Proceedings of the Annual Meeting of the Cognitive Science Society

Cognitive Workload and the Motor Component of Visual Attention

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Ralph, Jason
Gray, Wayne
Schoelles, Mike

Publication Date
2012

Peer reviewed

Cognitive Workload and the Motor Component of Visual Attention

Jason Ralph, Wayne D. Gray, & Michael J. Schoelles
Rensselaer Polytechnic Institute

Abstract

Outside the laboratory, the ability to control visual input during multiple task performance by controlling *where the eyes look and when* is an obvious component of multiple task performance. However, inside the laboratory researchers either obviate the control of the eyes by presenting information from one task at a time or are oblivious to the need for *just-in-time* control of the motor component of visual attention. We investigate the effects of cognitive workload on eye movements in a paradigm that controls the demand on the eyes as an input channel while increasing workload by increasing the demand on working memory. Despite constant visual demands, we find that fixations become more scattered with increasing working memory load.

Keywords: dual mechanisms of control, cognitive control, cognitive workload

Introduction

Our ability to switch among multiple tasks has been the subject of extensive research for many decades (Allport, Styles, & Hsieh, 1994; Altmann & Gray, 2008; Broadbent, 1952; Cherry, 1953). People generally exhibit a multitask effect in which they are slower and commit more errors when performing multiple concurrent tasks than a sequential series of the same tasks. Most theories of multitasking propose a resource capacity explanation for this multi-task effect. These theories, including Wicken's multiple resource theory (MRT) (Wickens, 2002; Wickens & Colcombe, 2007), and Salvucci and Taatgen's threaded cognition (Salvucci & Taatgen, 2008; Salvucci & Taatgen, 2011), propose a variety of resources which must be shared by concurrent tasks. They argue that when the capacity of one of these resources is exhausted, the cognitive system must wait for the bottleneck to clear, which causes slower task execution.

MRT identifies input modalities (visual or auditory), response modalities (motor, vocal), and cognition as resources responsible for multi-task effects (Wickens, 2002; Wickens, 1992; Wickens, Goh, Helleberg, Horrey, & Talleur, 2003). For instance, if two tasks require visual processing, performance will be slower because the capacity of the visual processor is limited. MRT has not been instantiated in computational form but its assumptions appear compatible with other capacity theories such that multiple tasks can be performed without interference until some capacity limit is reached (e.g., see Just, Carpenter, & Hemphill, 1996a, 1996b).

Threaded cognition is a model based procedural theory of task switching, which is implemented within the ACT-R cognitive architecture (Anderson et al., 2004). In this theory, "concurrent multitasking emerges from the interaction of autonomous process threads in conjunction with a straightforward mechanism for resource acquisition and conflict resolution," (Salvucci & Taatgen, 2008, p. 102). Although we see Threaded Cognition as a major advance in the modeling of complex cognition, at present it focuses on the control of task switching per se, not on the control of tasks.

As detailed by the Dual Mechanisms of Control (DMC) theory (Braver, Gray, & Burgess, 2007; Braver, 2012), the brain's ability to prepare ahead of time (proactive control) is limited. Proactive control requires that the strategy and information for a task be kept active. Hence, proactive control places strategies such as the verbal-articulatory loop (Baddeley, 2012) and updating, shifting, and inhibition (Miyake et al., 2000) under control of the PFC and subject to limits in PFC processing. The importance of proactive control depends on the nature of the task being performed. In situations with a limited number of stimuli and responses (i.e., most experimental psychology tasks), then proactive control seems best. However, in tasks that require the subject to respond to one of a number of different stimuli in a number of different ways (e.g., driving on the freeway during rush hour) then reactive processing seems required. Cognitive control is recruited by a mix of proactive and reactive influences, and understanding how this mix changes based on task demands is key to understanding performance of multiple tasks.

We hypothesize that differences between single-task and multi-task performance depend on how the brain manages the demands on proactive control. A major limiting factor in multitasking is our ability to distribute our attention to perform multiple concurrent tasks. In visual tasks, the capacity of the eyes to focus on one part of the visual scene at a time is a factor constraining our performance. Many tasks require the ability to control eye movements to capture required information from the world, or to maintain sustained focal attention.

