Effects of Working Memory Capacity, Task Switching, and Task Difficulty on Multitasking Performance

Shijing Liu[1], Amy Wadeson[1], Na Young Kim[1], & Chang S. Nam[1]
[1]Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University, Raleigh, NC 27695, USA

Multitasking requires human operators to handle the demands of multiple tasks through task switching at the same time and this ability is required in many jobs. Previous studies showed that different levels of working memory capacity (WMC) and task switching abilities can lead to differences on multitasking performance. With increased complexity of tasks, maintaining task performance is challenging. This study sought to find the relations of WMC, task switching, task difficulty, and multitasking performance. Multi-Attribute Task Battery II (MATB-II) was employed in this study as a platform to assess multitasking. Automated OSPAN and Trail Making Tasks (TMT) were used to assess WMC and the task switching ability, respectively. Results indicated that there were significant effects of these three parameters on multitasking performance. Other dimensions of multitasking performance will be addressed in future studies.

INTRODUCTION

Multitasking performance requires the ability to execute multiple actions in the same time period while frequently switching between individual tasks (Colom, Martínez-Molina, Shih, & Santacreu, 2010; Delbridge, 2000). Multitasking consists of two different phases of activity: attending to individual constituent tasks, and switching between tasks (Judd & Kennedy, 2011; Rubenstein, Meyer, & Evans, 2001). It also requires management of the overall system at once. Several studies have found that there is a measurable cognitive cost associated with frequent task switching during multitasking. This causes significant degradation in performance and increases the time required to complete the tasks (Fischer, Morris, & Joslyn, 2003; Pashler, 2000; Rubenstein et al., 2001).

It has also been found that as cognitive workload increases with the increase in task complexity, maintaining task performance within an acceptable range becomes more difficult (Smith, Gevins, Brown, Karnik, & Du, 2001; Wickens, 2008). Cognitive functions such as changing recently stored information and inhibiting certain cognitive procedures are critical for multitasking performance. Working memory is responsible for such cognitive functions (Konig, Buhner, & Murling, 2005; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000; Oberauer, Süß, Wilhelm, & Wittmann, 2003), as it is the ability to store and process sensory input or information in order to perform multiple tasks (Gohar et al., 2009; Heitz, Unsworth, & Engle, 2004). It contains the passive system related to temporary storage of information as well as the active system responsible for the executive control of cognitive processes (Konig et al., 2005), which are critical for multitasking performance. Konig et al. (2005) proposed that working memory capacity (WMC) can be conducive to switching from one task to another when working on a number of tasks simultaneously. As such, WMC has been reported as a predictor of multitasking performance (Bühner, König, Pick, & Krumm, 2006; Konig et al., 2005).

Previous studies found that both WMC and the task switching ability had effects on multitasking performance. In addition it has been reported that changes in user performance can be a consequence of varying levels of task difficulty (Kane & Engle, 2003; Konig et al., 2005; Redick & Engle, 2006; Sanchez-Cubillo et al., 2009). However, there is a lack of research on the relations among WMC, task switching, task difficulty, and multitasking performance. Therefore, these relationships are the focus of our study.

This study sought to find the influence of WMC, task switching, and task difficulty on multitasking performance. Multi-Attribute Task Battery II (MATB-II) is a computer-based task designed to evaluate operator performance and workload (Santiago-Espada, Myer, Latorella, & Comstock, 2011) and was employed in this study as a platform to assess multitasking. In the following sections, we will introduce the experiment, present results, and discuss our findings.

METHOD

Participants

A total of 10 healthy university students (female = 4, male = 6, mean age = 25.4 years, SD = 2.46) participated in this study. All participants were right handed, had normal hearing and normal or corrected-to-normal vision, and had no previous experience with MATB-II. In addition, they were all native English speakers.

Apparatus and Materials

All participants were required to complete three sessions during the experiment: (1) An automated OSPAN task to measure their working memory capacity (WMC); (2) a set of

Trail Marking Tasks (TMT) to estimate their task switching ability; and (3) a set of Multi-Attribute Task Battery II (MATB-II) tasks to investigate their performance in multitasking.

