DISSERTATION

SWITCH CHOICE IN APPLIED MULTI-TASK MANAGEMENT

Submitted by

Robert Gutzwiller

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2014

Doctoral Committee:

Advisor: Benjamin Clegg

Christopher Wickens
Kurt Kraiger
Stephen Hayne

ABSTRACT

SWITCH CHOICE IN APPLIED MULTI-TASK MANAGEMENT

Little to date is known concerning how operators make choices in environments where cognitive load is high and they are faced with multiple different tasks to choose from. This dissertation reviewed a large body of literature concerning basic research into choice in task switching, as well as what literature was available for applied task switching. From this literature and a prior model, a model of task switching choice that takes into account specific task attributes of difficulty, priority, interest and salience, was developed. In the first experiment, it was shown that task difficulty and priority influenced switching behavior. While task attributes were hypothesized to influence switching, a second major influence could be time on task. In the second experiment, it was shown that tasks indeed vary in their interruptability over time, and this is related in part to what task is competing for attention as well as the cognitive processing required for the ongoing task performance. In the final third experiment, a new methodology was developed to experimentally assess the role of diminishing rate of returns for performing a task. This declining rate was expected (and did result in) a general increase of switching away from an ongoing task over time. In conclusion, while task attributes and time on task play a major role in task switching in the current studies, defining the time period for theorized effects appears to be the next major step toward understanding switching choice behavior. Additionally, though the experiments are novel and certainly make a major contribution, to the extent that behavior is only represented in them, the methodology may miss some amount of ‘other’ task behavior, such as visual sampling.

TABLE OF CONTENTS

INTRODUCTION

"We choose to go to the moon. We choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one which we intend to win." - John F. Kennedy

JFK famously delivered his "moon speech" at Rice University to help unite a nation under a common task – to land men on the moon. As can be surmised from the quote, several factors marked this as worthwhile goal in comparison to other tasks that might have been competing for attention in the political climate of 1962. Going to the moon was chosen, according to JFK, because it was difficult, of urgent priority and of immense interest. In this case, difficulty was played down in light of the priority and interest of the task. These factors appear to sway how humans may make decisions to switch tasks away from others that may be ongoing. At the time of the speech, NASA was requesting 5.4 billion dollars, and the resources and effort could have easily been directed toward any number of the other major cultural and political issues, including racial tensions and the emergence of political troubles in Vietnam.

Similar to allocating resources from one task to landing on the moon, there is a distinct need to understand the conscious allocation of human attention. One can imagine a range of magnitudes of time for the switching to occur – in the case of going to the moon, this time may span months and years, perhaps decades. On the other end of the spectrum, switch decisions may need to be made rapidly in quick succession at the millisecond level.

As technology increases in utility and scope, demands for human monitoring are at an increased premium. In other words, attention is needed more often and in more places, over the course of any given real-world task performance. For example, with the addition of GPS systems in vehicles, drivers are not just looking at the road ahead, but also need to look at the display for the GPS, and are allocating attention to cell phones as well (Caird, Willness, Steel, & Scialfa, 2008; Drews & Strayer, 2008; Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013). Another prescient example can be found in the operation and supervision of unmanned drones. These operators are under immense attentional demands, in part the result of continuing technological changes in data presentation and increases in automation (Cook & Smallman, 2013; St. John & King, 2010). Humans are asked to do more and more attention-demanding work for reasons including technological demands and changes in work environments. Operators increasingly need to be able to engage in task management. Task management can be defined as the need to sequentially direct attention to performing several tasks, each of which may have separate demands and components (see Dismukes & Nowinski, 2007; Dismukes, 2010). Thus, task management is similar to concurrent multitasking (an attempt to perform multiple tasks in a short timeframe simultaneously) but is focused on the sequential allocations of attention to tasks and under some conditions represents attention acting as if limited by a single channel (e.g., Liao & Moray, 1993). Therefore, task switching is a more appropriate analogy for task management.

When operators must perform multiple tasks that require a switch of attention, they are naturally forced into tradeoff situations pitting allocating attention toward one task at the expense of removing attention from another. This frequently results in accidents, most notoriously revolving around prospective memory failures - the result of frequent interruptions and thus frequent task switches (e.g. Dismukes & Berman, 2010; Loukopoulos, Dismukes, & Barshi, 2009). And, these accidents occur even when operators possess large amounts of experience (Chou, Madhavan, & Funk, 1996; Funk, 1991). What aspects of tasks influence these multi-attribute decisions, and under what circumstances do they occur? What influences the weighting between factors, once these factors are understood, and what makes someone more or less likely to switch?

Chapter 1 of this dissertation begins by briefly covering what is known about task switching, what conditions make the applied realm an interesting case for limited parallel processing, and what the existing experimental literature has to offer about choice in task switching.

Chapter 2 covers the available literature on a relatively new paradigm, voluntary task switching, which enables study of task switching choice.

Chapter 3 details the applied study of task switching choice, surveying the additional studies on how operators choose tasks in more complex task-based, multi-task management scenarios. Prior literature is utilized to help build and update the Strategic Task Overload Model (STOM) developed by Wickens, Santamaria & Sebok (2013), which strives to predict operator switch choice behavior under load. Gaps between what is known or well supported, and what remains to be understood are discussed. Justifications and rationale for the upcoming experiments are then laid out.

Chapter 4 addresses the main platform of investigation – the Multi-Attribute Task Battery II, and how it was altered to fit the needs of task switching experimentation. Following, an experiment which evalutates manipulations of task difficulty and priority on switching behavior and choice is presented.

In **Chapter 5**, the extended literature and predictions of time-based task switching theories are presented and discussed. Then, several predictions are made and tested in a second experiment, which uses task interruptions to assay how switch tendencies change over time in an ongoing task.

In **Chapter 6**, an additional experiment is performed to look more closely at one aspect of time-on-task based switching explanations, that of diminishing returns.

Chapter 7 concludes the dissertation by summarizing the main findings and situating the projects in the scope of task switching, STOM modeling efforts, and application. Potential future issues for theory and practice are described.

CHAPTER 1: WHAT IS KNOWN ABOUT TASK SWITCHING CHOICE?

This dissertation begins by briefly covering what is known about task switching, what conditions make the applied realm an interesting case for limited parallel processing, and what the existing experimental literature has to offer about choice in task switching.

Task Switching

In a concurrent multitasking or timesharing situation, two or more task demands for attention must be met, but these demands overlap in time and in some cognitive sense (i.e., demand similar attentional resources, as in two visual tasks that occur at overlapping times). When two tasks compete for attention because they occur closely in time, the limitations of single channel processing may be invoked. Additionally changes in motor activities demanded by different tasks, such as movements of the hand to different response mechanisms or large eye movements toward separated displays also results in an attention more closely attuned with single channel processing (Liao & Moray, 1993; Salvucci & Taatgen, 2011).

Whatever task is first tends to have priority for attentional processing, while the other task is delayed in a ‘pipeline’, waiting to be serviced (see Fig 1.1 below). In an illustrative example, consider two tasks that may need to be performed quickly, and close in temporal proximity. Each task has a stimulus that elicits some response. For many experiments, the stimuli are simple – numbers and letters – and the tasks are also simple, perhaps determining whether the number is even or odd or the letter is a consonant or vowel. Thus, there are two stimuli presented in this conceptual example, one for Task 1 and one for an upcoming Task 2. Tasks in this context typically require a rapid, discrete response, rather than a continuous or time-intensive one. Initial perception of each of the stimuli is needed, and the information processing continues into further stages of response selection and decision-making, followed by the actual response processing time. The stimulus-onset asynchrony (SOA) represents the time interval between the presentation of the stimulus for Task 1 and the presentation of the stimulus for Task 2 in an experiment. In general, the shorter SOA is, the more 'cost' to reaction time (to the stimulus presentation) is observed in Task 2[*] (Monsell, 2003, Pashler, 1994; ; Salvucci & Taatgen, 2011; Welford, 1967; Wickens, Hollands, Banbury & Parasuraman, 2013).

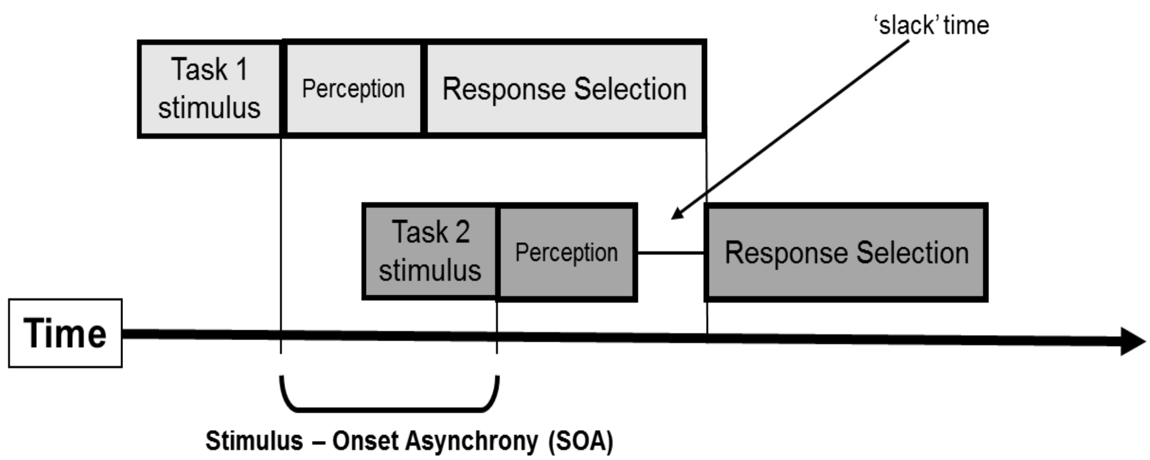


Figure 1.1. Representation of the psychological refractory period (PRP) as related to dual-task performance. The common finding is that as SOA decreases, more of a cost is incurred on the second task performance because stages in processing cannot be utilized by both tasks simultaneously.

Most of what is known about dual-task performance has been primarily concerned with examining interference between two tasks given overlaps in time (e.g., Pashler, 1994). These experiments provided evidence that interference between two tasks can be partially attributed to an attentional bottleneck but also to interference between modalities and task codes (Ruthruff, Hazeltine, & Remington, 2006), as well as a prominent bottleneck at the response stage as an

[*]Attention limit is from a theorized central bottleneck in processing. There could be limits imposed by dual demands on a given capacity (i.e., resources in multiple resource theory; Wickens, 2002) needed for both of the two tasks.

explanation for effects (De Jong, 1993; Pashler, 1994). A few examples are present that suggest concurrent multitasking can occur without cost, however, in all of them immense practice time must be undertaken, leading to this ability (e.g., Schumacher et al., 2001). Salvucci and Taatgen (2011) provide some evidence that concurrent multitasking can be done without a significant decrement in performance. In their conceptualization, tasks must be 'interleaved' at optimal natural breaks during which a task does not require attention for enough time that attention can be switched to a new task, and then back again, in a productive manner. Hence a task switch is involved, even in a more traditional and anecdotal conceptualization of multitasking.

Single Channel Theory in Application

The importance of invoking single channel theory is clear from an applied psychology perspective. Many attention allocation conditions in the real world are similar. Three environmental properties were outlined in Wickens and McCarley (2008) that set up a human performance scenario constrained by single channel theory. These are (a) times when task relevant visual displays are separated by considerable distance, making the bottleneck primarily perceptual and potentially inducing WM load; (b) when two or more tasks have to be processed rapidly in sequence due to their simultaneous or near simultaneous onset, such as a road hazard appearing at the same point as another event such as a conversational request or cell phone alert; and (c) when two or more tasks currently require a high degree of engagement or involvement, such as attempting to hold two conversations simultaneously. Liao and Moray (1993) suggest required motor movements, such as moving hands from one task response apparatus to another, may also invoke single channel processing.

These situations seem to adhere to the rules of single channel processing. Many of the applied scenarios where task management is required possess at least one of these characteristics (e.g., aviation scenarios wherein the pilot must deal with pre-flight checklists and may simultaneously be receiving instructions from air traffic control) and single channel framework has been applied to fault management, and to driving while using cell phones (Wickens & McCarley, 2008).

It should be mentioned that there are two potential exceptions: when there is either a clear benefit from automaticity, and/or, when there is a clear benefit from the different resource demands in regard to multiple resource theory (Wickens, 2002; Wickens, 2008).

These single channel applied domain properties help distinguish the targeted realm of study (task switching choice within task management) from related realms of concurrent multitasking and timesharing. Salvucci and Taatgen (2011) in fact speculate about a ‘range’ that spans from *sequential* to *concurrent* multitasking (also discussed in Wickens & McCarley, 2008). The range helps to split characteristics of task coordination and management, and sequential allocations of attention from dual- and simultaneous task performance. Sequential multitasking characterizes situations when task information may be presented at various times, for various tasks, and each task performed may be suspended and resumed because attention is required to be allocated between the tasks separately, instead of in parallel. Operators must retain goal states and a running queue of tasks ‘to do’, thus requiring significant memory maintenance of states of problems, and their relative priorities (Altmann & Trafton, 2002; Salvucci & Taatgen, 2011) to ease resuming tasks if they are left, and in order to deal with task interruptions. In comparison during *concurrent* multitasking, which is more akin to the colloquial representation, the operator handles tasks simultaneously because information is available for both tasks, and, the relative absence of resource conflicts may allow for it (because of automaticity, for example, or a lack of resource-limited issues; Norman & Bobrow, 1975).