The ability to control our eyes during task performance is an overlooked but critical variety of control. Unless we assume that pointing-the-eyes at potentially informative areas of visual interest is somehow both automatic and effortless, then the need to control gaze to optimize task performance adds yet another burden to our control mechanisms. Hence, it may be possible to determine when a switch in control mode has occurred by examining patterns of eye fixations. We hypothesis that effective proactive control results in long, constant fixations on the area of the visual scene related to a task. As the demands on proactive control increase, eye movements should become more scattered, evidenced by short fixation durations spread across a wider area. After presenting our paradigm and results, we return to speculative discussion of this issue in our Discussion section.

Previous Research

In prior research, we found that performance in a dual-task paradigm was affected by requiring subjects to look in different locations for information needed for different tasks (Ralph, Gray, & Schoelles, 2010). Subjects performed a continuous visual tracking task while concurrently performing an n-back style memory task. (The paradigm will be explained in detail in the methods section below.) Consistent with an MRT prediction, subjects who received information aurally outperformed those who received it visually. However, we could not determined whether auditory instruction relieved the cognitive demands of the task, or simply the visual demands.

The DMC account implies that proactive control of behavior relies on a common set of PFC mechanisms. Viewing the eye as something that needs to be controlled, we hypothesize that it is the extra control of the eye that contributes to the auditory vs visual tasks differences, not necessarily the differences in auditory vs visual processing (as implied by MRT).

To test this hypothesis, the current study compares two visual conditions which share the same visual demands, but differ in the cognitive (working memory) requirements. This differs from most MRT research, which typically increases cognitive load by increasing the difficulty of visual tracking (Wickens et al., 2003). In the current study, differences in eye movement behavior and performance cannot be attributed to increased demands on the eye as these demands are identical across conditions. Thus the focus of the current study is on how increased demands for proactive control affect eye movement strategies and performance in the absence of additional demands for the control of the eye.

The Study

The NavBack paradigm was designed to collect detailed empirical data in a task at the approximate complexity of those used by Wickens in testing MRT (e.g., Martin-Emerson & Wickens, 1992, 1997). The NavBack task combines a tracking task with a working memory task. The tracking task requires the subject to keep an arrow centered as it *jitters* (i.e., moves randomly) from side to side. Concurrently, subjects perform a continuous working memory span task, which is similar to the n-back memory task (Gevins & Smith, 2003; Jaeggi et al., 2010; McEvoy, Smith, & Gevins, 1998). In the high workload conditions, subjects must maintain a list of instructions in memory of how to *turn* in the next three intersections (e.g., “left”, “right”, or “forward”). After each intersection the subject has to delete the just completed instruction and add a new instruction to the end of his mental list. In the low workload condition, only one instruction must be maintained.

Methods

Subjects

22 undergraduate students of Rensselaer Polytechnic Institute (mean age = 19) volunteered to participate in this study for course credit. Eleven subjects were randomly assigned to each of two conditions: *High Memory* or *Low Memory*.

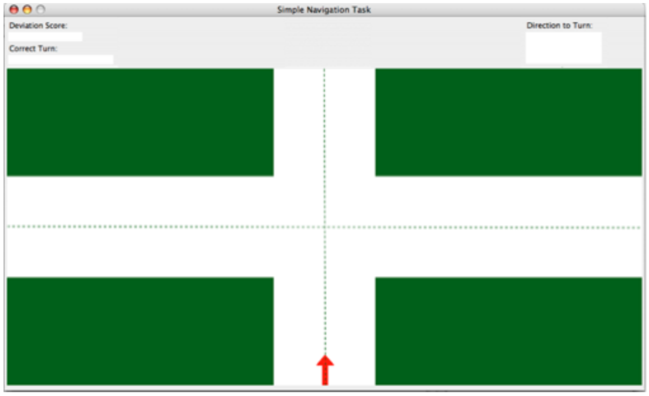

Figure 1. Screenshot of the NavBack Paradigm. Subjects must keep the arrow in the center of the road as they receive turn instructions in the box at the upper right corner. The arrow remained at the bottom of the screen, and the ”road” scrolled downward, simulating forward movement. Subjects could turn when the arrow was in the intersection.

Apparatus and Materiels

The experiment was run on an Apple Mac Mini computer (running Mac-OS 10.4) at a 1024x768 screen resolution. Eye fixation data were collected using an LC Technologies tracker at a 120 Hz sampling rate. A chin-rest was used to stabilize head movements and ensure a fixed viewing distance of 60 cm. The NavBack software is a custom application implemented in Lispworks 5.1.

