 13
  Participant and Switch Data ......................................................................................... 13
  Task Performance ...................................................................................................... 14
  Time Data .................................................................................................................. 14
  TLX Data ..................................................................................................................... 15
Discussion .......................................................................................................................... 16
References ............................................................................................................................ 21
Appendix ................................................................................................................................. 25
List of Tables

Table 1  Mean(sd) Cognitive Test Score by Task Switching Group…………………………26
Table 2  Mean(sd) Performance Data by Task and Group……………………………………...27
Table 3  Mean(sd) Reaction Time (in seconds) Following Task Switch by Task Switch
        Group …………………………………………………………………….……....28
Table 4  Mean(sd) Reaction Time (in seconds) Following Task Switch by Task Order
        Condition ………………………………………………………………………...29
Table 5  Mean(sd) Raw NASA-TLX Ratings of Workload Dimensions and Total
        Weighted Scores by Task Switch Group………………………………………...30
List of Figures

Figure 1  Sample Yoking of Task Switching Groups…………………………………………31
Introduction

Task switching, the decision to cease performing one task before its completion in favor of performing another, has become a regular occurrence in many people’s daily lives (Czerwinski, Horvitz, & Wilhite, 2004). Despite its popularity, most recent research suggests that people should not switch between tasks. The idea of a “task switching cost” is well documented (Cellier & Eyerolle, 1992; Gilbert & Shallice 2002; Rogers & Monsell, 1995; Schneider & Logan, 2007; Wylie & Allport, 2000). A task switching cost is typically defined in terms of a slower reaction time during a trial in which a different task was completed in the previous trial, compared to a trial where the task remained the same. Wylie & Allport (2000) suggested that this task switching cost was because of persisting activation of inappropriate task rules; the rules for the first task remain activated even after the switch to another has occurred and it takes time to correct the error and adjust to the new rules before taking any action. When the rules remain consistent (i.e. the task does not switch), this error does not occur.

Task switching is inherent in many of the tasks that people perform in their daily lives. For example books are not read from cover to cover without stopping; they can spend weeks on night tables with people reading one chapter at a time. College courses are broken up into lectures occurring once or twice a week for a semester. In these cases, people have few problems picking up where the task was last left off; one book is seldom confused with another, and students do not need to hear the previous lecture again prior to understanding the current one (although a quick review to reinstate the context of the new material never hurts). People take study breaks rather than study for hours without stopping, and this kind
of behavior is actually recommended by research for both motor skills (Shebilske, Goetl, Corrington, & Day, 1999) and learning content (Bahrick & Phelps, 1987). If task switching research says that people should not engage in this sort of behavior because it is detrimental to their performance, then why is it so common?

Closely examining previous research on issues related to task switching reveals a variety of problems with the methodology. For example, these studies typically bear little resemblance to real life task performance. Specifically, the nature of the tasks that participants perform and the way in which they are performed differ markedly from the performance of most tasks. Therefore, one should be cautious before generalizing results from that research to a broad argument against task switching.

The first issue with many task switching studies is that the tasks that participants perform are relatively arbitrary. They have included color identification (Hartley & Little, 1999; Leacock, 1997), shape identification (Leacock, 1997), odd/even or high/low number classification (Lien & Proctor, 2003; Monsell, Sumner, & Waters, 2003; Rogers & Monsell, 1995; Schneider & Logan, 2007), consonant/vowel letter classification (Lien & Proctor; Rogers & Monsell; Schneider & Logan), identifying changes in light intensity (Tombu & Jolicoeur, 2004), and the Stroop task (Gilbert & Shallice, 2002; Wylie & Allport, 2000). These tasks are minute in comparison to everyday work tasks. For example, in a diary study of task switching behavior, Czerwinski, et al. (2004) found that participants tended to define tasks on higher level, e.g. create charts for meeting, matlab coding, respond to e-mail. These comparatively larger tasks obviously consist of smaller elements, and the experimental tasks mentioned above bear more resemblance to those subtasks than actual tasks.
A similar argument can be made for the amount of time that it typically takes to complete these simple tasks—less than 2 seconds to perform one trial of many of the tasks described above. Rather than performing a task that takes a significantly longer amount of time to complete, participants perform series of repetitions of short, arbitrary tasks. Switching can occur as often as every trial (Wylie & Allport 2000), or as relatively infrequently as every 8 trials (Monsell, Sumner & Waters, 2003), creating a maximum time of approximately 16 seconds between switches, which is also not typical of real world performance. Czerwinski, et al. (2004) found in their diary study that the average task length during a workday was 53 minutes with a standard deviation just above 90 minutes. Once again, the experimental methodology does not necessarily match how most users define their own daily tasks. The switching behavior required in these experiments is more characteristic of the behavior seen in multitasking (actively trying to complete at least two ongoing tasks at the same time), than of a person switching between tasks.