Automated OSPAN. WMC can be predicted by working memory span tasks (Kane & Engle, 2003; Redick & Engle, 2006). For instance, the operation span task (OSPAN) (Turner & Engle, 1989) has been established as a reliable and valid marker of WMC and an automated version of OSPAN is employed in current study (Unsworth, Heitz, Schrock, & Engle, 2005). This version of OSPAN is a computerized test that includes items (letters) to remember and a distracting activity in the form of math problem solving and allows the participant to complete the task independently of the experimenter (Unsworth et al., 2005). There were three practice sessions (letter span, math problem, both of them combined) and one experimental session. During the experimental session, participants were required to complete the tasks as accurately and quickly as possible. There were 75 trials in the experimental session and score range of this test was 0 – 75. The task took approximately 20 minutes to complete for each participant.

Trail Marking Tasks (TMT). TMT have been widely used to test the ability of task switching and other cognitive processing abilities (Arbuthnott & Frank, 2000; Bowie & Harvey, 2006; Kortte, Horner, & Windham, 2002; Sanchez-Cubillo et al., 2009). The TMT is a two part paper and pencil task and the comparison of the scores from these two parts have revealed a significant correlation to switch-costs in task-switching paradigms (Sanchez-Cubillo et al., 2009). Therefore, task switching ability is measured and predicted by TMT in our study.

The TMT includes two parts: A and B. In TMT-A the numbers 1-25 are strewn across a page and the participant connects them in numerical order, from least to greatest, as quickly as possible. The second part (TMT-B) requires the participant to switch between numbers (1-12) and letters (A-L), connecting them in numerical and alphabetical order, as quickly as possible. Before each part of the TMT, a small example was given to the participant. The score of each part is recorded separately and reported in seconds. If during the task the participant skipped a number or a letter, the mistake was verbally addressed by the experimenter and the participant was required to fix it before continuing. The score was recorded in seconds only, accounting for the fact that any mistakes would automatically increase the total time to complete the task.

Multi-Attribute Task Battery II (MATB-II). The MATB-II is a computer-based task designed to evaluate operator performance and workload (Santiago-Espada et al., 2011). It was employed in this study as a platform to assess multitasking. Figure 1 shows the interface of MATB-II.

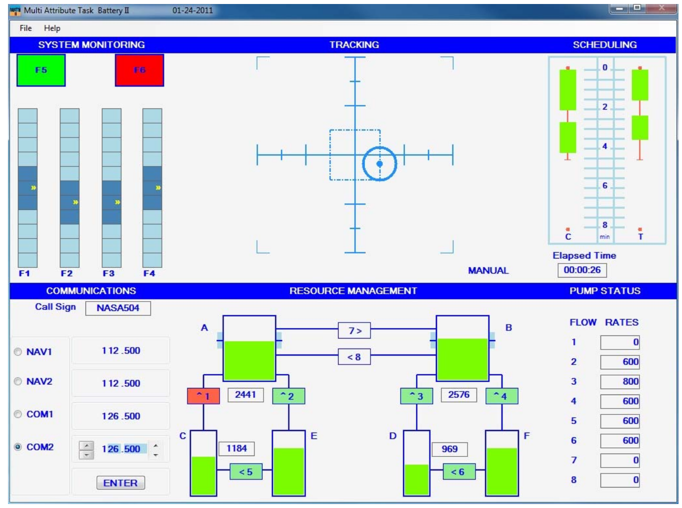

Figure 1 Multi-Attribute Task Battery II (MATB-II) tasks

The tasks in MATB-II are analogous to activities that aircraft crewmembers perform in flight, yet are accessible to populations with no aviation experience (Comstock & Arnegard, 1992; Santiago-Espada et al., 2011). There were four MATB-II tasks applied in this study: Light, Scale, Tracking, and Communication (COMM). For the Light task, participants needed to respond to the absence of the green light and the presence of the red light (top left of Figure 1). For the Scale task, participants needed to monitor the four moving scales (left, middle of Figure 1) and click on the bars when they detected a deviation from the midpoint. For the COMM task (bottom left of Figure 1), audio messages with particular "callsign" were played and participants were required to listen to the content and respond by selecting the appropriate radio channel and frequency on the display. When the Tracking task was under manual mode, participants were asked to use a joystick by their left hands to keep the circular target in the center of the tracking window (top, middle of Figure 1).