Design

Between-Subject Condition: Memory Load. Each memory load condition received a new direction (the turn to make in the future) while traveling through a “city block”. For the *High Memory* condition, the new instruction specified how to turn three intersections in the future. In the *Low Memory* condition, the new instruction specified how to turn at the next intersection. Hence, the memory load for the High vs Low Memory conditions was three versus one items.

Within-Subject Condition: Instruction Presentation Time. All subjects received the new turn instruction *early, middle, or late* during their *travel* through a city block. The instruction appeared for 2-s beginning either 1-s, 3.4-s, or 5-s after the arrow exited the intersection and entered the next city block. On any given episode cycle, whether the instruction was presented early, middle, or late was determined randomly.

Procedure

Each subject completed a 2-min practice session to familiarize them with the demands of the task, followed by eight 5-min experimental blocks. Each 5-min block consisted of a continuous series of episode cycles. Each cycle began when the tip of the arrow left an intersection and entered the next city block (city blocks are the green areas in Figure 1). At one of three randomly chosen times a new instruction appeared in the direction box on the upper right of the screen (see Figure 1). Travel time through each city block was 6-s.

Although, once in the intersection, subjects could turn at any time, minimizing the jitter score required the subject to turn at the exact center of the intersection. The animation for the turn added 1,500 ms to the time spent in the intersection when subjects made left or right turns. During each episode cycle subjects had to do two related tasks: the jitter task and the turn direction task. The jitter task is a visual-motor task requiring constant attention. The turn-direction task requires monitoring for the appearance of a new turn direction while performing the jitter task. Depending on condition subjects needed to hold either one or three turns in memory. Each new episode cycle required them to update the list of items held in memory.

Jitter Task: Visual-Motor

Subjects were instructed to keep the arrow in the center of the road (on the dotted line in Figure 1) as the arrow *jittered* actively from side to side, every 200 ms, based on a pseudo-random function. Subjects corrected the arrow's horizontal position by pressing the *a* and *d* keys on a standard keyboard. Their goal was to keep the arrow as close to the center of the lane as possible. The arrow's position at the beginning of each city block was determined by the timing of the previous turn. If the previous turn (left or right) was initiated at the exact vertical center of the intersection, then the arrow began the next city block in the center of the lane. If the turn was initiated early or late, it began the next city block deviated from the center by an amount proportional to the distance from the turn point to the center of the intersection. (Ss were not instructed on this aspect of the task.) The computer logged the absolute value of the number of pixels deviated from the center every 200 ms. We refer to this value as "jitter score".

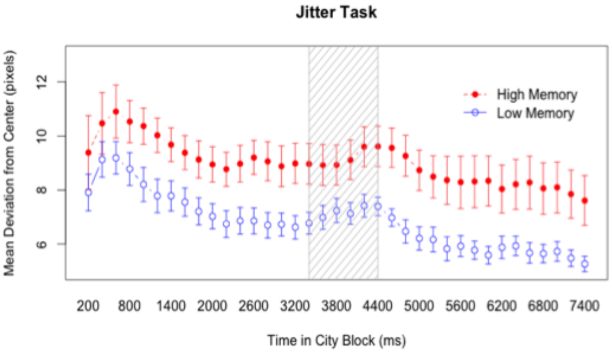

Figure 2. Mean jitter scores for middle instruction city blocks. The gray rectangle denotes when the instruction was on-screen.

Turn Direction Task: Working Memory Updating

Concurrent with the Jitter Task, subjects were required to keep either one or three turns in memory (e.g., left, right, forward). At the beginning of a 5-min set, high memory subjects were presented with the initial three turns that were to be made in the first, second, and third intersection. After the set began, new turn directions appeared at one of the three times described earlier. The direction presentation time was randomly chosen for each city block.

As discussed previously, success in the high memory condition essentially required subjects to rehearse, update, and maintain a list of three instructions. Subjects had to mentally delete the instruction for the just completed turn and to append a new instruction at the end of their mental list. Subjects in the low memory group had to remember only the most recently presented direction. Feedback on the correctness of the most recent turn was available at the top left of the screen (see Figure 1).

Results

Jitter Scores. An analysis of variance was performed on jitter scores by memory load and instruction time and revealed main effects of memory load $F(1,20)=7.96\ p<.02$, and instruction time $F(2,40) = 17.31\ p<.001$. High memory subjects (M=10.42 pixels SD=2.03) had higher mean jitter scores than low memory subjects (M=8.38 pixels, SD=1.05). Figure 2 shows the mean jitter scores throughout an average city block (for middle instruction times). Low memory subjects outperformed high memory subjects throughout the city block.