Because of the short time spent on these tasks, this may pose a high working memory load. For example, Baddeley, Chincotta, and Adlam (2001) asked subjects to perform a verbal task, reciting aloud the days of the week and months of the year alternatively. To prevent learning, a random starting point was chosen. Subjects would be presented with “Saturday, March” which required the continuation “Sunday, April, Monday, May.” These verbal trials were alternated with a list of arithmetic problems switching between addition and subtraction. It is not surprising that performance on these tasks was poor given the number of rules participants were required to retain simultaneously. The implications of this study do not appear to argue against task switching so much as they once again demonstrate
limited working memory capacity, even in the face of relatively simple tasks. Also
supporting the role of working memory in task switching, irrelevant information presented
between two tasks has been found to aid in disengagement from the first task (Leacock,
1997). That is to say, the need to keep so much information in working memory caused an
overload, and a short interval allowed the first task information to be purged from working
memory.

Another pressing concern about the task switching research is that participants are
typically forced to switch from one task to another. In contrast, Czerwinski, et al. (2004)
found that 40% of switches in the workplace were reported as being self-initiated rather than
forced. The second most common reason for task switching was task completion at 19%. In
task switching experiments where the switch is forced, the second task may be more of an
interruption than another task. Though many reasons provided for task switching would be
dependent on the situation, 23% of reported task switches can be considered in this category.
Interruptions have been shown to negatively affect performance (Bainbridge, 1984; Hodgetts,
& Jones, 2006; Monk, Trafton, & Boehm-Davis, 2008; Zijlstra, Roe, Leonora, & Krediet,
1999) by leading to a resumption lag in the time taken to return to the interrupted task.

From a methodological standpoint, the task interruption paradigm does not differ
greatly from the task switching paradigm, as has been recognized earlier (Monsell, 2003;
Allport & Wylie, 2000). In this case, is not surprising that a similar negative effect is found.
In addition to simply increasing the time on task, research suggests that interruptions also
negatively affect a person’s state of mind, including a decrease in positive feelings about the
task and increased rating of mental effort (Zijlstra, et al., 1999).
The major difference between the interruption scenario and the task switching scenario is semantic. In the task interruption paradigm, researchers distinguish between a primary and secondary task. The primary task is interrupted by the secondary task. The resumption lag refers to the time it takes for the participant to resume the primary task after completing the secondary task. This secondary task is largely ignored when it comes to performance measurement. In contrast, task switching research makes no such distinction. Participants perform two separate tasks without any indication that one is more important than the other. Accordingly, the task switching cost measures the time that it takes for participants to resume either task.

In forced task switch situations people may try to exert some control over the situation, if they can. For example Zijlstra et al. (1999) found that when participants perform a document editing task and are interrupted with a phone call, they would often wait several rings before answering to deal with the interruption, perhaps trying to schedule the interruption at a more convenient time in their primary task progress. Likewise, even in random task switching conditions, older adults became less engaged with a task in anticipation of an impending task switch (Kray, 2006) suggesting that the participants were preparing to switch tasks. These findings imply that people employ strategies to deal with interruptions and that, if given the opportunity to use these strategies, people can learn to manage them more effectively.

There are a few studies that have examined this last issue by providing participants with information about when task switches will occur. Several studies on predictive task switching, where the switches occur at regular intervals, show virtually no task switching
cost after the first switch when compared to random task switching (Rogers & Monsell, 1995; Milán, Sanabria, Tornay, & González, 2005). Cued task switches, which explicitly signal participants in advance of switches, appear to reduce the typical task switching cost dramatically (Schneider & Logan, 2007; Meiran, 1996; Monsell, Sumner & Waters, 2003; Yeung & Monsell, 2003).