One clear analogy to task management then is the study of task switching, which originated as early as Jersild (1927). Task switching is a concept typically married to a paradigm used in cognitive psychology to examine the mental mechanisms and processes surrounding a change from performing a given task (such as classifying a number as greater, or less than, five) to performing another given task (such as classifying a letter as a vowel or a consonant). In a standard task switching experiment, participants perform independent tasks that each requires attentional processing. As the examples suggest, this is fairly simple in that it may require only a simple identification and response, and the 'tasks' can be better described as rules (for task switching reviews see Kiesel et al., 2010; Monsell, 2003).

These tasks are generally first practiced in isolation. After single task practice, the two tasks are then executed in close sequential or temporal proximity to each other, sometimes in blocks of 'switch' or 'repeat' trials, other times in mixed orders within blocks (e.g., Strobach, Liepelt, Schubert, & Kiesel, 2012). Switch trials are usually defined as pairs of trials where the task performed on trial $n - 1$ differs from the task performed on trial n (n representing the current trial). The performance on a task following a switch is then compared to trials of repeated performance, or against a single task baseline depending on the purpose of the experiment. This can be examined in both speed of response (reaction time, RT) or accuracy of the response. The most common finding is that switch costs are present in reaction time. In other words, there is a "switch cost" to alternating or switching task rules from trial to trial, compared to performing a given task multiple times in a row on repetition trials (e.g., Monsell, 2003, Kiesel et al., 2010). A diagram of a generic task switching experiment is provided below in Figure 1.2.

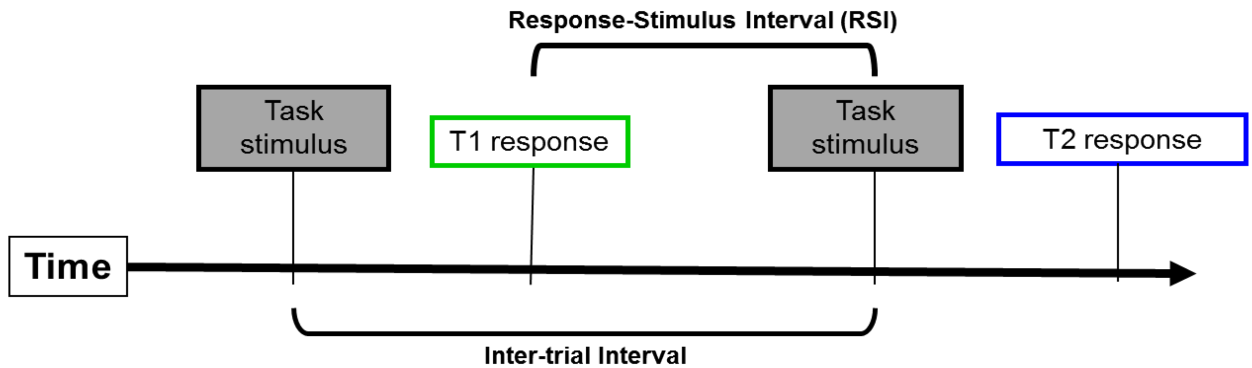


Figure 1.2. A simple task-switching experimental procedure. T1 and T2 represent the response outcome of two different tasks being performed. Represented here is a switch set; a repetition set would be two consecutive performances of task 1.

A related interval of time, the response-stimulus interval (RSI), is the time in between a Task 1 response and the subsequent presentation of the stimulus for Task 2. RSI controls the time an operator has to prepare for a switch, and is independent of any cuing. Longer RSI usually increases and shorter RSI decreases performance on task switching trials (Monsell, 2003). An RSI-increase benefit to task performance may also depend on the predictability of the task-switching paradigm. In other words, does the participant know which task is coming up next? Monsell (2003) observed that costs are evident not just at the switch event when the switched-to task suffers RT and accuracy declines. The cost can also be seen over the course of *time following the switch*, lending more credibility to a general view of memory as involved in task performance preparation. Eventually task performance reaches nearly the same point as when the task is repeated without any switch: but when upcoming task predictability is eliminated, task switch cost persists for several trials only gradually returning to near normal performance (Monsell, Sumner, & Waters, 2003). In contrast, using predictable task switching resulted in the cost to performance over time disappearing after a single trial of the switched-to task (Monsell et al., 2003).

Combined, the evidence considered from the basic task switching literature above suggests that increases in cue-stimulus interval, response-stimulus interval, and response-cue interval aid task switch performance in terms of reaction time, decreasing or reducing the switch cost, and benefiting repetition trials as well. The benefit to performance appears to be due to the memory benefits of cues, and the time for encoding and retrieval of task performance sets, such that better cues and longer time for successful retrieval enable better switch task performances. Nevertheless, the effectiveness may be limited to conditions in which an operator is exposed to multiple different interval times in the case of CSI (Altmann, 2005). Despite the benefits of increasing amounts of time it does not eliminate switch costs, and neither does extensive practice (Rogers & Monsell, 1995), and neither do cues even when cues are informative about the duration of the upcoming interrupting task, and its priority through peripheral cues (Hameed, Ferris, Jayaraman, & Sarter, 2009).

The applied relevance of these all of the summarized effects to task switching choice, other than the robust nature of the switching cost, is less clear. By focusing on reaction time, other questions are left unresolved. For example, what is the impetus of a switch? About half of switches that occur under free choice conditions originate from the operator who possesses a goal state (e.g., Adler & Benbunan-Fich, 2013) and clearly these differ in some respects from exogenous interruptions (McFarlane & Latorella, 2002).

What elements determine which task could be chosen and when a current ongoing task is abandoned? It would be specious to suggest that a single factor determines task switch choice; there are likely multiple complex traits that govern what task is performed following a switch, and these span properties of both the ongoing task (OT) being performed, and, the properties of the alternative tasks (AT) which could be selected and switched to at any point. Yet, none of the literature surveyed so far covered task choice, much less the factors influencing it.

Scaling Research From Basic To Applied

Determining how operators make task-switching choices in real-world situations is useful to understand. For example, understanding operator choices in switching may help change the design of platforms to make certain task properties exemplify one or more of the presupposed factors that influence switching, making some tasks more or less likely to be switched toward under various conditions. Prior research as a whole has constrained the applicability of task switching basic research to real world task management performance, where choice is an integral component. Clearly, in the search for applied task switching knowledge, several factors preclude the natural approach of pulling from the basic task switching findings.

For one, basic laboratory research typically uses tasks that are not commonly found in an applied context, and are potentially much less cognitively demanding and complex. Consider as one applied context, the technology-enabled driving experience. A clear contrast can be shown between an example of a basic “task” used in task switching experiments - the task of identifying whether a number is even or odd - compared to a task an operator might perform in a driving simulator experiment, such as hazard avoidance where an operator must engage in visual monitoring for obstructions, predict upcoming threats to safety, and potentially execute maneuvers to avoid them (Fisher, Pollatsek, & Pradhan, 2006; Gugerty, 1997; Horrey, Wickens, & Consalus, 2006). These two ‘tasks’ are certainly different enough to warrant speculation that results compiled using them may not be transferrable.

Second, the typical task-switching paradigm in cognitive psychology research is limited in terms of the opportunity for a participant to make a free choice to switch – and to which task to be switched – in addition to being limited in general to two tasks, versus the multiple tasks that may be available to an operator in the real world. The typical task-switching paradigm has stuck almost entirely to paradigms that *force* switching between tasks, cue switches and repetitions of tasks using colors or shapes, and implement task rules such as asking participants to memorize an ordering of tasks to be performed (e.g., AABB).

Finally, experimental manipulation of more than one facet of task characteristics may also reveal different patterns of task factor influence. For example, it could be that task characteristics occur in a hierarchical fashion – one factor if present will dominate the others in the hierarchy below it. Another possibility is that each attribute is independent; they may be additive, such that a task with priority <u>and</u> interest would rank higher in attractiveness than a task with only priority or only interest. Finally, a true interaction of these factors is possible. For example, perhaps the difficulty and the interest in a task interact; it is not that the task is attractive because of ease of performance, or because it is interesting, but rather because it is both.

In other words, the relative weights of each of the factors in a decision making context should be examined, and must be determined experimentally in order to properly predict what tasks may be switched to, and what other tasks are more likely to be ignored. The basic task-switching paradigm clearly does not allow this to be evaluated.

A literature search was conducted to see if choice had been addressed in any task switching domains. This search resulted in the discovery of a missing, but relevant area of work, which will be subsequently included in an updated form of the model. The next section summarizes this area as a review of a laboratory task-switching paradigm called Voluntary Task Switching (VTS).

CHAPTER 2: VOLUNTARY TASK SWITCHING

The goal of this chapter is to review the body of literature that can contribute to the development of a model of task switching choice. This means literature should additionally take into account the unconstrained, high workload, and switches that occur in the real world. There is a lengthy history of studying task-switching behavior in cognitive psychology. Some basic task switching literature has focused almost exclusively on reaction time measures of switching, switching costs, and examinations of executive functioning (Monsell, 2003; Kiesel et al., 2010).

The Paradigm

A recent paradigmatic development by Arrington and Logan (Arrington & Logan, 2004, 2005) has changed the direction of basic task switching research to emphasize and explore switch choice through the voluntary task-switching (VTS) paradigm. VTS has led to a wealth of new experiments that could be informative to discovering the factors that influence task switching choice.

Arrington and Logan created and developed a task-switching paradigm called Voluntary Task Switching (VTS) that has now been used in over 30 publications. The new paradigm attempts to examine task switch choice under a variety of experimental conditions. Unlike all prior task switching paradigms, wherein participants adhere to an experimenter-determined ordering of tasks (e.g., AABB) or a cued task structure (cued for task A, must perform task A, etc.), their paradigm allows participants to choose the task to perform from trial to trial. However, in almost all of the VTS studies reviewed here, two basic constraints were imposed on participants', as described below, which limit the generalizability of their findings to most applied switching situations.

Organization of VTS Literature Review

The questions that need to be asked are primarily related to how task switch choice is affected by specific task characteristics. In determining how to review the availed literature, a 'features of relevance' checklist was created and used to determine which, if any, studies provided information on how task switch choices are made.

Tasks or Rules. First, each experiment in each publication was determined to have used either *tasks*, or as was more common, different *rules* for one stimulus. For example, two tasks may be completing a Sudoku puzzle, and reading a paragraph of text; whereas switching that takes place between two rules may be switching between judging a single number stimulus to be even/odd, versus judging the number to be greater than/less than five. Only 14 of 44 studies used tasks instead of rules according to this definition.

Task Choice as a Dependent Variable. The experiments reviewed had to have a reported measure of task choice. This was typically represented in global measures about whether tasks were repeated, or switched across blocks of trials. Although some interesting data may result from examining reaction time or task accuracy when choice is allowed, it did not provide information about how *often* different choices were made, and what task was chosen as a result: possessing a choice metric checked to see if the literature could be informative in such a regard. Overall, the fact that most studies included some choice measures was promising in the context of understanding task switch choice.

Task Choice Freedom. Next, the issue of operator 'task choice' was examined for each experiment. Choice was recorded as either *free choice*, meaning that the participant could choose to switch to perform any task at any time with no penalty other than any induced in switching attention; or, the choice could be *constrained*, as was inherently the case with any study that used the Arrington and Logan (and any sub-set of) instructions. Choice under the Arrington and Logan (2004) instructions was allowed with two caveats, a ‘cost’ of sorts that is likely not present in the real world, and thus drives those conditions further away from the true relevant conditions. These instructions generally also ask participants to randomly switch between tasks (or, rules), and to keep their individual task (or rules) performances equivalent in number between the two. In other words, participants are asked to perform each task an equal amount of times over every block while keeping the order of these performances as random as possible. These *constrained* choice instructions were utilized in 33 of the 44 studies surveyed, suggesting their permeation in the paradigm, and leaving only eleven studies that examined task switch choice under *free choice* conditions.

Task Characteristics. Finally, the task properties were examined for any manipulation of the proposed characteristics that could influence task-switching choice, such as task difficulty, priority, salience and interest/engagement. This investigation applied to both “tasks” and “rules”. If these factors were not experimentally manipulated, but discussed or appeared as related to task choice, they were of interest and was noted along with what property was present, and the circumstances under which the study took place. In the review, 17 studies included at least a measure, or a manipulation, of one task characteristic as related to task switching choice.