Turning Task. The turn results were as expected with the Low Workload group more accurate (94.64%) than the High Workload group (85.31%). These differences were significant ($F(1,20) = 12.57\ p<.01$) but will not be discussed further in this short report.

Fixation Locations. Analysis of the eye data was performed using the areas of interest displayed in Figure 3. Only data recorded while the arrow was within a city block is included (e.g. not in an intersection). The data yielded several differences in fixations between the high vs low memory conditions. Figure 4 shows the proportion of time spent looking at the arrow, direction box, and road areas. High memory subjects spend less time fixated on the arrow and more time looking at the direction box throughout the city block. This difference is most pronounced before the instruction appears (compare Figure 5 top and bottom for those times before the appearance of the instruction), suggesting that low memory subjects employed a more economical strategy to monitor the direction box for the appearance of the arrow. As Figures 4 and 5 demonstrate, compared to low memory subjects, high memory ones devoted proportionately more time looking at the direction box and less time looking at the arrow[1].

The fixation pattern for low memory subjects likely reflects a proactive strategy designed to maximize time spent fixating on the arrow while maintaining the ability to monitor for instructions. It is a strategy that high memory subjects were either unwilling or unable to pursue.

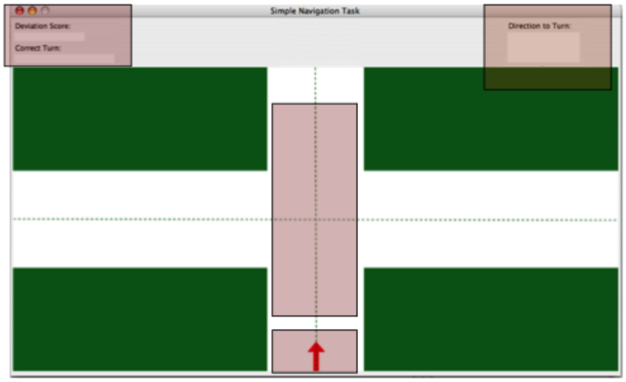

Figure 3. Areas of Interest used for the analysis of eye fixations. Shaded areas represent the arrow, direction box (upper right), feedback (upper left), and road areas.

Eye Movements Reveal Task Priority Differences with Workload.

Figures 4 and 5 limit the data they present to the four task relevant areas shown in Figure 3. We find it both interesting and important that the two conditions spent different proportions of time gazing at screen areas relevant to different parts of the task. The low workload condition spent more time on jitter related areas (e.g. arrow), and the high memory condition spent more time looking at areas relevant to the turning task (e.g direction box, road).

We further examined the degree of task focus by conducting an analysis of the variability of eye fixations throughout the screen area for the two conditions (see Figure 6). For this analysis, we split the screen into 30x30 pixel boxes and computed the standard deviation of the number of fixations per box for each subject. By this measure, increasing standard deviations show that subjects focus more on some of the pixel boxes than others. The lower the standard deviation, the more evenly spread the fixations are across all pixel boxes. An analysis of variance performed on this measure shows that the fixations of high memory subjects were significantly

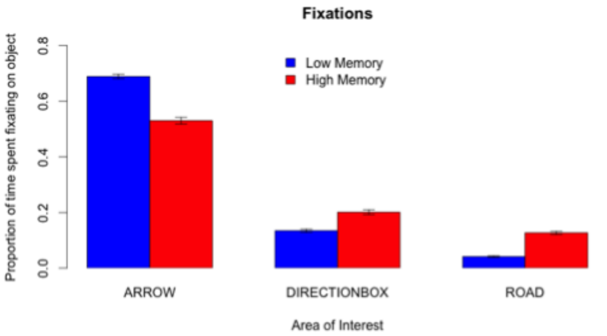

Figure 4. Time spent fixating on the arrow, direction box and road areas, respectively. High Memory subjects spent more time looking at the road and direction box.