Research in the interruption literature also supports the benefits of cued task switching. Altmann and Trafton’s (2002) goal-activation model is a formal model of goal encoding and retrieval that can be used to describe interruption scenarios. When the primary task is interrupted, the primary task goal must be momentarily suspended in favor of the interruption task. This goal immediately begins to decay, and must later be retrieved once the interruption is over. However, if a delay is introduced between the cue for an interruption and when the interruption task begins, then the participant has an opportunity to strengthen that goal and encode cues that will make resuming the task easier.

The studies reviewed above only provided participants more information about the timing of stitches switches, but still did not provide direct control over those switches. Studies examining free task switching performance are virtually nonexistent in the literature. Arrington & Logan (2005) claimed to allow voluntary task switching in their experiment. At each trial, participants were allowed to choose one of two tasks to complete, with the instruction that they should try to perform each task on about half of the trials, but in a random order. This type of switching is not truly free; even though the instructions were not as explicit as in other experiments. Participants were only allowed to switch at certain times, and were told how much time they should be spending on each task. However, even under
these slightly relaxed rules, they did find that the task switching cost decreased when compared to a group that was forced.

Other research suggests that, when participants have more control over the task performed and when it is performed, then the task switching costs are reduced (Schumacher, Seymour, Glass, Kieras, & Meyer, 2001). In a concurrent task paradigm, Schumacher et al. observed vast individual differences between participants. Some participants showed virtually no task interference where others were adversely affected by the introduction of a second task. These results suggest that not all participants were using the same strategy to manage the dual task scenario, with some seeming to prefer a larger processing overlap, whereas others seem to have preferred less overlap, thus leading to an increase in reaction time. Their research suggests that interference between tasks can be mediated by individual preferences in scheduling.

The above review points out many of the issues surrounding previous research. Tasks have been typically short and arbitrary, with a forced task switching schedule that does not accurately represent real world behavior. To get a better idea of what goes on cognitively when people are actually switch between tasks, the research methodology needs to grow closer to what people actually do. This will require (1) having participants perform tasks that more closely resemble a typical task and (2) having participants spend a more typical amount of time on those tasks. Most importantly, participants should be allowed to task switch freely. This would provide a more accurate understanding of how people actually switch between tasks and give more accurate results regarding how this affects performance when they do so.
In order to examine this issue, the current study allowed participants to voluntarily switch between two tasks, and compared their results to yoked participants who were forced to switch at the same time, and those who did not switch at all. The tasks that participants performed were chosen to meet the previously discussed requirements. One task was a document editing task, which may be performed in a variety of situations. The other task was a Sudoku puzzle.

These tasks were chosen for a variety of reasons. Firstly, they more closely resemble typical work tasks, because they consist of smaller steps which work towards an overall goal; secondly, they require more than mere seconds to complete. They were also chosen because they seem less likely to overlap in terms of cognitive resources (Wickens, 1984), thus aiding task switching behavior. Lastly, they were chosen as tasks that might encourage participants to switch naturally. Simply allowing voluntary switching does not guarantee that participants will actually chose to switch. These two tasks seemed to naturally provide moments where a switch may be appropriate. The document, as will be described below, was divided into several paragraphs, which may serve as break points. Sudoku is a puzzle game which may have moments where the participant is “stuck” as to what their next move should be. This also may provide the participant with an opportunity to switch tasks that remains natural.

The main hypothesis of the current study was that allowing people to switch tasks freely would reduce the task switching cost. If people freely chose to switch from one task to another, they had some internal reason for finding that choice desirable and planned to get some benefit out of it. They also would not experience the interruption effect of being told to task switch and would not show as much of a cognitive deficit when switching attention from
one task to another. Additionally, control over switching allows for prediction of switch and preparation for it. Accordingly, the effects of cued task switching were also examined. By permitting participants to switch between tasks freely, or introducing cues to allow participants to prepare for a task switch, the present research focused on the optimization of task switching.