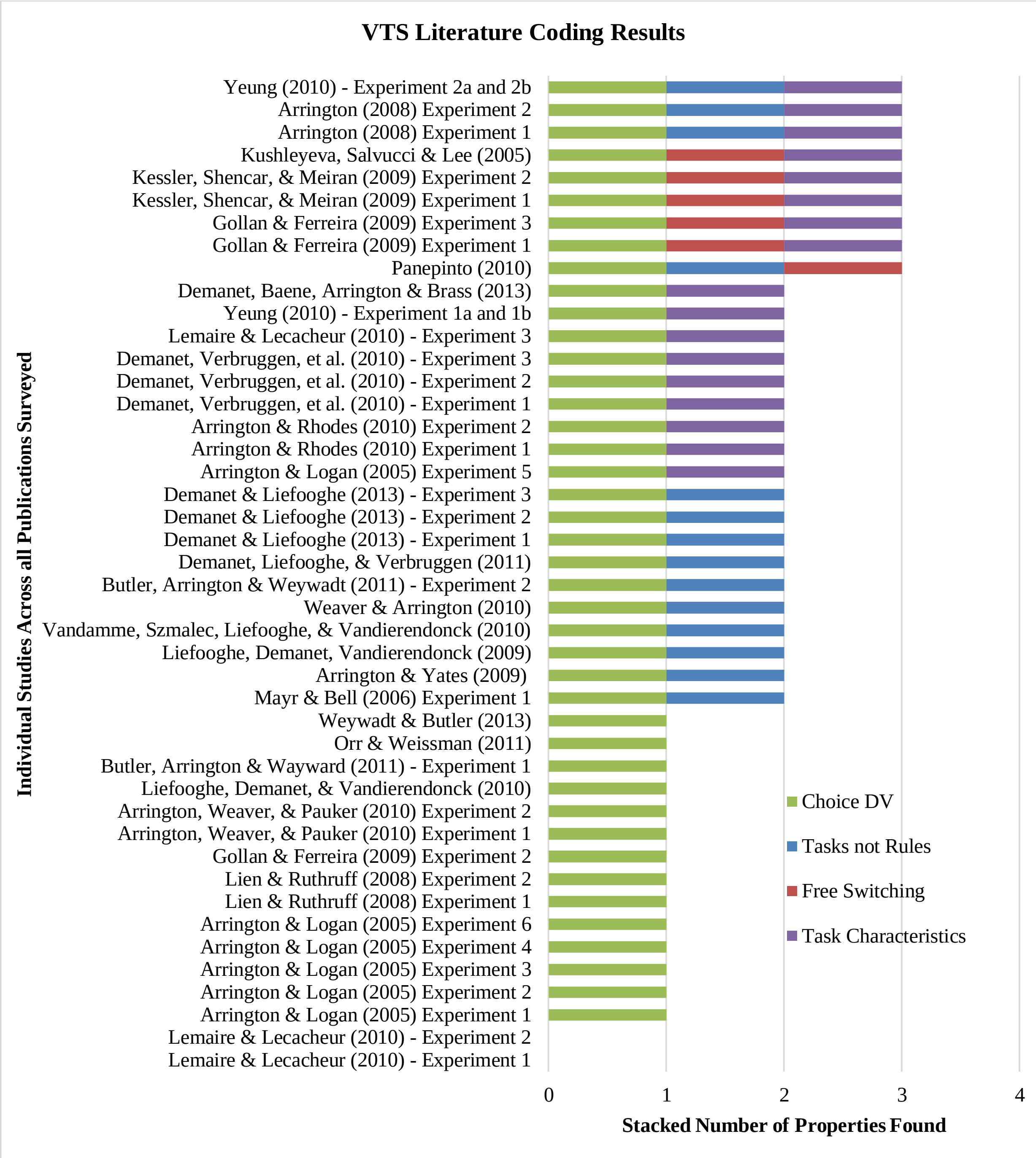


Figure 2.1. Individual studies as coded for each feature in the literature review, each bar indicating the presence of the factor in the study. The graph and underlying matrix helped to showcase the lack of VTS literature that addresses the question of truly free choice in task switching decisions (unconstrained by Arrington and Logan's instructions; red bars) between actual tasks (blue bars) instead of rules or classifications. While choice was a dependent variable (DV) measured in the majority of the articles (green bars), a large gap also exists in manipulating or describing task-characteristic factors that may influence task choice, such as task difficulty, interest, priority, salience, and other factors such as working memory load (purple bars).

Using the above general classification studies were organized for easy of interpretation (Figure 2.1). Each division established previously organizes the VTS literature, and special attention was given to studies that addressed multiple features. In each section, effort was made to highlight findings that contribute to a model and understanding of task switch choice, though this sometimes meant a finding was couched within a less than ideal *other* characteristic (such as using *rules*, or possessing *constrained choice*). Nonetheless, these effects are potentially valuable in painting a complete picture of what is known about task switching choice.

Specific Findings of the Review

Tasks or Rules. In many respects, the VTS literature parallels the basic task switching literature in that often tasks are very simplistic and may be better characterized as rules (e.g., asking participants to make parity judgments of numbers, consonant/vowel judgments for letters, and sometimes shape or size judgments). As a good example, six studies found in Arrington and Logan (2005) all use two rule-based 'tasks', in which upon seeing a stimulus (a single digit number) the participant must execute one of two tasks, either a parity judgment (e.g., number is even or odd) or a magnitude judgment (e.g., number is greater or less than 5). Throughout each experiment, while other variables were manipulated such as cueing type, these same tasks were used in each case and the stimuli were always numbers. Across these experiments, a trend emerged that participants tended to repeat, rather than switch between the 'tasks'. The effect, deemed the 'repetition bias' by the authors, can be found across many of the VTS experiments. Thus, despite the constrained nature of the choice (due to instructions), and the lack of true tasks (use of rules instead), *some* evidence is found that suggests participants view of switching is much like a cost or at least is not preferred, under a variety of experimental conditions.

Additional studies found similar results based on using different rules for a number stimulus (Butler, Arrington, & Weywadt, 2011, exp 1; Demanet, Verbruggen, Liefooghe, & Vandierendonck, 2010, exp 1 and 2; Weywadt & Butler, 2013), although some used noun judgment rules, such as whether a word represented something living, or whether an object was small/large (Arrington, Weaver, & Pauker, 2010; Demanet, Baene, Arrington, & Brass, 2013; Demanet, Verbruggen, Liefooghe, & Vandierendonck, 2010, exp. 3).

Other studies using *global versus local* letter identification tasks (Arrington & Rhodes, 2010), shape location or shape type judgments (Yeung, 2010), two visual search tasks (Kushleyeva, Salvucci, & Lee, 2005), two strategies for solving similar math problems (Lemaire & Lecacheur, 2010), two languages as a potential response (Gollan & Ferreira, 2009), or a Stroop task variant (Liefooghe, Demanet, & Vandierendonck, 2010).

Some studies used more than two rules, such as Lien and Ruthruff (2008) who asked participants to make a parity, size, or distance (near or far from 5, with near representing 3, 4, 6 and 7, and far representing 1, 2, 8 and 9) judgment on a number. Similarly, Kessler, Shencar and Meiran (2009) had participants make three different judgments on a shape stimulus in their first experiment, and four different judgments in experiment two, including size, fill state, shape and color. All suffer the same conceptual issue, however, in that the stimuli themselves do not change, and therefore the tasks being switched between are more easily classified as *rules* rather than true tasks. As previously discussed, such results are of low overall value for developing an understanding of task switch choice.

Nevertheless, 14 studies reviewed used tasks that were not rules for a single stimulus, but more often were separated in stimuli and rule or task (see Figure 2.1). Some of these studies simply used letters and numbers for two separate tasks to switch between, instead of just one type of stimulus (Arrington & Yates, 2009; Arrington, Weaver, & Pauker, 2010; Arrington, 2008; Butler, Arrington, & Weywadt, 2011, exp. 2; Vandamme, Szmalec, Liefooghe, & Vandierendonck, 2010; Weaver & Arrington, 2010). Other studies used letters and shapes (Demanet & Liefooghe, 2013; Liefooghe, Demanet, & Vandierendonck, 2009), or symbols and numbers (Demanet, Liefooghe, & Verbruggen, 2011). In Demanet, Liefooghe and Verbruggen (2011) for example, participants were shown either symbols or numbers in green or purple, categorizing either one.

A few studies went beyond two different tasks; for example, Demanet and Liefooghe (2013) examined participant responses to four distinct stimuli in experiment 3, asking participants to classify a letter, number, shape, and color. One study also used unique *true* tasks – participants switched between solving a Sudoku puzzle, and editing a word document (Panepinto, 2010), an important find as the attempt to scale the literature to any applied domain is made.

Although understanding the makeup of each study is important in determining how it scales up to applied research, additional steps were taken to extract general relationships from all of the studies. From this and several additional applied papers reviewed later, it was found that there is a general tendency to "stay" with a task (in other words, repeat it) of 60% (95% CI; 58-62%). Thus, the switch aversion (or repetition bias) appears to be a reliable finding in this literature.

Task Choice Freedom. Studies that incorporated free choice were few. Several of the studies that did, also used constrained choice as a comparison condition, however none of them were able to show differences between such groups in task switching free choice. Only six studies reviewed allowed participants to make an actual *free,* unconstrained switch choice:

Gollan & Ferreira, (2009, exp. 1 and 3); Kessler, Shencar, & Meiran (2009, exp. 1 and 2); Kushleyeva, Salvucci, & Lee (2005); and Panepinto (2010). All others surveyed used the instructions for participants outlined in Arrington and Logan (2004), and although sometimes these instructions were modified slightly (Liefooghe et al., 2010), choice was still severely constrained, which limited the scaling of findings to the real world.

Gollan and Ferreira (2009) were interested in the voluntary task switch choice occurring when bilinguals use different languages to respond to pictures with the name of the picture contents. They examined a 'mixed' block condition where participants were asked to say whichever word came to mind quickly. English-dominant participants switched less often than the balanced-bilingual group (24% vs. 35%), suggesting a naturally higher rate of switching when two languages are known approximately equivalently. Additionally, English-dominant participants were less likely to "stay" in Spanish (by giving repeated Spanish responses). In another experiment participants included older and younger bilinguals, and were allowed to switch or repeat language in picture naming as before. Older bilinguals switched marginally more often than younger; but younger strong Spanish speakers switched more often than older speakers. The task of naming picture items is more of a rule – use either English, or Spanish, to name the object - than a true task. Nevertheless, it could be suggested that speaking an unfamiliar language is more difficult, and would be avoided. These results model such a hypothesis well, and implicate age-related effects of switch cost avoidance, even when a language may be well learned.

In Kessler et al (2009) experiment 1, participants judged shape stimuli on three dimensions (shape, color, and size) in one of two groups. Upon close inspection, it appears that the variance far exceeded the means in both groups; accordingly, it cannot be concluded that switch freedom is unimportant from this study.

Kessler et al. (2009), exp. 2 and Kushleyeva et al. (2005) are both summarized under the difficulty feature of relevance section later, but it should be noted that they represent a few of the rare times in which a feature of a task is examined under free switch conditions. One of the other studies that fit into this category is Panepinto (2010), in which participants in one of three conditions switched back and forth between a Sudoku style puzzle, and a word document-editing task. No differences between groups were apparent in Sudoku performance or first-switch time to complete an action. In this case, no differences were shown related to this condition and not enough data were reported to determine whether the frequency or direction of switches had a clear impact based on group.

One VTS study (Liefooghe et al., 2010) examined more closely the limitations of the Arrington and Logan (2004) instructions that were found to be so common in the VTS paradigm. In this experiment, participants completed a Stroop task variant, in which words had to be either read, or the name of the color the word was printed in was named. Participants completed one of two conditions; in one, the typical VTS instructions were given. In the other condition, the Arrington and Logan (2004) instructions were specified, but importantly no instruction to produce random sequences of task performances were given. The logic of this was that this component of the instructions might be tied closely to executive functioning. In the standard condition, a repetition bias was shown (.59), and more pronounced in the modified version (.69). Additionally repetition bias differed by task – the proportion of repetitions was higher in the color-naming task. This can be interpreted as supporting the task difficulty relationship: color naming is probably more effortful in a Stroop task. Additionally participants tended to repeat color-naming more often than word naming.

Thus in the few studies available it would appear there is at least some difference between constrained and free switching results – potentially as a function of the Arrington/Logan instructions requiring additional executive control by specifying random performance of tasks. Nevertheless, repetition biases are still evident even in highly constrained switching, suggesting some continuation of a ‘stay’ preference. In other words, participants preferred to stay with the easier task. Not many studies have examined free choice for task switching. In the following section, those studies that have done so are summarized to the extent that certain task characteristics influence the task attractiveness of a voluntary task switch.

Task Characteristics

Difficulty. Few studies considered the role of task difficulty explicitly. For example, in Lemaire and Lecacheur (2010) participants switched between two strategies for performing mental math problems. In their third experiment a significant switch avoidance effect was shown (.41), and, participants were less likely to repeat a strategy when the target problem required carrying.