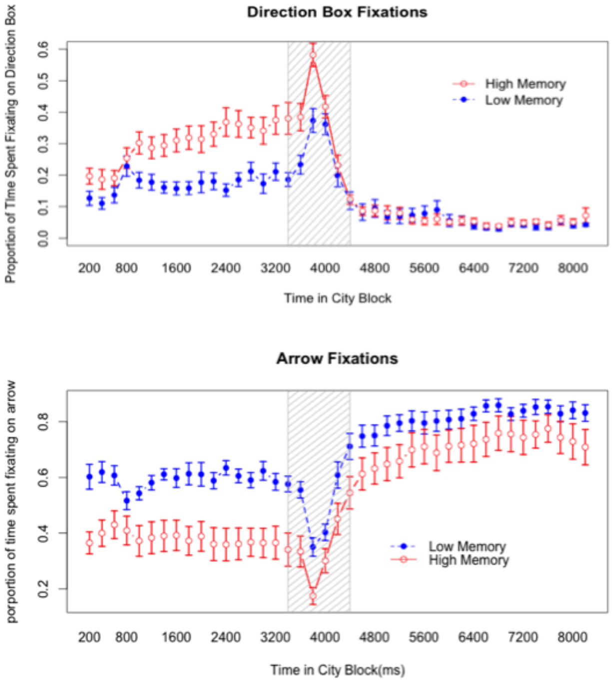

Figure 5. Proportion of time spent fixating on the direction box (top) and arrow (bottom) for middle instruction city blocks. The gray area represents the time period when the instruction is on-screen.

---

[1] Less than 4% of fixations for *high memory* and 3% of fixations for *low memory* subjects were to the *feedback* area with large variations between-Ss within the same group. A one-way ANOVA on that area was not significant.

more evenly distributed (i.e., more scattered) than those of low memory subjects *F(1,20)=8.31 p<.01*. But, when separated into the relevant areas of interest, individual ANOVAs showed that high memory subjects were more scattered in the *arrow* area *F(1,20) = 4.18 p<.05*, but not in the *direction box* or *road* areas.

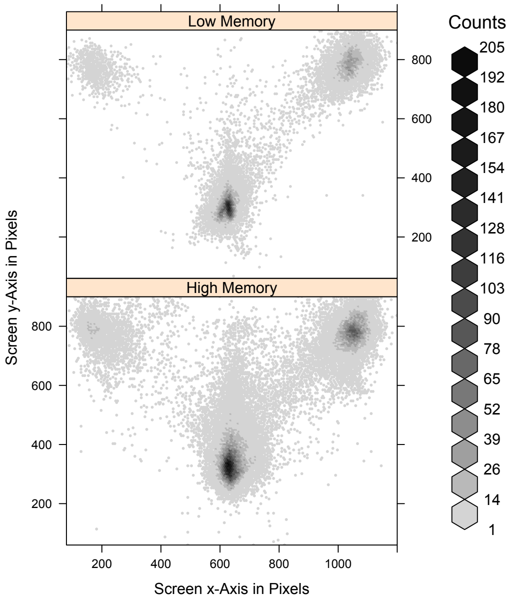

Figure 6. Density plot of eye fixations on screen for Low Memory (top) and High Memory (bottom) groups. Scale on the right shows the density count for the number of fixations centered at each pixel of the screen. Note that despite the same nominal task requirements for where-to-look-when the High Memory group scattered their fixations more across the entire screen than did the Low Memory group.

.

Discussion

As the jitter and memory tasks do not share any obvious resource conflicts, it is unlikely that resource based theories of multitasking (such as MRT) would predict the effects of memory load on jitter task performance. Likewise, as there are only two tasks being interleaved, it is also unlikely that Threaded Cognition would predict differences in how they are controlled. Yet our results show a significant difference in *inattention* to task relevant areas between conditions.

A general theory of cognitive control must explain the control required to switch among multiple tasks as well as the control required to perform each single task. For many tasks, a bridge between these two types of control may lie in understanding the role of eye movements. In our study, low memory subjects performed better on the jitter task than high memory subjects throughout the average city block. In addition, our microanalysis revealed that visual attention on the arrow is a very important factor in predicting jitter performance. Jitter performance in both conditions suffered in the time period before instructions were presented (as per Figure 2), as visual attention had to be split between the arrow and direction box.

Eye fixation data suggested major differences in the allocation of control caused by memory load. Low memory subjects were able to remain fixated on the arrow for much longer periods of time than high memory subjects. Fixations to the direction box increased until an instruction appeared for high memory subjects, but remained relatively constant for low memory subjects[2].

The demands on the proactive working memory system had an impact on the fixation strategy employed. The data suggest that high memory fixations reflect a different proactive strategy meant to focus on the turn instruction at the cost of a loss of focus on the arrow. Despite this difference in strategy, our variability measure showed that memory load also caused the high memory condition to lack focus on the arrow *while they were looking at it.* It can be argued that Increased memory load indirectly affected jitter performance by causing a lack of task focus. Task that require constant focus for optimal performance (like our jitter task) will be particularly susceptible to this effect.