In the operation of dual or multiple tasks, mental workload can reflect user performance and it is multidimensional and multifaceted (Cain, 2007; Wickens, 2008). Subjective rating of mental workload can be assessed with the NASA-task load index (TLX) (Hart & Staveland, 1988). This rating scales include six parts: mental demand (MD), physical demand (PD), temporal demand (TD), and the individual's perceived level of performance (PE), effort (EF), and frustration (FR). NASA-TLX is used to estimate user mental workload during multitasking in present study.

All participants were required to use a mouse with their right hand to respond to Light, Scale, and COMM tasks. A NASA-TLX questionnaire was presented after each trial.

Experiment Design and Procedure

Each participant completed three sessions: TMT, OSPAN and MATB-II. These three test sessions were independent. The order of three sessions were randomly assigned to each participant. There were training sessions before OSPAN and MATB-II tasks and no training session for TMT.

For MATB-II tasks, three levels of task difficulty (low, medium, and high) were set up by adjusting the event rate of each subtask and, in the case of the Tracking task, difficulty level was set up by manipulating speed and number of directional changes. There were 6 (3 levels × 2 replicates) MATB-II trials for each participant. Each trial of the tasks lasted 5 minutes and participants were required to perform four tasks simultaneously during each trial. Participants completed the NASA-TLX questionnaire after each trial and rested for 3 minutes. All participants were tested individually in a laboratory and the experiment took 1.5 hours for each participant.

RESULTS

For each participant, their TMT scores, OSPAN scores, and MATB-II performance were recorded. MATB-II performance includes average response time and response ratio (correct response per trial) for Light, Scale, and Communication task, mean root mean square (MRMS) of the distance between target and the center point for Tracking task, and mean NASA-TLX rating score per trial. Since one of the defining feature of regression and correlation is the ability to assess unique or partial relationships (Cohen, Cohen, West, & Aiken, 2013), regression and correlation analyses were carried out to test the effect of WMC (OSPAN score), task switching ability (TMT score), and level of MATB-II task difficulty (Level) on multitasking performance.

Results of overall multitasking performance showed that all three main factors (Level, OSPAN, and TMT) were significant ($p = 0.000$, 0.000, and 0.005, respectively). Meanwhile, two interaction factors (Level × OSPAN and OSPAN × TMT) were found to be significant or marginally significant on overall multitasking performance ($p = 0.010$ and 0.054, respectively).

Results of response time (RT) indicated that level of task difficulty was significant on the Scale task and OSPAN was significant on the COMM task. No significant interaction factors were found significant on response time. Table 1 shows the parameter estimates of regression analysis for response time and parameters of each model.

Table 1
Parameter estimates of regression analysis for response time (RT) of Light, Scale, and COMM tasks

	RT_Light	RT_Scale	RT_COMM
Level	0.139	0.548**	0.190
OSPAN	-0.004	0.006	-0.015*
TMT	0.006	-0.001	-0.011
Level × OSPAN	-0.003	0.0005	-0.009
Level × TMT	-0.004	-0.003	0.002
OSPAN × TMT	0.0003	-0.001	0.002
Level × OSPAN × TMT	0.0002	-0.0005	0.001
R^2	0.093	0.264	0.204
Adj R^2	-0.032	0.163	0.095
F(7, 52)	0.7439	2.6148	1.8649

Note. * significance level: $p < 0.05$. ** significance level: $p < 0.01$.

Results of response ratio (RR) indicated that OSPAN had significant effects on both Light and COMM tasks. No significant interaction factors were found significant on response ratio. Table 2 shows the parameter estimates of regression analysis for response time.