In Kessler et al (2009), exp. 2, participants judged shape stimuli on four different properties. Judgment was deemed easy or difficult for the tasks, and in general participants more switches compared to staying within easy or hard task execution. There was also a slight bias to perform the easier tasks (53%) overall, compared to the difficult.

Kushleyeva et al. (2005) utilized a pair of tasks, where participants had to alternate between two visual search tasks (looking for the presence or absence of a target letter in strings of letters). One of the search tasks was always twenty characters in length, and thus the most difficult, but the other task was manipulated to be either 5 or 9 characters, thus relatively easier and harder. The participants received limited time for each task to be performed, and were penalized if a task was left undone. It was anticipated that because points were awarded for correct answers, participants would spend more time in the easier searches. Participants were found to switch tasks more often after completing a 9-letter, than a 5-letter task (.57). Interestingly, they were also more likely to switch than stay if the next task presented was a 9-letter search instead of a 5 (.54). A clear effort avoidance effect was exhibited here.

Finally, across two sets of experiments, Yeung (2010) asked participants to switch between responding to shapes and locations. In experiment 1, participants responded to spatial locations with a spatially corresponding key, which was suggested to be easier than responding to shape judgments. Choices were restricted in task switch freedom by modified Arrington and Logan (2004) instructions, in which participants were either told to use time before stimulus presentation to prepare for the upcoming task in a cued version, or to decide what task to perform (VTS condition). The VTS condition showed the repetition bias (switch avoidance) effect, being significantly more likely to repeat a task, and this was especially so at short RSIs (.63 overall). There was also, importantly, a significant bias to perform the shape identification task more often than location (.53 versus .48). In experiment 2, participants switched between two location tasks, one that was compatibly, and one that was incompatibly mapped in one condition (VTS condition), and the repetition bias disappeared, although there was a small and consistent bias for performing the more difficult task (incompatible mappings). Shape repetition increased the likelihood to repeat a task, even after controlling for repeated locations. Yeung (2010) paints a picture of task difficulty that appears to somewhat contradict that shown so far (and what will be discussed later), in that participants seem to be slightly, though significantly, biased *toward* the more difficult task.

Overall across the experiments located, and including the results of some further applied studies reviewed later, a measure of preference for switching to an easy task (63%; Wickens, Gutzwiller & Santamaria, under review). However the reasons for this are rarely given. Interestingly, it could be that preference for easy tasks may be reflecting the rate of marginal return received by performing the task (Duggan, Johnson, & Sørli, 2013). Other studies show some evidence that delays in tasks also encourage switching (Katidioti & Taatgen, 2013), which again may be tied to a reduction in rate of return. The rate of return is revisited theoretically (Chapter 5) and then experimentally (Chapter 6) in this dissertation, slightly altered to distinguish diminishing (over time) returns of a task.

Role Of Working Memory And Executive Functioning In VTS. One a priori assumption concerning switching is that it requires mental resources and may also require some amount of working memory and executive control. To the extent this is the case, individual differences in either could have various effects on switch frequencies; it may be expected that lower capacity would result in fewer switches, because the underlying process is resource limited. For similar reasons, it may be expected that under high WM load, that may switch less frequently. A secondary effect may be that tasks with high WM demands may be avoided, again because of the demands of switching – and in line with the general attractiveness of easy tasks as reported above.

Regarding an individual differences account, Butler, Arrington and Wayward (2011), exp. 1, had participants task switch between a parity and a magnitude judgment on a number, and found that working memory capacity was unrelated to switch probability. One other study

(Arrington & Yates, 2009) showed a significant negative relationship between executive control measures of the Attentional Network Test (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005) and task switching probabilities. The ANT task executive control measure correlated with switch probabilities at low and high RSIs (r = -.27, and -.37, respectively). In other words, with more executive control, as measured by the ANT, there was less likelihood to switch tasks. Increasing the time between response and stimuli presentation appeared to strengthen this relationship. These results suggested that some individual differences in attention captured by ANT are related to VTS paradigms; however, this study was subject to the constrained task switch choice. A question thus emerges about the role of executive control in free task switching choice conditions – this is returned to and tested in Chapter 5 of the dissertation.

In addition to individual differences, one effect of WM on switch behaviors may be whether an operator is currently under significant WM load. Demanet et al. (2010) presented two experiments in which a WM load increased the repetition bias across two different types of voluntary task switch paradigms from .51 to .58. In Weywadt and Butler (2013), participants completing voluntary task switches were less likely to switch under articulatory suppression compared to a foot-tapping condition. A take-home message from these experiments is that taxing WM resources increases the likelihood of ‘staying’ with a task. However, the categorization scheme for the studies showed that all of them used restricted choice instructions, and the tasks being switched between were characterized as ‘rules’ than true tasks.

Priority. No VTS studies considered task priority or manipulated it. Implicit in studies that examined more applied tasks was the idea that each task should be prioritized equally, so there is some recognition priority has been controlled in these cases. This control is exhibited by operators basing any choices between different tasks “as if” the priorities are equal, similar to an idea discussed in task scheduling in Raby and Wickens (1994). Thus, priority represents an area of much-needed measurement and experimental investigation based on this review. Priority has been previously discussed and examined in the skill acquisition literature and the interruption management literature. Markedly, Gopher and colleagues (Gopher, Weil, & Siegel, 1989) revealed priority instructions were influential in training to play a game called Space Fortress. There, it was shown that priority manipulations via simple instructions altered the way that participants allocated attention, and that these changes led to effective training especially in dynamic task performance that required timesharing. Further, in the interruption management literature, Iani and Wickens (2007) underscore the importance of an interrupting task's importance. An interrupting task related to a weather pattern, affecting navigation, exerted a negative influence on ongoing task performance (error in a designated flight path) when a tunnel-based display was used. Their results suggest that some tasks may be deemed a high priority or of high importance (these values can also be assigned), and as a result can interrupt ongoing performance. Similarly, certain ongoing tasks have high priority: in the driving realm, hazard avoidance is an ongoing task that appears to be 'protected' even in the face of other interrupting tasks (Horrey & Wickens, 2004).

The findings summarized here suggest task priority is important in two ways, despite its lack of experimental manipulation within VTS: (1) it may bias attention toward a given task and could be related to or serve as a proxy for task importance, and (2) priority may influence how participants learn to perform in a task management context – most likely increasing their performance and ability to learn timesharing strategies between multiple aspects of a task.

Interest. Interest in the task to be performed was only assessed in one experiment. Interest broadly encompasses the emerging conceptualization of engagement (Wickens, Hollands, Banbury & Parasuraman, 2013; Montgomery, Sharafi, & Hedman, 2004). In studies of interruption management, for example, task interest could be exogenous or 'optimal' interrupting task importance (such as a weather alert, or a hazard in the roadway being deemed interesting because of its relationship to the task; Iani & Wickens, 2007; Horrey & Wickens, 2004) or simply be innate interest found in engaging in certain tasks. If serving as a proxy for task engagement, high interest may even be related to cognitive tunneling in performance (e.g., Wickens, 2005).

Within the VTS review however, interest was not specifically discussed, but could be inferred from the discussion of mind wandering and task engagement, found in Demanet et al (2013) as an example of subjective interest. Participants performed animacy or size judgments on words, and conditions compared voluntary choice to cued task switching. Trials were either compatible, or incompatible, with prior task performance exposure to the given stimuli during training (similar to Arrington, Weaver & Pauker, 2010). Participants were more likely to repeat tasks on compatible trials (.58), and this chance increased from .54 to .63 during the trials before an episode of mind wandering was measured by subjective report. It was suggested, then, that mind wandering could be framed as loss of interest or engagement in the ongoing task, resulting in a switch to the alternative 'mind wandering' activity. This results in more task repetitions within the main task, and agrees with anecdotal sense that lack of interest/engagement in the ongoing task may spur a task switch to something different. However, it must be kept in mind that Demanet et al. (2013) used rule-based tasks, and did not allow free unconstrained switching.

In summary, interest is associated with task switching, perhaps most saliently in cases of attentional "tunneling", it may also be inferred that interests in tasks may increase their attractiveness. Though only one VTS study incorporated mind wandering, a presumable loss of interest, the results suggested mind wandering increased repetition or the ‘stay’ preference. Tying this idea back to the need for executive functioning, it may be that mind wandering is a *failure* of the executive functioning needed to switch tasks (McVay & Kane, 2010).

Salience. Task salience serves as a common experimental variable in many studies of interruption management. A general finding is that the role of task salience as an influence on switch behavior may depend on modality of presentation (e.g., Lu et al., 2013; Sarter, 2013). For example auditory alerts work well when the competing tasks are visually dominated, or when visual alerting is potentially difficult to detect (Iani & Wickens, 2007). Further, visual and auditory signals may be assumed to be more salient than memories to switch to a task based only on prospective memory (Dismukes, 2010), and sometimes garner faster responses (Lu et al., 2013). Operators occasionally ignore switching back to tasks that require prospective memory remembrance, which also represents a decay of a goal state (Altmann & Trafton, 2002; Dismukes & Nowinski, 2007).

Arrival time as used in an altered VTS paradigm may be a proxy for task salience effects. In general, earlier arrival corresponds to a task being more likely to be chosen unless a long period of waiting is experienced before the competing tasks are presented (Arrington, 2008; Mayr & Bell, 2006). Though a bottom-up feature like task arrival may influence switch, if more time is given between response and the next stimulus presentation the effect diminishes. This understanding is consistent with the ability of participants to protect some driving tasks from interfering activities (Horrey & Wickens, 2004).

Summary

Despite the rise of research using VTS, it suffers from severe problems as a proxy for answering applied questions about multi-task switch choice. VTS experiments mainly examined *rule* switching; this is distinct from switching between real *tasks* (for example, between a compensatory tracking task and a monitoring task), as elements such as the stimuli, task locations and relevant information rarely change in the former but tend to be dynamic in the latter. Second, the choice that participants in VTS studies make has been almost universally restricted by experimental instructions developed in Arrington and Logan (2004). These instructions are at odds with an effectual scaling to the real world where operators are routinely given free choice between tasks. These gaps are not filled by most studies in interruption management, because although the tasks there are true tasks, methodologies do not always allow for free choice between multiple tasks (instead of just ongoing or interrupting).

Finally, critical gaps of knowledge exist in determining the influence of many of the factors used to make task-switching decisions, even within the VTS literature. The lack of empirical work, especially taking into account characteristics that influence this choice, is apparent from the VTS review. For example, while some work has addressed the issue of task priority in determining task-switching behavior, no task switching study has addressed the tradeoffs between priority and task difficulty, or other pairs of factors, so little is known about how task switching choice operates in situations which include tasks that differ on *both* factors. Attempting to bridge these gaps is vital to informing a model or theory of task switch choice in applied domains, and can help address some of the applied issues (such as why some tasks are chosen over others that might be more optimal).

CHAPTER 3: DEVELOPING AND TESTING MODELS OF SWITCHING CHOICE

Allocation of attention is an important component in many applied task-switching domains, where operators are performing task management more than parallel processing. This task management involves the need for operators to weigh options and make decisions about what tasks to switch to, if they switch away from an ongoing task. An expanded review found that a new paradigm, voluntary task switching, allowed for some measurement of task switching choice. However this choice was in general constrained. Many studies limited the use of true tasks, and only explored or manipulated potential task characteristics cursorily. Therefore, while some information regarding task switch choice was unearthed, much remains to be done to address major gaps.

A Model of Task Switching Choice

Recently, drawing from an ongoing literature review of applied task switching and interruption management, Wickens et al (2013) developed a preliminary model of applied task switching choice under load that incorporated several properties shown in the literature to influence switch choice. These factors were developed as reviewed above, from studies in applied task switching and switch ratios were calculated, so that one could suggest, for example, that given an ongoing task, a person would be more likely overall to stay with that task, than to switch to any alternative task (a preference ratio of .60 according to Wickens et al., 2013, a value consistent with the repetition bias findings in the VTS literature (see Ch. 2; Wickens, Gutzwiller et al., under review).