In currently planned studies, we hypothesize that these results are not a special case, but rather a demonstration of how cognitive control affects the use of all the brain's resources. We can no longer assume, as Threaded Cognition does, that task performance is only subject to the control required to switch tasks. We must also consider the effects that overall control demands have on how we perform each task. Recent research points to differences in the use of proactive control in affecting behavior in many tasks, including the attentional blink (Taatgen, Juvina, Schipper, Borst, & Martens, 2009), AX-CPT (Braver et al., 2007), and n-back (Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011) tasks. Our future research will focus on the mode of control (e.g. proactive/reactive) to help us explain these complex multitask effects that do not fit into a strict resource (MRT) framework.

Acknowledgements

Research reported in this paper has been supported, in part, by grant N000141010019 to Wayne Gray from the Office of Naval Research, Dr. Ray Perez, Project Officer. Julia Van Cleve and Matt Nebel assisted in data collection.

---

[2] We remind the reader that the data presented in Figure 5 is for those trials on which Ss received their instruction in the middle of the city block. These data are representative of the early and late conditions in that those conditions also show increased fixations on the Direction Box until the instruction is presented for both Low and High load conditions with High > Low.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance iv* (pp. 421–452). Cambridge, MA: MIT Press.

Altmann, E. M., & Gray, W. D. (2008, July). An integrated model of cognitive control in task switching. *Psychological review*, *115*(3), 602–39. doi:`10.1037/0033-295X.115.3.602`

Anderson, J. R., Bothell, D., Byrne, M. D., Douglas, S., Lebiere, C., & Quin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: dual mechanisms of control. In C. Conway, M. J. K. Jarrold & a. N. T. A. Miyake (Eds.). *Variation in working memory* (pp. 76–106). Oxford University Press.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106 –113. doi:`10.1016/j.tics.2011.12.010`

Broadbent, D. E. (1952). Speaking and listening simultaneously. *Journal of Experimental Psychology*, *43*(4), 267–273. doi:`{10.1037/h0058014}`

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. doi:`10.1121/1.1907229`

Gevins, A., & Smith, M. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1), 113–131.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010, November). The relationship between n-back performance and matrix reasoning — implications for training and transfer. *Intelligence*, *38*(6), 625–635.

Just, M. A., Carpenter, P. A., & Hemphill, D. D. (1996a). Constraints on processing capacity: architectural or implementational? In D. Steier & T. M. Mitchell (Eds.), *Mind matters: a tribute to allen newell* (pp. 141–178.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996b). The capacity theory of comprehension: new frontiers of evidence and arguments. *Psychological Review*, *103*(4), 773–780.

Martin-Emerson, R., & Wickens, C. D. (1992). The vertical visual field and implications for the head-up display. In *Proceedings of the 36th annual meeting of the human factors and ergonomics society* (pp. 1408–1412). Santa Monica, CA: Human Factors and Ergonomics Society.

Martin-Emerson, R., & Wickens, C. D. (1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors*, *39*(4), 581–601.

McEvoy, L. K., Smith, M., & Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. *Cereb. Cortex*, *7*(8), 563–574.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex ”frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Ralph, J, Gray, W., & Schoelles, M. (2010). Squeezing the balloon: analyzing the unpredictable effects of cognitive workload. In *Proceedings of the human factors and ergonomics society 54th annual meeting*: (pp. 299–303).

Salvucci, D. D., & Taatgen, N. a. (2008, January). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological review*, *115*(1), 101–30.

Salvucci, D. D., & Taatgen, N. a. (2011, April). Toward a Unified View of Cognitive Control. *Topics in Cognitive Science*, *3*(2), 227–230.

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011, February). Control of interference during working memory updating. *Journal of experimental psychology. Human perception and performance*, *37*(1), 137–51.

Taatgen, N. a., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009, August). Too much control can hurt: a threaded cognition model of the attentional blink. *Cognitive psychology*, *59*(1), 1–29.

Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177.

Wickens, C., & Colcombe, A. (2007). Dual-Task Performance Consequences of Imperfect Alerting Associated With a Cockpit Display of Traffic Information. *Human Factors*, *49*(5), 839–850.

Wickens, C. D. (1992). *Engineering psychology and human performance* (Second). New York: HarperCollins.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. a. (2003, January). Attentional models of multitask pilot performance using advanced display technology. *Human factors*, *45*(3), 360–80.