Method

Participants

Participants were 64 undergraduate students (23 male, 41 female) enrolled in psychology courses at a large Southeastern American university. The sample had an average age of 20.7 years, with a range of 18-47. All but four participants played Sudoku less than once a week. Those four participants reported playing more than once a week but not daily. Fifteen participants were enrolled in an introductory psychology course and were recruited through a website used to sign up for participation in psychological research experiments. They received partial course credit for their participation and had the option of writing a brief research paper to receive the same credit. The remaining forty-nine participants were enrolled in one of two upper level psychology courses. These students were offered extra credit in this course for their participation.

Apparatus/Materials

Computer. The two main tasks were completed on a Macintosh iMac G4 with a full keyboard and mouse. The experimental trials were presented using a HyperCard-based program specifically designed for this experiment.
Document. The document consisted of five unrelated paragraphs. The subject matter of the five paragraphs were as follows: evolution, human speech recognition, imaginary numbers, SAT scores, and walnut trees. It contained 52 errors in spelling (e.g. “psycology”), grammar (e.g. “there” in place of “their”), capitalization (e.g. “new York”) and punctuation (e.g., not including closing quotation marks around a direct quote.) It was always one word/punctuation mark at a time that required editing; there was no reorganizing of sentence structure. All five paragraphs appeared on the screen at once, so no scrolling was required. Above each paragraph was the statement “You have corrected x out of y errors in Paragraph z.” The variable x changed each time an error was located and corrected. Samples of the types of errors were included in the instructions.

Sudoku. Sudoku is a logic-based number placement puzzle. The layout consists of nine 3x3 squares. Each horizontal row, vertical row and 3x3 square must contain the digits from 1-9 without any duplicates. The matrix was presented partially completed and the player had to logically complete the rest without making a mistake. As with all Sudoku games, there was only one way to complete the puzzle. The puzzle used in the experiment was obtained from Web Sudoku (Greenspan & Lee, n.d.) It was of a “Medium” level and consisted of 52 blank cells to be filled in during the experiment. Directions and a sample puzzle were included in participant instructions.

NASA-TLX. The NASA-TLX (Hart & Staveland, 1988) is a measure of workload. It allows participants to rate their workload in several different categories such as frustration and time demands, and to rank order which aspects they believe to be most important.
Cognitive Tests. Two cognitive tests were included in the design to test for equality within groups. The first, the Digit-symbol substitution Task (DSST) (Wechler, 1997) is a measure of perceptual speed, and includes an optional implicit learning test, which was used here. It presents participants with digit 1-9 and a symbol next to each, and a number matrix with blank boxes underneath. Participants were given 90 seconds to fill in as many of the symbols as they could without skipping any boxes. Afterwards, the participants were asked to recall as many of the symbols as they could. The second test was the Shipley Institute for Living (SILS) vocabulary subtest (Shipley, 1940). This test consisted of 40 multiple choice questions in which participants identified which of four words is closest in meaning to a target word.

Demographic Survey. After all other tasks were completed, participants were given a short demographic survey. In addition to asking age, gender, year in school, and major, the survey also asked participants how often they played Sudoku, as a test to make sure that participants did not play regularly.

Procedure

General. All participants were run individually. They first filled out an informed consent form and were told that they may quit the experiment at any time without penalty. Participants first completed the SILS followed by the DSST, which were done on paper. Following this, participants were seated at a computer where they read instructions about both tasks and the mechanisms used to switch between them. Samples of each task were also provided for participants to complete. Following the instruction period, participants were invited to ask questions. If there were none, participants began the 30 minute timed period of
the test. Half of the participants in each group started on the Document task and half of them began with Sudoku. In each yoked set of participants, the starting task was always the same. The four groups’ separate switching procedures are described below. After the tasks were completed, all groups completed the NASA-TLX, and a demographic survey.

**Voluntary Switch Group.** Participants were told that they had 30 minutes to accomplish as much as they could of both tasks in the time period allotted, and were told that they were free to switch whenever they felt it was convenient. The written instructions stated “If at any time during the experiment, you want to switch from one task to the other, just click the ‘Switch’ button. Then, work on the other task until you want to switch again. Use this button as many times as you like during session.” Participants switched between tasks by clicking a button at the bottom of the screen. The computer recorded the times that the participants switched, and the experimenter used a stopwatch to enforce the 30 minute time period.