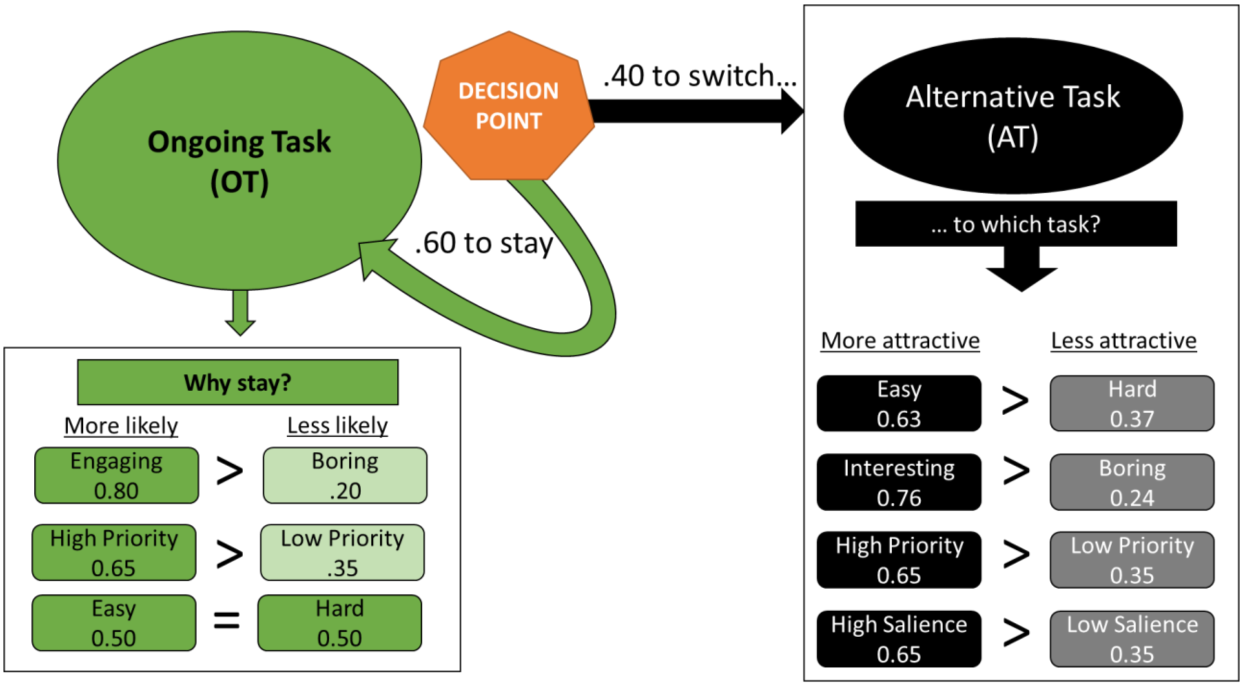


Figure 3.1. Choice probabilities represented in the Strategic Task Overload Model (STOM) developed in Wickens et al (2013) shown as updated by the VTS review (Wickens, Gutzwiller, & Santamaria, under review). Green represents the ongoing task, with its likelihood to continue, "stay" preference of 0.60, and the associated task attributes that influence the strength of the decision. Black represents the alternative task, a 0.40 chance to switch away from the OT, and its associated attributes that make the task more, or less attractive.

It can be seen that one of the main findings was a propensity to stay with an ongoing task (Figure 3.1), versus switch to an alternative task. Overall, the chance of switching away from an OT was determined to be an odds ratio of .40:.60 that any given moment, the odds of switching away from an OT to an AT is close to 1:2. Several factors should influence whether an operator 'stays' with an ongoing task, including engagement (a factor seen in cognitive tunneling), priority, and difficulty. Task salience does not apply for the ongoing task.

However, when a switch is chosen, several characteristics of a task (in bold, below the alternative task in Figure 3.1) were determined to influence the chance of a given task being switched to. Once having switched away from an OT, for example, which AT is chosen? Four general factors, as covered in Ch. 2, are revisited here.

Task Difficulty. Extrapolating from the VTS review, task difficulty effects are shown, with operators more likely to switch to easy tasks.

Payne, Duggan, and Neth (2007) showed participants avoid difficult tasks in general or at least gravitate toward tasks that are easier to perform and provide a higher rate of return on attentional investment. In the first two experiments, participants switched between two ongoing Scrabble tasks given different sets of letters (one was easier than the other to generate words) and asked to create as many words as possible within a time limit. Participants spent more time (60% of overall time) in the easier task than in the difficult, and the number of switches was low (Exp. 1). The same basic effect was shown using a 'medium' difficult task in place of the easy (Exp. 2). In the third experiment, participants attempted to work on two separate word search puzzles. Groups either completed the search puzzles, which varied in difficulty, in consecutive order, or could switch back and forth between them. Both groups spent the most time in the easy puzzle; this difference was greater for an easy-hard task combination, and switches were in general low.

In an observational study, Jin and Dabbish (2009) observed information technology workers using computer monitoring software, and by asking workers to note any time they had a reason for switching tasks. These explanations were broken into seven categories; including reasons related to tweaking the ongoing task (e.g., adjustment, inquiry) and switching to new things (e.g., break). It was shown that workers switched more often to rest (break), than switching to do a remembered task or perform inquiry about the ongoing task, two points that suggest effort avoidance at about a 2:1 ratio here.

Kool et al. (2010) also showed over six experiments that participants in general exhibit avoidance of effort when presented with low or high demand tasks, having to choose between a series of tasks, which required more switching and one that did not. In experiment 2, participants were asked to solve easy or hard two-digit subtraction problems. Requiring or not requiring a carry when solving the problem determined task difficulty; participants avoided the difficult problems in favor of the easy ones, selecting low demand 73% of the time.

In a memory study examining the learning of Spanish-English word pairs, Metcalfe (2002) used a unique design for word acquisition. In this phase, participants were shown three English words and had to click on each of them to get the Spanish translation. Words were ordered left to right based on increasing difficulty. Participants had 5 seconds to decide which word to select to see its translation, and completed blocks of learning, each followed by a test in which the English word was presented and the Spanish word was asked for. Overall, participants preferred to allocate learning time to easy, versus difficult levels at about a 2:1 ratio. In part, this was the result of poor initial learning of these items; in fact the trend was to increase time to the difficult trials. A further point seems to be that participants are capable when given feedback, to change allocation of attention based on updating subjective difficulty and goals. Overall the preference was for easier vocabulary, but once mastered they moved on to harder words.

Task interest. Only one study was found that examined this factor in a task-switching domain. Interest was a factor addressed in a few interruption management studies although these do not always look at performance in the context of multiple tasks switching choice. Spink et al. (2006) had participants search through three different web sites for information on a variety of problems, and participants were free to search between sites however they wanted. After searching, participants rated the attributes associated with their search technique, including the ease of information access, their interest in the information-seeking problem, and their prior knowledge of the search topic. High personal interest was the major factor in information search order, with 45% of participants selecting the most interesting search problem first. Prior knowledge was also rated as a major factor in information problem ordering for 25% of the participants; and 20% of participants said that the ease of finding the problem information was a major factor. While this study addressed in part the issue of interest, it did not attempt to manipulate it; and it was clear that while interest accounted for a large percentage, participants still made decisions using other information such as task difficulty.

Task Priority. Three studies were discovered that examined the role of task priority on task switching choice; one was an unpublished study in the lab as reported by Wickens et al. (2013). The others (Janssen & Brumby, 2010; Raby & Wickens, 1994) suggested the important role of priority in determining what task will be done and switched to. In Janssen and Brumby (2010), participants completed driving and dialing trials in a simulator. Within subjects, focus of attention was instructed to be either on the dialing task, or on lane keeping in the driving task. Importantly, the numbers that were dialed were in larger chunks than the standard US telephone (5+6 digits, versus 3+4). Participants showed high preference for the prioritized task, especially for steering during chunks of dialing responses, resulting in an estimated 2.2**:**1.2 ratio, or about 67% likelihood, of doing the high priority task.

In Raby and Wickens (1994), after completing a set of simulated landing attempts under three different workload levels, pilots rated aviation tasks. Participants seemed to alter their task completion strategy when under higher levels of workload, generally doing those tasks that were rated as of highest priority, and under high workload, shedding those rated lower priority.

The available literature on the influence of task priority in applied contexts, especially when tradeoffs are present, is minimal. Combined with the lack of studies containing this factor from the review of VTS, this area deserves attention in any forthcoming experiment, because we simply do not know how priority and difficulty may trade off as, for example, in choosing between high priority difficult tasks versus a low priority easy task. Which combination dominates?

Task Salience. Salience effects for the model were largely taken from the interruption management literature (see Wickens et al., 2013). Two main generalizations were assumed: (1) some reminders are more salient than others (auditory over visual, for example, Sarter (2013), and (2) any reminder is a helpful input when considering the prospective memory limitations of task management (Wickens et al., 2013). In the upcoming experiments, it is assumed generally that tasks involving auditory stimuli are more salient than visual, but that visual elements are more salient than prospective memory in terms of an event being performed.

Summary

In summary of these initial findings and model, there is a clear lack of empirical findings that allow any specifications or building of a foundational model of task switching choice. However, those that do exist suggest operators should show an effort avoidance pattern, choosing easy over difficult tasks. Switch avoidance may be high especially under high workload when resources are scare, though this may be mitigated in part by individual differences in executive functioning. Further, once a switch is initiated, an interesting or engaging task will be more attractive than a boring one, a high priority task will be more likely to be performed than a low priority task, and tasks with high salience are more likely to be chosen compared to those with low salience. Thus a multitude of factors are at play in environments where operators are switching tasks in complex real-world situations that have high cognitive demands. The experiments to support weighting or unifying these factors appear to be absent from the literature.

The following experiments endeavored to address task-switching choice in relevant ways, in order to contribute to continued model and theory-building efforts (e.g., Wickens et al., under review; 2013; Gutzwiller, Wickens, & Clegg, 2014). The outcome of continuing to inform both of these efforts is clear: an expanded understanding of a critical aspect of human performance that is likely only to increase in importance with the increase of technology that continually demands switching of attention.

Significant gaps of knowledge are addressed in the three experiments undertaken below. In Exp. 1, an important question is asked about how task factors, such as task priority and task difficulty are weighted relative to each other, and, interact to influence decisions and switching when manipulated under the high workload conditions of the Multi-Attribute Task Battery. Additionally, how do other task characteristics play a role, especially in circumstances wherein two tasks are directly competing for attention through simultaneous events? In Exp. 2, I examine what effect, if any, time duration of an ongoing task has on the likelihood to switch (a potential time-on-task effect). The rationale and literature for Exp. 2 proposes multiple hypothetical outcomes, and is covered in depth at the beginning of Ch. 5. In Exp. 3, I investigate another potential influence for time-on-task effects, specifically; the role of a diminishing rate of return on performing a specific task in MATB may have on task switching choice. Prior work only supports the role of rate of return, but many tasks diminish in this rate and are not static. Following the experiments, a general discussion of applied multi-task switching choice is undertaken, in which combined findings are tied into the existing literature.

CHAPTER 4: EXPERIMENT 1[†]

Based on reviews of the role of choice in task switching literature (Ch. 2 and Ch. 3) experiment 1 was devised to address two factors that were described in the STOM model (Ch. 3): task difficulty and task priority. Each of these factors was expected to exert an appreciable influence on task switching choice by increasing the likelihood of being chosen by a switch, and decreasing the likelihood that a task with priority would be abandoned. However, due to the restricted nature of the basic work they derive from, the influence of these factors does not have much applied support on their own, and even less when they are examined under joint manipulation.

Experimental Rationale

Task difficulty has been shown to influence the target of a switch, in that an easy task is chosen more often than a more difficult task (e.g., Wickens, Gutzwiller, & Santamaria, under review; see also Payne et al. 2007, Kool et al., 2010; Jin & Dabbish, 2009). Task priority, as shown in Raby and Wickens (1994) and in Janssen and Brumby (2010), also influences task switch choice because operators are in general motivated to perform high, versus low, priority tasks. One of the things we do not know is how these factors interact and weigh against each other, when paired against other model factors like task interest and saliency. A multitasking simulation, the *Multi-Attribute Task Battery* (MATB) was used to examine performance under multi-task situations (though importantly constrained, as outlined in the methods).

[†] *Experiment 1 was completed as part of work funded by NASA and an initial write-up was accepted for publication and presentation in the 2014 Human Factors and Ergonomics Society meeting (Gutzwiller, Wickens, & Clegg, 2014).*

As the MATB had not been used previously as a task-switching paradigm, one goal was to determine its viability and sensitivity as a platform. In this experiment, it was determined that features of relevance could be manipulated (difficulty), and that specific tasks could be pitted against each other for operators to choose from under a multi-task situation. This was a unique contribution, since choice of which of <u>multiple</u> alternative tasks is relatively underdeveloped, and previous work tends to focus only on pairs of tasks (Bogunovich & Salvucci, 2008; Brumby, Salvucci, & Howes, 2007; Janssen, Brumby, Dowell, Chater, & Howes, 2011; Janssen & Brumby, 2010).

Below, the MATB environment is described. The platform was also used for Exp. 2 and Exp. 3.

MATB II Overview

MATB II is a multitasking research tool designed to assess operator performance on four main, concurrent tasks (tracking, monitoring, resource management, and communications), and is an updated version of the original MATB (Comstock & Arnegard, 1992). Although MATB is a concurrent multitasking simulation, it can be made to serve a constrained, sequentially multitasking and task-switching purpose for examining task management in two important ways. First, instructionally participants are not allowed to multitask in their responses, as they are only able to operate the joystick or mouse with a single, dominant hand. They are not allowed to use two hands, or to switch back and forth between two different hands to speed responding. Secondly, performance in MATB is able to be tracked between four distinct tasks, so switches and switch choice is a measureable outcome variable.