Table 2
Parameter estimates of regression analysis for response ratio (RR) of Light, Scale, and COMM tasks

	RR_Light	RR_Scale	RR_COMM
Level	-0.004	-0.001	0.005
OSPAN	-0.0004*	-0.0002	0.002**
TMT	0.0002	0.0001	0.002
Level × OSPAN	-0.0002	0.0009	-0.0008
Level × TMT	0.0004	-0.0006	-0.001
OSPAN × TMT	0.0000	0.0000	-0.0001
Level × OSPAN × TMT	0.00004	-0.0000	0.0000
R^2	0.175	0.047	0.180
Adj R^2	0.062	-0.084	0.067
F(7, 52)	1.5491	0.3603	1.5992

Note. * significance level: $p < 0.05$. ** significance level: $p < 0.01$.

Results of MRMS for Tracking task showed that main factors, OSPAN and TMT, had a significant effect on the performance of Tracking task. Interaction factor of Level and OSPAN also indicated significant effect. Analysis of NASA-TLX rating scores implied that all three main factors had significant effects. No significant interaction factors were found. Table 3 shows the parameter estimates of Tracking task and NASA-TLX rating score.

Table 3
Parameter estimates of regression analysis for Tracking task (MRMS) and NASA-TLX rating score (TLX)

	MRMS	TLX
Level	16.191**	14.097**
OSPAN	-0.251**	-0.512**
TMT	-0.085	-0.775**
Level × OSPAN	-0.207*	-0.035
Level × TMT	0.136	0.006
OSPAN × TMT	0.016	0.031
Level × OSPAN × TMT	0.021	-0.019
R^2	0.829	0.557
Adj R^2	0.806	0.496
F(7, 52)	35.4381	9.1612

Note. * significance level: $p < 0.05$. ** significance level: $p < 0.01$.

The relation between three main factors and multitasking performance were also investigated with correlation analysis. Table 4 shows the correlation coefficient.

Table 4
Correlation between Level, OSPAN, TMT, and multitasking performance

	level	OSPAN	TMT	RT_ Light	RT_ Scale	RR_ Light	RR_ Scale	RT_ COMM	RR_ COMM	TRCK_ MRMS	WRS
Level	1.00	0.02	-0.01	0.02	0.49	-0.20	0.01	0.19	0.05	0.86	0.58
OSPAN		1.00	-0.18	-0.12	0.07	-0.21	-0.06	-0.19	0.28	-0.18	-0.25
TMT			1.00	0.09	0.03	0.00	0.03	-0.19	0.21	-0.06	-0.31
RT_Light				1.00	0.58	0.12	-0.03	0.47	-0.15	0.37	0.35
RT_Scale					1.00	-0.16	0.04	0.46	-0.18	0.56	0.42
RR_Light						1.00	0.32	0.19	-0.23	-0.10	0.00
RR_Scale							1.00	-0.21	-0.11	-0.01	-0.07
RT_COMM								1.00	-0.37	0.46	0.48
RR_COMM									1.00	-0.08	-0.27
TRCK_MRMS										1.00	0.67
WRS											1.00

DISCUSSION

The results revealed the relations among the level of task difficulty, WMC, task switching ability, and multitasking performance. The overall model indicated that these three parameters had significant effects on multitasking performance. Correlation analysis implied there was a significant relationship between these parameters and user performance. With the increase in level of task difficulty, user performance on each subtask of MATB-II decreased (e.g., response time, reaction ratio and accuracy) and their mental workload increased. Users with a higher level of WMC (OSPAN scores) showed a better performance on most subtasks in MATB-II. Stronger task switching ability (TMT) also indicated a better overall performance on multitasking.

The findings of response time on Light, Scale, and COMM tasks predicted that users with a lower level of task difficulty or higher WMC were able to make a response to the stimuli quicker. These results matched our expectation, although it did not show significant effect of the interactions between these parameters.

The results of response ratio of the subtasks demonstrated that only WMC had significant effect on the performance. Users with a higher WMC had completed more COMM tasks correctly, as we expected. The result of response ratio for the Light task was the opposite, where higher levels of WMC produced a lower performance on Light task.