**Forced Switch Group.** Each participant in this group was yoked to a participant in the Voluntary switch group in that they were forced to switch after the same amount of minutes and seconds that the Voluntary participant switched. They were told that they had 30 minutes total to work on the tasks and that switches would occur randomly and without warning. Time was kept with a stopwatch. Each time a switch was scheduled, the experimenter said “Switch” and participants were required to stop what they were doing immediately and switch tasks by clicking the button at the bottom of the screen.

**Cued switch group.** Participants in this group were also yoked to a participant in the Voluntary group. They were told that they had 30 minutes total to work on the tasks and that
switches would occur randomly but there would be a warning. Time was kept with a stopwatch. Ten seconds prior to a scheduled switch, the experimenter said “10 seconds” followed by “Switch” when time was up, at which point participants were required to click the button at the bottom of the screen and switch.

*No Switch Control Group.* Participants in this group also spent 30 minutes total on the two tasks, divided into two, not necessarily equal, sessions. This amount of time spent on each task was determined by calculating the total time that their partner in the Voluntary group spent on the task. Time was kept with a stopwatch. Participants were offered a break between the two tasks. Though this change from one task to the other can be thought of as a task switch, methodologically, it is not considered one, because they did not switch mid-task, but rather at the end of the task session.

**Results**

Participant and Switch Data

Group means for the cognitive test scores are presented in Table 1. There were no differences found between the four task switch groups on the Shipley Institute of Living vocabulary subtest \[F(3)=.564, p > .05\]. There were also no differences on the Digit-symbol substitution task, either in terms of the amount of symbols filled in \[F(3)=.988, p > .05\], or the number of symbols recalled \[F(3)=1.39, p > .05\].

The mean number of switches in the 30 minute period was 4.75 with a sd of 2.65 and a range of 2-11. The mode was 5 and the median was 4. On average, participants spent slightly more time on the Document task, with an mean time on task as 16:28 (sd = 4:19). The mean time on the Sudoku task was 13:31 (sd = 4:19).
Task Performance

A MANOVA was performed with the performance on each of the two tasks as the dependent variables. Performance was measured by the number of errors corrected on the document, or the number of cells completed in Sudoku, out of a maximum of 52. The independent variables included were four task switching groups and two task orders. The performance by task switching group data can be seen in Table 2.

The overall MANOVA was not significant for task switching group \( F(6,112)=1.02, p > .05 \), nor were the individual ANOVAs on the Sudoku \( F(3,56)=1.13, p > .05 \) or Document performance \( F(3,56)=.90, p > .05 \) data. Likewise, the MANOVA for task order was not significant \( F(2,55)=1.40, p > .05 \); the ANOVAs for Sudoku \( F(1,56) = .096, p > .05 \) and Document performance \( F(1,56)= 2.64, p > .05 \) were also non significant. The interaction was not significant on the overall MANOVA test \( F(6,112)=2.01, p > .05 \), nor were the individual ANOVAs on this interaction significant [Sudoku: \( F(3,56) = 2.24, p > .05 \]; Document: \( F(3,56) = 2.03, p > .05 \)].

There was no effect of task switching group on task performance, and because one of the groups included in this analysis was a No Switch Control group, this experiment observed no task switching effect. Those who switched between tasks performed no worse than those who completed the two tasks in two separate uninterrupted sessions.

Time Data

A MANOVA was performed to analyze the task switching cost observed between the groups. In this experiment, the task switching cost was defined as on the amount of time, in seconds, to complete a portion of either task directly following a switch. In the case of the
document, this refers to the time taken to locate and correct an error starting from the time a switch was made from the Sudoku task. For the Sudoku task, it was the time taken to fill in a cell in the matrix directly following a switch from the Document task. The dependent variables were (1) the reaction time following the first switch, and (2) the mean of the reaction times following all subsequent switches. This analysis replicates previous research analyses, which suggested that cued task switching can reduce the time switch cost after the first switch (Schneider & Logan, 2007; Meiran, 1996; Monsell, Sumner & Waters, 2003; Yeung & Monsell, 2003). The independent variables included were three task switching groups and two task orders. The No Switch Control group was excluded from this analysis because they made no mid-session switches.