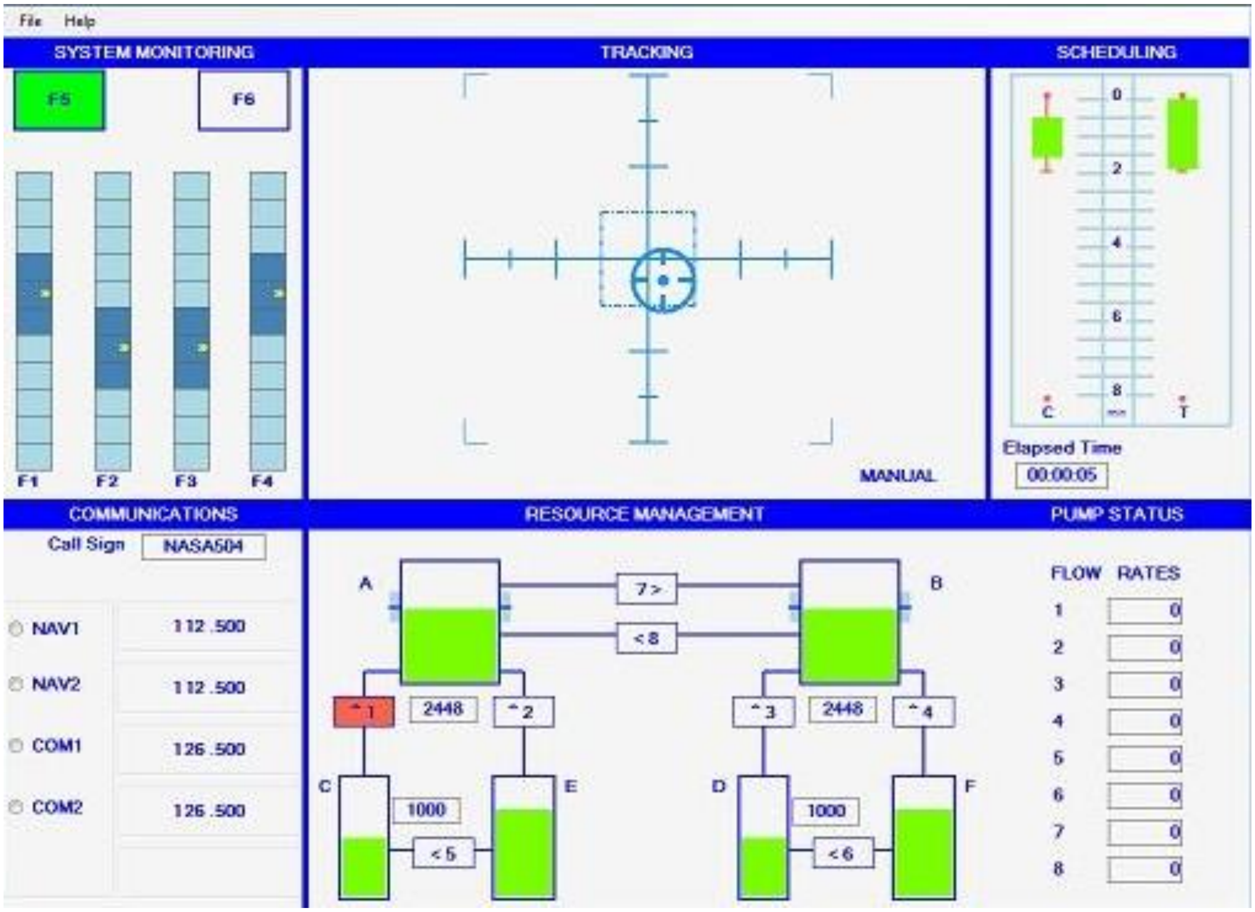


Figure 4.1. The basic MATB II simulation overview, with red circles indicating the four different task areas (clockwise from top left: monitoring (Mon), tracking (Trk), resource management (Rman) and communications (Comm).

All information about all tasks was present visually on screen, with the exception of the communications task that relies on operators to listen to simulated air traffic control messages. Each task, and the possible events that could occur within it, is described in more detail below, and a graphic of the task in operation (Figure 4.1) is provided that should help orient to the program as it is presented to participants.

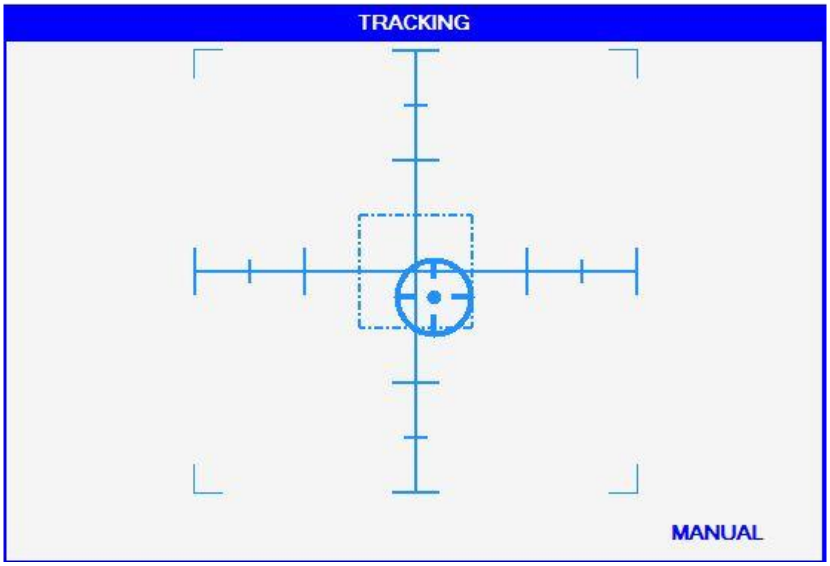


Figure 4.2. The tracking task in MATB II. The circular reticle is under the influence of the underlying random function, which moves it independently of operator action. Participants attempt to control it with a joystick with the goal of positioning it within the square (Exp. 1) and additionally with the small inner circle on top of the intersection of the crosshairs (Exp. 2 and Exp. 3).

The **tracking task (Trk)** represented a two-dimensional random input compensatory tracking task in which the participant attempts to keep a target circular reticle within a small, visible square box (see Figure 4.2). In Exp. 1, the tracking task was active for the entire trial in the *easy* and *difficult* conditions where difficulty was varied by tracking bandwidth, and was only inactive for 3 seconds in the *transition* condition as the program changed the tracking difficulty from easy to difficult at the halfway point.

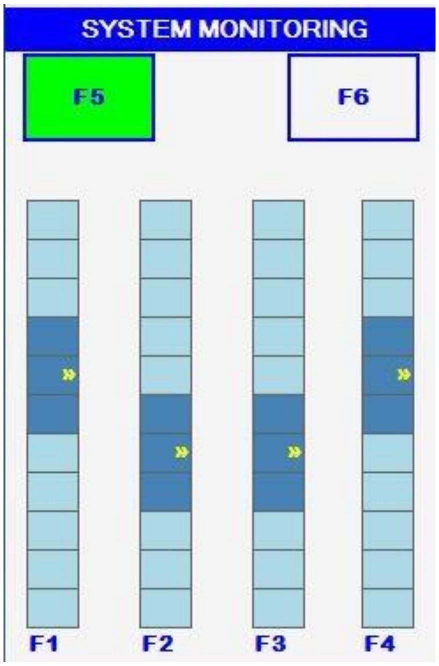


Figure 4.3. The monitoring task in MATB II. In the upper portion, the two light boxes are shown, and in the lower portion the four scales are visible (all readings normal).

The **monitoring task (Mon)** has two main components, lights and scales (Fig. 4.3). Operators responded by clicking the mouse on the affected component: for the lights, operators responded to the onset of a red light, and the offset of a green light. Four scales, with oscillating arrow indicators were also present as part of the monitoring task. Each of these scales can go out of range by registering either too high or too low for a short period. These events also require a click response.

In Exp. 1 participants were asked to detect events and respond by clicking on them. Across test trials, there were four green and four red lights, and eight scale events (two scale events for each scale – one of high and low indicator). The only exception was the transition trial, which had to drop a scale event (scale 3) due to programming constraints. Events occur at semi-random intervals approximately every 32s.

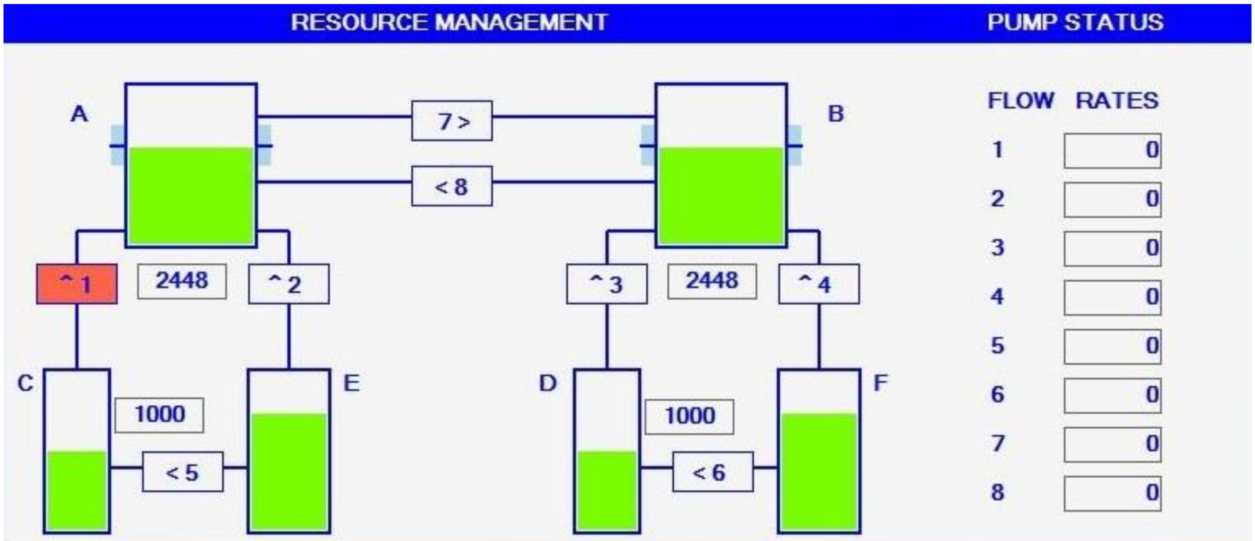


Figure 4.4. The resource management task. Tanks are labeled with letters A-F, while the pumps that allow flow of fluid between them are numbered 1-8, and are accompanied by arrows that show the direction of fluid transfer. The green areas represent the fuel present in the tanks.

The **resource management task (Rman)** represents fuel management aboard an aircraft (Fig. 4.4). Operators must maintain fuel levels in two main tanks which otherwise constantly deplete below critical target levels. These two tanks are connected via pumps to each other, and to four additional tanks. Pumps direct resource flow into or out of each of the tanks, and are controlled by the operator to regulate the levels in the main tanks. Events in the resource management task represent the failing of one of the eight interconnecting pumps. In each of the test trials in Exp. 1, each pump failed once in a randomized order. Pumps are repaired only through scripted events (i.e., automatically by the platform) after about 30s. A more extensive task analysis of the resource management task can be found in Appendix C.

Figure 4.5. The communications task in MATB II. In the upper portion, the participant's ownship callsign was always displayed (and was always the same). On the left, the four different radios that can be selected, and on the right the frequencies associated with them.

The **communications task** (Comm) simulated a pilot interacting with an air traffic controller request. Auditory messages begin with a callsign to denote their intended recipient, and convey an instructed action. Participants are only required to respond if the message is directed to their ownship callsign, "NASA 504". The instructions included during a communications event are to alter frequencies on one of four communications radios (*Com1*, C*om2*, *Nav1*, or *Nav2*) to a new five-digit frequency. Operators are instructed to select the appropriate radio and adjust it to the stated frequency using the mouse (see Fig. 4.5). For a detailed assessment of individual timeframes for information presentation across each Comm task, see Appendix D and Exp. 2. In Exp. 1, test trials contained equal numbers of ownship and other ship instructions (four each) occurring approximately one every 72s, and events occurred once for each radio type, with no overlap in the frequency that was required.

Task Performance Metrics

Performance for each task was measured in the following manner. Task accuracy for the monitoring task was scored as participants correctly responding to an event (whether a red light onset, green light offset, or any of the scale events) within a 10s timeout window. Monitoring reaction time therefore was measured only for trials in which the event was responded to correctly.

Task accuracy for the tracking task was measured by tracking error (measured in pixel deviations of the target reticle center, from the center of the crosshairs) periodically over the course of a trial. Measurements of this error were taken approximately every second and averaged over the entire trial.

Task accuracy in the resource management task was scored much like tracking, in that accuracy was represented by error in the task (deviation in tanks A and B from the target level of 2500). The error was measured about every 10 seconds, or less, as subject responses in the task also triggered a recording event. Absolute values of the error were averaged across the duration of the trial for each tank separately.