The findings of the tracking task showed significant effects from the level of task difficulty and WMC, but not the task switching ability. These results indicated that users who had a higher WMC were able to control their targets more accurately and steadily with a lower level of task difficulty. Since participants were required to control the joystick continuously with their left hands during experimental trials, their left hands did not need to switch to other subtasks and the ability of task switching did not show significance on this subtask.

The results of mental workload from NASA-TLX rating showed significance of all three main factors. With the increase in the level of task difficulty, participants experienced a higher level of perceived mental workload. Under the same setup of MATB-II tasks, users with a higher level of WMC or task switching ability showed less mental workload during the experiments.

CONCLUSION

Findings of the present study demonstrated the significant influence of the level of task difficulty, WMC, and task switching ability on multitasking performance. The results revealed the relations of these parameters and user performance from different dimensions. The current study is limited to participants consisting of only university students. Future studies could recruit a wider group of participants which will cover a wider range of WMC and task switching ability. Furthermore, the current study only included a short session of training on MATB-II operation and no participants had previous experiences of MATB-II or pilot operation. User performance tends to show interesting patterns after a long term of systematic training. Future studies should investigate the effect of training on user performance.

REFERENCES

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of clinical and experimental neuropsychology*, 22(4), 518-528.

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature protocols*, 1(5), 2277-2281.

Bühner, M., König, C. J., Pick, M., & Krumm, S. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human Performance*, 19(3), 253-275.

Cain, B. (2007). A review of the mental workload literature. Defence Research and Development Toronto (Canada).

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Colom, R., Martínez-Molina, A., Shih, P. C., & Santacreu, J. (2010). Intelligence, working memory, and multitasking performance. *Intelligence*, 38(6), 543-551.

Comstock Jr, J. R., & Arnegard, R. J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research.

Delbridge, K. A. (2000). Individual differences in multi-tasking ability: Exploring a nomological network.

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Word knowledge and working memory capacity. *Journal of Educational Psychology*, 82, 799-804.

Fischer, S. C., Morris, K. A., & Joslyn, S. (2003). Measuring multi-tasking ability: Report prepared for the Office of Naval research by Anacapa Sciences.

Gohar, A., Adams, A., Gertner, E., Sackett-Lundeen, L., Heitz, R., Engle, R., ... & Bijwadia, J. (2009). Working memory capacity is decreased in sleep-deprived internal medicine residents. *J Clin Sleep Med*, 5(3), 191-197.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52, 139-183.

Heitz, R. P., Unsworth, N., & Engle, R. W. (2004). Working memory capacity, attention control, and fluid intelligence. *Handbook of understanding and measuring intelligence*, 61-77.

Judd, T., & Kennedy, G. (2011). Measurement and evidence of computer-based task switching and multitasking by 'Net Generation'students. *Computers & Education*, 56(3), 625-631.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology*: General, 132(1), 47.

Konig, C. J., Buhner, M., & Murling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human performance*, 18(3), 243-266.

Kortte, K. B., Horner, M. D., & Windham, W. K. (2002). The trail making test, part B: cognitive flexibility or ability to maintain set. *Applied neuropsychology*, 9(2), 106-109.

Oberauer, K., Süß, H.-M., Schulze, R.,Wilhelm, O., & Wittmann,W.W. (2000).Working memory capacity facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017–1045.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31, 167–193.

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell, & J. Driver (Eds.), Attention and Performance XVIII: *Control of mental processes*. Cambridge, MA: MIT Press.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, 20(5), 713-721.

Rubenstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Cognitive processes in task switching. Journal of Experimental Psychology: *Human Perception and Performance*, 27(4), 763–797.

Sanchez-Cubillo, I., Perianez, J. A., Adrover-Roig, D., Rodriguez-Sanchez, J. M., Rios-Lago, M., Tirapu, J. E. E. A., & Barcelo, F. (2009). Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, 15(3), 438.

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock Jr, J. R. (2011). The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(3), 366-380.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent?. *Journal of memory and language*, 28(2), 127-154.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span ask. *Behavior research methods, 37*(3), 498-505.

Wickens, C. D. (2008). Multiple resources and mental workload. Human Factors: *The Journal of the Human Factors and Ergonomics Society*, 50(3), 449-455.