As can be seen in Table 3, the overall MANOVA found that task switch group did not have a significant effect, \[ F(4,84) = .357, p > .05 \] and this holds true for the individual ANOVAs on first reaction time \[ F(1,42) = .587, p > .05 \] and subsequent reaction times \[ F(2,42) = .289, p > .05 \] as well. The MANOVA did however reveal a significant order effect \[ F(2, 41) = 4.94, p < .05 \]. As can be seen in Table 4, participants who began with the Sudoku task showed both a shorter reaction time following the first task switch \[ F(1,42) = 4.51, p < .05 \] and subsequent switches. \[ F(1,42) = 7.78, p < .05 \]. The interaction was not significant for the overall MANOVA \[ F(4,84) = .537, p > .05 \], nor the first RT ANOVA \[ F(2,48) = .846, p > .05 \] or the subsequent RT ANOVA \[ F(2,48) = .220, p > .05 \].

**TLX Data**

An ANOVA was performed on total NASA-TLX workload weighted total scores with task switch group as the independent variable. No significant difference was found
A two-factor mixed measures ANOVA was performed on the ratings of the 6 workload dimensions, and task switch group. There were no differences between groups \( F(3,60) = .159, p > .05 \) nor was there an interaction between group and workload dimension \( F(15,174) = .378, p > .05 \). There was a significant main effect for workload dimension \( F(5,56) = 175.95, p < .05 \).

A post-hoc test revealed that Physical Workload was rated significantly lower than all other dimensions. Temporal Workload was rated as significantly higher than Performance; Frustration was not significantly different from either. Mental Workload and Effort were rated significantly higher than the other dimensions. Group and workload dimension means are presented in table 5.

**Discussion**

The hypotheses of this experiment were that those participants allowed to switch voluntarily would perform better than those who were forced, and that those given a cue prior to switching would perform better than those without a cue. These hypotheses were not supported by the data. Participants allowed to task switch voluntarily did not perform any better, either in terms of number of items completed, or reaction time, than those who were forced to switch. Likewise, this is true for those given a cue. Furthermore, there were no group differences in perceived mental workload. Despite not being significant, these results are still noteworthy.

Most interesting is that not only was there no difference between voluntary task switching and forced task switching, as the hypothesis predicted, but there was no overall effect of task switching at all. Those in the No Switch Control group performed no better or
faster than those in the three task switching groups. They also experienced the same amount of mental workload. In other words, there was no observed task switching cost.

Why did these results not replicate previous results and produce the task switching cost so often seen? One possible explanation is differences in methodology between this experiment and those conducted previously. As was discussed above, most studies involving task switching consisted of simple, short “microtasks” with mere seconds between switches. This experiment intentionally differed from that previous research, including tasks performed at a more macro level, to more accurately mirror real life task switching performance. The nonsignificant results here lend support to the notion that what often occurs in the real world and what goes on the typical task switching experiment are not the same thing. If they were, this experiment should have found evidence of a task switching cost.

Because participants spent longer periods of time on each task session, they switched tasks less times overall than participants in typical task switching experiment. Zijlstra et al. (1999) found that being interrupted several times is more disruptive than one interruption that lasts longer. Being interrupted several times is more characteristic of past task switching experiments, and longer interruptions are more characteristic of the current study.

The two tasks used in this study were largely self-paced. Participants were free to locate and complete portions of the task in their own time. Perhaps with longer lasting, self-paced task sessions, people have more time to disengage from one task before beginning another, and feel less time pressure to perform. This may also have provided people with more cognitive resources to devote to rehearsing goals and keeping both tasks cognitively active, as opposed to the typical experimental paradigm, where many resources are devoted
to the constantly changing task goals. This design decreases the possibility that participants will experience the working memory overload that is likely experienced in experiments involving more frequent switches. Participants need not hold two task rules in their working memory during task performance but can focus on one at a time. This also frees up attentional resources that were previously being very frequently moved between tasks.