Task accuracy in the communications task was scored on a complete basis. There were two components to this accuracy; the correct radio must be chosen as instructed, as well as the correct frequency (again only the event was an ownship event and thus required a participant to respond).

An assessment of time on task was also possible in the MATB platform. Time in task was measured slightly differently between tracking (accumulating total uninterrupted 1+ second intervals spent tracking), and the other tasks (subject actions were scored discretely as taking 500ms each and this time was summed). The value of 500ms was chosen as a lower limit on single action task response times because a more precise measure was not available in this experiment.

Experiment 1 Hypotheses

In Exp. 1, the relative task difficulty and priority of the tracking task component of MATB was manipulated both within (difficulty) and between (priority) participants in a 2 x 2 mixed design. Based on the STOM model and informed by the literature review, three switch-centered hypotheses (*H1*-*3*) were formulated and tested. Two of these were based on task difficulty and one was based on task priority.

First, because task switching is assumed effortful and resource limited, (*H1*) predicted that less switching should occur during difficult tracking compared to easy tracking conditions. Secondly, (*H2*) predicted a difficult task should garner proportionally fewer switches to it, whether difficulty is manipulated or measured subjectively. Finally, (*H3*) predicted that a higher priority task should lead to more switches toward it from a different ongoing task – thus high priority representing an increased AT attractiveness. Other task attribute ratings were assessed, but not manipulated. To the extent that these differed substantially between tasks, we could examine their influence on task switching.

Vitally, the manipulation of priority concurrent with manipulating difficulty represents the first integration of two of the task properties in a task switching paradigm that may influence which AT is chosen when a switch occurs (but see Gopher, Brickner, & Navon, 1982, for effects of priority and difficulty on *concurrently* performed tasks). This helps populate a model and addresses gaps in the literature on task management.

METHODS

Participants and Materials

Eighty-one students at Colorado State University participated in return for optional, partial course credit.

The experiments were performed on a Dell computer with a standard mouse. Operators were given stereo headphones and a Logitech Extreme 3D Pro joystick for performing the communication, and tracking tasks respectively. The multitasking simulation MATB II (Santiago-Espada, Myer, Latorella, & Comstock, 2011) developed from the original MATB program (Comstock & Arnegard, 1992) was used for all trials. MATB presents four independent tasks for operators to perform during a trial. The screens for each task were arranged in a square with two rows (a separation of about 1.6 degrees of visual angle) and two columns (with a separation of .19 degrees of visual angle). MATB instructions were adapted from the MATB II Manual (Santiago-Espada et al., 2011) and covered each of the four tasks present (see Appendix A for the instructions for Exp. 1).

Procedure

The experiment took approximately one hour to complete. Operators were told they would learn how to perform various tasks that were related to flying a plane, and were introduced to the MATB II simulation through a series of instructional, self-paced slides, which they had 10 minutes to review. They were required to view all of the slides one time before returning to prior material. The final slide provided performance goals; in the Equal Priority condition (**EQP**; n = 38), operators were told to perform all tasks as best as possible. In the Tracking Priority condition (**TRKP**; n = 43), operators were told to prioritize tracking over all of the other tasks, while still performing them as best as possible. Task priority was manipulated using instructions for operators to prioritize response to and accuracy of the tracking task (e.g., Gopher et al., 1989) while switching to other tasks only if possible.

Operators were instructed to perform the tasks with only their dominant hand. They were not allowed to use two hands at any point to respond. All responses were made using only the mouse or joystick provided, with the operator required to switch back and forth between them as input when necessary. This critical instruction allowed us to examine task switch behavior without the possibility of concurrent performance. Operators then completed the training trial which contained all of the elements of the MATB simulation used during later experimental test trials, including varying difficulty of the tracking task component, pump failures in the resource management task, own- and other-ship call signs in the communications task, and both light and scale events in the monitoring task.

Following the training trial, participants were allowed to ask questions and were provided answers by the experimenter. After the training trial, but before beginning the test trials, the experimenter reminded operators to perform all tasks quickly and accurately (the equal priority condition; EQP), or that the tracking task was the highest priority and that the other tasks should still try to be performed when able (the tracking priority condition; TRKP). The heightened priority of the tracking task was described to operators as the importance of tracking and aligning an airplane during a landing. Participants then performed three test trials of varying tracking task difficulty (easy, difficult, and transition). Easy and difficult trials were counterbalanced across operators, with all participants performing the transition trial last.

During each MATB test trial, the difficulty of one task in MATB (tracking) was manipulated within subjects by altering the update rate of the tracking task (i.e., changing bandwidth, Wickens, Hollands, Banbury & Parasuraman, 2013). In the transition trial, the tracking task switched from easy to difficult at about the halfway point. Multiple events in all four of the tasks were also presented during all trials, and participants attempted to respond to all task events. Additionally for the experiment several task event pair conflicts were created, wherein two events in two different tasks occurred close together in time (within 500ms of each other, as simultaneous presentation was not possible within MATB II). The arrival time of an event pair conflict varied randomly across trials. The presentation order of tasks within a pair was intended to be random, however the platform internal limitations prevented this. The issue is discussed further in the results section for the conflict events.

Three types of paired conflicts occurred, commensurate with a factorial combination of event pairs between monitoring, resource, and communications tasks (see Table 4.1).

Table 4.1. Graphical outline of the type and number of conflicting event pairs that occurred during each of the test trials in Experiment 1.

	TRK	**MON**	**RMAN**	**COMM**
TRK	x	x	x	x
MON		x	**2 conflicts**	**2 conflicts**
RMAN			x	**2 conflicts**
COMM				x

Each conflict pair type occurred twice over the course of each MATB test trial. These events represented a true free choice between two tasks. Therefore determining *which* of the two conflicting tasks was chosen when a task switch occurred helped determine the influences of the different factors of relevance in driving choice, as outlined in the STOM model.

After operators completed the final test trial a brief survey was administered which asked them to making paired comparison ratings and indicate which tasks were more difficult to perform, more interesting, and were higher priority. Comparison order was mixed between rating variables (see Appendix B). These ratings served two important purposes: (1) it allowed for a manipulation check on whether or not priority instructions were remembered, and (2) it gave some perspective for what MATB task characteristics may be subjectively. It created an opportunity to examine the extent to which the priority, difficulty and interest attributes of the four tasks differed, and influenced task switching choices.

RESULTS[‡]

Two participants' data were excluded from all performance and switching analyses based on outlier analyses (case 1, M = 158.18, group M = 42.35; case 2, M = 175.4, group M = 522) using greater than three standard deviations from the group mean used as a criteria. Additionally 5 participants had missing data from the surveys, and were not included in their calculation or in interpreting the results of the conflict events. Where sphericity has been violated in ANOVAs, based on Mauchly's W reaching significance at $p < .05$, Greenhouse-Geisser corrections were used.

Task Performance

Each task was scored in the manner discussed in the MATB description above. Task performance was only peripherally important in relation to the main issues of interest, in that if tasks are being systematically ignored, then participants may not be taking them into account in switch choice. However it appeared that in both conditions and trial types all tasks were performed (see Table 4.2).

[‡] *While the transition trial was originally of interest when the experiment was designed, it did not capture the comparisons of interest to the STOM model, unlike those between the easy and difficult trials which were counterbalanced. Interpretation of the paired conflict events and switching data were also only desired for effects of priority and difficulty manipulations, and therefore a detailed assessment of the transition trial was skipped.*

Table 4.2. The mean performance on each of the tasks in MATB subdivided by the between subjects condition of tracking priority.

	Tracking Error		Rman Error		Comm Accuracy		Mon Accuracy	
	Easy Trk	Difficult Trk	Easy Trk	Difficult Trk	Easy Trk	Difficult Trk	Easy Trk	Difficult Trk
Equal Priority	21.74	47.00	377.91	482.49	0.28	0.32	0.36	0.36
Tracking Prioritized	21.25	42.35	581.01	675.82	0.28	0.30	0.34	0.30

Note: Tracking and resource management are reported in terms of root mean squared deviations, while communications and monitoring tasks are reported in terms of overall accuracy to events.

As a manipulation check, the impact of tracking difficulty was also examined. Increasing the difficulty of tracking should impact the performance of the tracking task. Difficulty of the tracking task may have also influenced the ability to perform the other tasks. Separate analyses were run for tracking, and the combination of resource management error, accuracy in the communication, and accuracy in the monitoring task.

Although the influence of counterbalancing was found periodically throughout the results, it was tangential to the main issues of interest, and not of clear theoretical or practice importance to the present purposes. The discussion and interpretation of counterbalancing effects is however presented in Appendix G.

A 2 (priority condition: EQP or TRKP) x 2 (tracking difficulty: easy or difficult) x 2 (counterbalance: easy or difficult trial performed first) repeated measures ANOVA was run on overall averaged error in the tracking task. As expected, difficult tracking trials had more tracking error (M = 44.02) than the easy trial (M = 21.51; $F(1,75) = 237.94$, $p < .001$, $\eta_p^2 = .76$). The difficulty manipulation was successful in increasing error in tracking.

To assess the influence of tracking difficulty on *other* task performance, an exploratory 2 (priority) x 2 (difficulty) x 3 (task type) x 2 (counterbalance) repeated measures MANOVA was conducted using resource management error, and accuracy in the monitoring and communications tasks. After checking for multicollinearity (measures correlated moderately, r = -.23 to r = .64), the analyses showed that tracking task difficulty, priority condition, and counterbalance condition exerted no main effects. Only one interaction between tracking difficulty condition and counterbalancing condition was significant in the multivariate analysis (Wilks' λ = .675, $F(3,73)=11.70$, $p < .001$), discussed in Appendix G.

In summary, tracking difficulty appears to have been successfully manipulated, with some residual influences related to the resource management task – a task that likely shares a common resource (e.g., Wickens, 2002). There was no main effect of, or interaction with, priority on task performance in tracking.

Task Switching

Task switching was measured by examining actions taken in each of the four tasks in MATB over the course of each test trial. To restate the hypotheses, there should have been fewer switches between tasks in general under difficult tracking conditions. Second, there should have been more switches to tracking when it was easy, rather than difficult. Third, higher task priority should result in more switches to tracking compared to lower priority tracking. As tracking was the manipulated ongoing task, all comparisons reported are focused on switches to and from the tracking task, which were averaged together. Two hypotheses were related to task difficulty, and were addressed in the following ANOVAs.

A 2 (tracking difficulty) x 2 (priority group) x 2 (counterbalance condition) repeated measures ANOVA was conducted. All sphericity assumptions were met. A main effect for tracking difficulty was found, confirming the first hypothesis – there were fewer task switches under difficult tracking (M= 43.26) than under easy conditions (M= 48.75; $F(1,75)$= 17.22,

$p<.001$, $\eta_p^2=.19$). No main effect of priority ($F<1$) or counterbalancing group ($F<1$) on number of switches related to tracking was found.

Although the confirmation of *H1* was attained and there were more switches overall in easy than difficulty tracking conditions, to specifically address *H2* (fewer switches to difficult tasks), a further comparison of switches when it was only "to tracking" was undertaken. Participants switched to the tracking task less when it was difficult (M = 43.54), than easy (M = 48.28; ($t(78)=3.55$, $p=.001$). These results confirmed the second hypothesis.

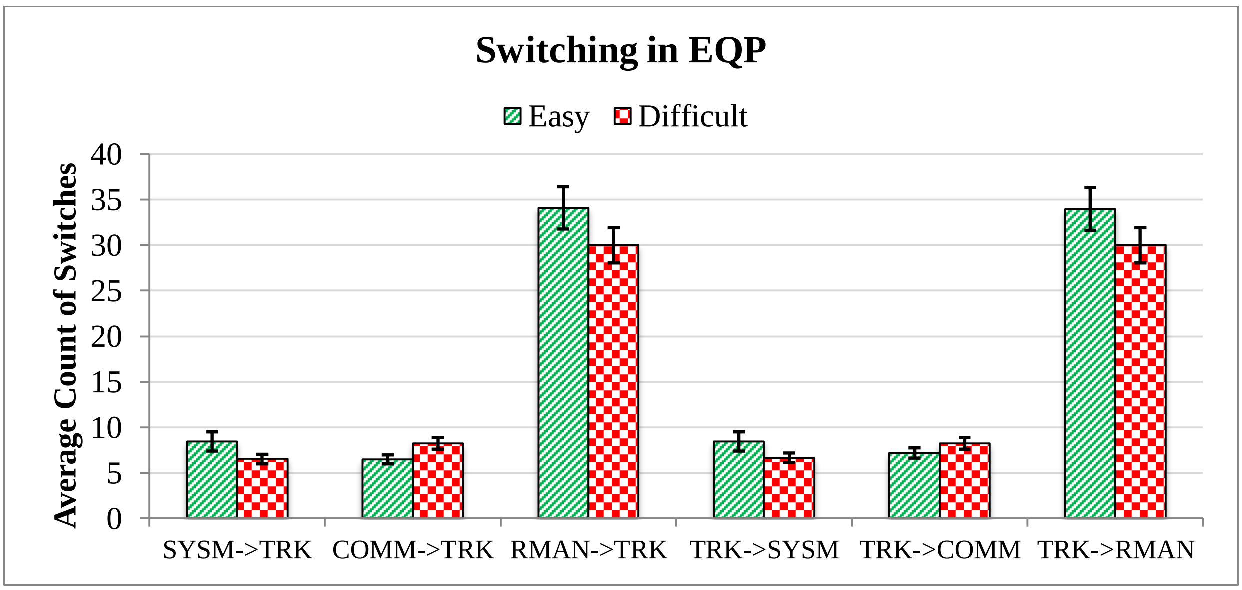


Figure 4.6. Representing the switches to and from, the tracking task overall in the easy and difficult trial in the **EQP** condition. Along the horizontal are the tasks switched from TO tracking (first three pairs of bars) and FROM tracking to (last three pairs of bars). Error bars represent standard error of the means for each category of switch.