Or, perhaps, the cost of reactivating the task goal after a switch has occurred takes very little time relative to the overall time spent on task, resulting in a virtually undetectable time decrement. It can also be noted that the total task performance time in this study of 30 minutes is still much shorter than the workday seen in Czerwinski et al. (2004). People may be capable of managing task switching in these larger tasks for longer periods of time without experiencing any performance decrement, but may still experience trouble when switching every few seconds. That kind of micromanaging might be what leads to the task switching cost.

Another explanation for why a task switching cost may not have been found is that the two tasks chosen for this experiment were selected specifically because they draw on different cognitive resources; the Document task drawing on the verbal and Sudoku drawing on the numeric/logical. This may have helped participants to disengage from the tasks more easily, since the tasks were not competing for resources. Perhaps two tasks that draw upon the same cognitive resources would lead to more task interference and produce a noticeable task switching cost, even with these longer lasting tasks.

When people know that they are going to be interrupted, they can sometimes compensate for the impending interruption by working faster than they would normally
(Zijlstra et al., 1999). Participants in the current study’s task switching conditions knew that they would be asked or allowed to switch between tasks. This might have lead participants to work faster on their tasks than they would have normally, decreasing any performance decrement that might have occurred.

The sample in this study consisted solely of undergraduate university students. This younger, student population may have contained more experienced task switchers than studies done in previous years. They may be more used to task switching; both switches that occur voluntarily and those that result from an interruption. People who are accustomed to being interrupted frequently can display strategies to better manage the interruption so as to not interfere with their current task, whereas people inexperienced in this area may not do so (Zijlsta et al., 1999). It is likely that this population is experienced with this kind of situation and may have employed strategies to manage it.

Due to constraints of time and access to participants, the sample size utilized in this experiment was relatively small. Even with this limitation, the estimated power was $\beta = .75$, suggesting that a larger sample size would likely not have produced significant results. There was also a large amount of variability within this sample. The standard deviations in the samples were very large in comparison with the mean values. For example, as seen in Table 2, the Voluntary and No Switch Control groups differed in mean Document performance by 4.3. However, the standard deviations for those groups were 8.6 and 9.3 respectively. Though there were noticeable mean differences in the performance of task switching groups, the standard deviations were large enough to often exceed these
differences. Individual differences between participants could have contributed to the lack of significant findings.

This study does open up several possibilities for future research. This experiment differs greatly from previous task switching experiments both in the tasks chosen, and the voluntary task switching condition. With these differences it is no great surprise that similar results were not found. How can the benefits of voluntary task switching be examined with tasks closer to that of previous research? Perhaps this experiment could be modified to use many microtasks, rather than one larger task with many parts. Participants can be given a series of many microtasks, with a similar series of different microtasks to switch to. Some could be allowed to switch voluntarily, while other groups are forced to switch at the same time, or after a fixed amount of trials have been completed.

In conclusion, this experiment lends support to the belief that previous task switching experiments may not be generalizable to real world situations. Though the original hypothesis that voluntary task switching would improve performance over forced task switching was not supported, the tasks performed at a more macro level than those studied in the past did not produce the expected task switching cost. Further research with tasks that more closely mimic real world scenarios must be done to determine whether or not people experience this task switching cost in their daily lives, or if it only occurs in the laboratory. This line of research will produce better recommendations for workers regarding their task switching behavior.
References


APPENDIX
Table 1

*Mean*(sd) Cognitive Test Score by Task Switching Group*

<table>
<thead>
<tr>
<th>Group</th>
<th>Shipley Institute for Living Scale</th>
<th>Digit-symbol substitution task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symbols Completed</td>
<td>Symbols Recalled</td>
</tr>
<tr>
<td>Voluntary</td>
<td>30.9 (3.3)</td>
<td>66.1 (4.5)</td>
</tr>
<tr>
<td>Forced</td>
<td>30.4 (2.6)</td>
<td>66.53 (8.5)*</td>
</tr>
<tr>
<td>Cued</td>
<td>29.8 (2.9)</td>
<td>68.8 (7.1)</td>
</tr>
<tr>
<td>No Switch</td>
<td>29.7 (3.9)</td>
<td>65.0 (6.7)</td>
</tr>
</tbody>
</table>

*Note.*  *1 score missing.*  **2 scores missing.*
Table 2

*Mean(sd) Performance Data by Task and Group*

<table>
<thead>
<tr>
<th>Group</th>
<th>Document</th>
<th>Sudoku</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary</td>
<td>40.9 (8.6)</td>
<td>30.8 (11.2)</td>
</tr>
<tr>
<td>Forced</td>
<td>37.5 (8.3)</td>
<td>33.6 (13.0)</td>
</tr>
<tr>
<td>Cued</td>
<td>39.7 (8.2)</td>
<td>27.4 (13.3)</td>
</tr>
<tr>
<td>No Switch</td>
<td>36.6 (9.3)</td>
<td>26.6 (11.6)</td>
</tr>
</tbody>
</table>
Table 3

*Mean*(sd) Reaction Time (in seconds) Following Task Switch by Task Switch Group

<table>
<thead>
<tr>
<th>Group</th>
<th>First Switch</th>
<th>All Other Switches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary</td>
<td>14.4 (10.4)</td>
<td>29.3 (16.6)</td>
</tr>
<tr>
<td>Forced</td>
<td>19.9 (23.1)</td>
<td>35.2 (24.0)</td>
</tr>
<tr>
<td>Cued</td>
<td>19.0 (10.9)</td>
<td>34.4 (32.5)</td>
</tr>
</tbody>
</table>
Table 4

*Mean (sd) Reaction Time (in seconds) Following Task Switch by Task Order Condition*

<table>
<thead>
<tr>
<th>Group</th>
<th>First Switch *</th>
<th>All Other Switches*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sudoku First</td>
<td>13.1 (5.4)</td>
<td>23.3 (15.4)</td>
</tr>
<tr>
<td>Document First</td>
<td>22.5 (20.8)</td>
<td>42.6 (28.8)</td>
</tr>
</tbody>
</table>

* $p < .05$
<table>
<thead>
<tr>
<th>Group</th>
<th>Mental Demand</th>
<th>Effort</th>
<th>Temporal Demand</th>
<th>Frustration</th>
<th>Performance</th>
<th>Physical Demand</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary</td>
<td>14.6 (2.7)</td>
<td>14.6 (3.0)</td>
<td>11.12 (2.5)</td>
<td>10.5 (5.3)</td>
<td>9.3 (3.2)</td>
<td>1.7 (.8)</td>
<td>64.6 (10.4)</td>
</tr>
<tr>
<td>Forced</td>
<td>14.8 (2.8)</td>
<td>14.0 (3.0)</td>
<td>11.3 (5.7)</td>
<td>9.0 (5.1)</td>
<td>8.8 (2.8)</td>
<td>1.9 (1.2)</td>
<td>62.3 (10.0)</td>
</tr>
<tr>
<td>Cued</td>
<td>13.81 (3.2)</td>
<td>13.5 (3.1)</td>
<td>11.2 (4.0)</td>
<td>10.3 (4.9)</td>
<td>9.9 (3.8)</td>
<td>2.4 (2.3)</td>
<td>63.4 (8.3)</td>
</tr>
<tr>
<td>No Switch</td>
<td>15.63 (2.9)</td>
<td>14.2 (2.6)</td>
<td>11.3 (3.9)</td>
<td>8.9 (5.2)</td>
<td>9.7 (4.4)</td>
<td>2.4 (2.7)</td>
<td>65.8 (10.8)</td>
</tr>
<tr>
<td>Average</td>
<td>14.72\textsubscript{1} (2.9)</td>
<td>14.1\textsubscript{1} (2.9)</td>
<td>11.2\textsubscript{2} (4.3)</td>
<td>9.7\textsubscript{2,3} (5.1)</td>
<td>9.4\textsubscript{3} (3.6)</td>
<td>2.1\textsubscript{4} (2.9)</td>
<td>63.7 (9.78)</td>
</tr>
</tbody>
</table>

*Note.* Means having the same subscript are not significantly different at $p < .01$ in the Tukey HSD test.
Figure 1. Sample Yoking of Task Switching Groups