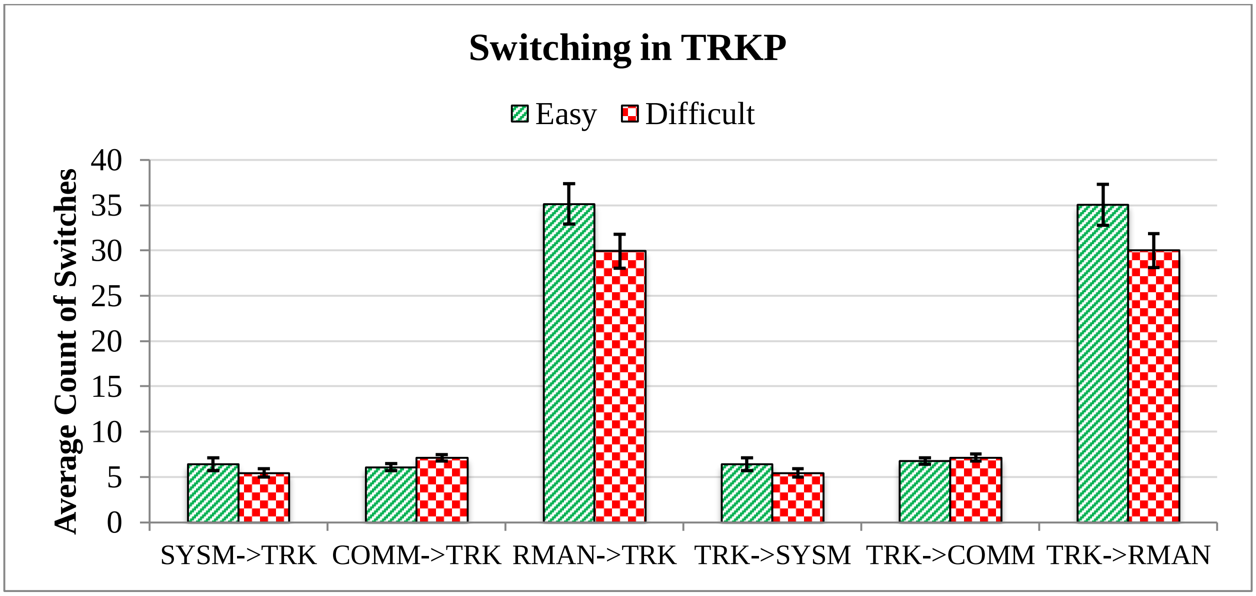


Figure 4.7. Representing the switches to, and from, the tracking task overall in the easy and difficult trial **in** the **TRKP** condition. Along the horizontal are the tasks switched *to* from tracking (first three pairs of bars) and *from* tracking to (last three pairs of bars). Error bars represent standard error of the mean for each category of switch.

In Figure 4.6 and Figure 4.7, total switches to (first three bar pairs) and from (last three ar pairs) tracking are graphed for both the EQP and TRKP condition. Although no effect of priority was found when analyzing tracking switches above, a visible difference can be seen between tracking difficulties on number of switches for the task pair of tracking and resource management when tracking was prioritized (Fig. 4.7), but not equal priority conditions (Fig. 4.6). Additionally, following from the performance analyses earlier, the effect of difficulty was only evident for the tracking and resource management tasks, conceptually linking them together. Therefore, two separate exploratory 2 (difficulty) x 2 (counterbalancing) ANOVAs were run on combined tracking and resource management switches, one for each priority group. Counterbalancing had no effect in the EQP or TRKP initial analyses, and was not included in the final ANOVAs.

In the analysis of the equal priority group, a marginally significant effect of tracking task difficulty was revealed, and fewer switches were found on the difficult trial (*M*= 74.8) than on the easy tracking trial (*M* = 83.38; $F(1,36) = 3.88$, $p = .06$, $\eta_p^2 = .10$). The same effect was found for the tracking prioritized group with fewer switches on difficult (*M* = 72.5) compared to the easy tracking trial (*M* = 83.05; $F(1,41) = 13.27$, $p = .001$, $\eta_p^2 = .24$). The analyses did not reveal an influence of task priority in terms of switching.

Despite the lack of an effect of task priority for switching frequencies, it could have been the case that participants simply stayed longer with tasks under higher priority or higher difficulty tracking conditions (in other words, spent less time on other lower priority, or easier tasks). An isolated analysis of task switching may miss the influence of task priority and difficulty relative to 'staying' with ongoing performance in the task. To address this, an exploratory analysis of time spent in the tracking task was undertaken.

Time Spent in Task

A 2 (priority condition) x 2 (difficulty condition) x 2 (counterbalance condition) repeated measures ANOVA was run on the summed time spent in the tracking task on each trial. A main effect of difficulty revealed that significantly more time was spent in the tracking task under difficult (*M*= 526s) compared to easy (*M*= 518s) conditions ($F(1,75) = 11.53$, $p = .001$, $\eta_p^2 = .13$). No main effect of priority ($F<1$) or counterbalancing was found ($F<1$). Critically, a marginally significant interaction between difficulty and priority emerged ($F(1,75) = 3.78$, $p = .06$, $\eta_p^2 = .05$) which suggested the difference between time spent in the easy and difficult tracking conditions was larger when tracking was prioritized (M_{easy} = 515s; $M_{difficult}$ = 528s), than when it was not (M_{easy} = 520s; $M_{difficult}$ = 522s). In fact, the difference was *only* significantly different in the prioritized tracking condition ($t(41) = -4.51$, $p < .001$) and failed to reach significance for the equal priority condition ($t(36) = -.43$, $p > .05$).

The above results provided limited evidence of a successful priority manipulation on time in task.

Task Attributes Survey

Participants provided paired task comparison ratings for three main categories of relevance to the STOM model: priority, difficulty and interest. Salience was not addressed in the surveys because it was decided a priori, based on the interruption management literature, that the Comm task with its auditory nature was more salient than the other visual tasks.

To showcase this data globally, each task was scored across all three of its ratings for priority, interest and difficulty. Higher ratings indicate a more positive global score for that variable (e.g., priority was rated higher for the tracking task in the TRKP condition than in the EQP condition). The results are shown below in Table 4.3 and 4.4.

Table 4.3. Subjective ratings summed for each of the four tasks in MATB for each condition.

	Priority		**Interest**		**Difficulty**	
Task	**EQP**	**TRKP**	**EQP**	**TRKP**	**EQP**	**TRKP**
Monitoring	-3.13	-3.98	-3.91	-4.29	-2.91	-3.56
Communication	0.06	-2.20	-0.09	-2.10	-2.03	-2.24
Resource	2.65	-0.24	3.34	1.51	4.69	3.10
Tracking	0.53	6.41	0.66	4.88	0.25	2.71

Note: Higher values represent higher attractiveness overall for that variable, except in the case of difficulty as it is reversed. **EQP**=equal priority; **TRKP**=tracking priority.

Task priority. Participants rated tracking as higher priority than the three other tasks in the TRKP condition (Table 4.3). In the equal priority condition, the task with the highest priority rating was resource management (2.65) with the next highest, tracking (0.53); compared to the

TRKP condition where tracking was highest priority (6.41) and the next highest priority was resource management (-.24). The Mon task was never rated as a high priority task in comparison to the other tasks, even under equal priority instructions when ostensibly equal priority should be given to its performance.

Task interest. The resource management task was the most interesting for the EQP condition (3.34), but for the TRKP condition, tracking was of greatest interest (4.88), well above the next highest task in interest (Rman = 1.51). Monitoring again was indicated as one of the least attractive tasks overall, perhaps additionally contributing to its poor performance. Based on Table 4.3, a pattern begins to emerge, in that communications and monitoring tasks take a back seat to resource management and tracking in terms of priority and interest ratings.

Task difficulty. Each group rated the resource task as the most difficult task, with EQP rating it highest (4.69) and the next highest rated task, tracking, only receiving 0.25. In the TRKP group, resource management was rated lower but still the top rated task on difficulty (3.10), and the second highest was also tracking (2.71). The easiest task was rated as the Monitoring task for both EQP (-2.91) and TRKP (-3.56) groups, with the next easiest rated as the communications task (-2.03; -2.24, respectively).

Paired Conflict Events

Although assessment of each factor could be useful in isolation, an equally useful way to understand tasks is to provide an amalgamation of all three factors for each task (see Table 4.4). The global ratings of hypothesized task attractiveness allow a comparison and a prediction between each task based on overall attractiveness.

Table 4.4. Subjective ratings across all three attributes, with equal weighting for priority, interest, and difficulty (a negative attribute).

	Global Subjective Rating	
Task	**EQP**	**TRKP**
Monitoring	-4.13	-4.71
Communication	2.00	-2.05
Resource	1.30	-1.83
Tracking	0.94	8.59

Note: Higher values represent higher attractiveness for that taskoverall; **EQP**=equal priority; **TRKP**=tracking priority.

Task switching related to each of the 12 paired conflict event was assessed in number of switches made following each task arrival. Thus for each of the three types of conflict events, it was possible to determine the overall percentage of switches to one event of the pair in relation to the other. These percentages are shown in Table 4.5.

Table 4.5. Focused conflict analysis results table.

	EQP		TRKP	
	Mon	Comm	Mon	Comm
MON VS COMM (other)	37.55	4.26	29.24	3.54
	Mon	Rman	Mon	Rman
MON VS RMAN	33.00	39.33	30.34	38.59
	Comm - Other	Rman	Comm - Other	Rman
COMM (other) VS RMAN	12.65	65.77	7.59	70.93
	Comm - Own	Rman	Comm - Own	Rman
COMM (own) VS RMAN	66.22	28.38	56.62	31.30

Note: Only the switches for tasks involved in the pair are shown. Numbers are percentages of total switches, averaged across difficulty level, as no clear differences exist between them. **EQP** = equal priority tracking, **TRKP**= tracking prioritized.

Each of these conflict event pairs represented a potential critical overload point for the operator, as two tasks need attention in addition to the ongoing tracking task, and fits in most closely with a condition that elicits single channel processing of information (Liao & Moray, 1993; Wickens & McCarley, 2008). In the experimental design, 12 total opportunities were provided for two simultaneous alternative tasks (AT) to arrive while subjects were engaged in the OT of tracking, in order to establish the tendencies of subjects to choose one of these ATs over the other. The 12 paired "switch opportunities" or conflicts were defined by four replications each of a choice between Rman and Mon, between Rman and Comm, and between Comm and Mon.

With each of these pair types, objective data on which task of the pair was chosen to switch to (e.g., degree of switch preference), could be compared with the predicted data from the global task attractiveness ratings, shown in Table 4.4. The task attractiveness ratings suggested both Rman and Comm would "dominate" Mon during these conflicts, since the former have more attractive (more positive and/or less negative) values than the latter. Averaged across both EQP and TRKP conditions, the Rman and Com tasks appear to be approximately equally attractive (-0.25 and 0 respectively). Unfortunately, two aspects of the data collection procedure prohibited us from collecting truly unbiased switch preference data.

First, due to constraints in the program, the Mon event onset always preceded the other conflict task (Rman or Comm) by 500ms. As a result, the Mon task has an effective "head start" (and the other task a "slow start"), and so any data showing greater switch tendencies toward Mon, may be due to the head start, and not because of other STOM model attributes (e.g., difficulty, which would also predict a switching preference to monitoring).

Second, the Comm task is divided into two very different classes of events, depending on whether the call sign designates *ownship,* in which case the information is highly relevant and requires further processing, or *othership,* in which case it is essentially irrelevant. In terms of STOM attributes, it could be said that this separates the two types of Comm events into high priority, and zero priority versions within the same task. However in collecting the attribute ratings for the Comm task (Tables 4.3, and 4.4), participants did not make this distinction, and hence we cannot decompose their overall ratings of the Comm task into separate ratings for the two different classes of Comm events. A further complication was that an equal number of ownship and othership events were not collected for each of the four replication pairs involving the communications task. In the case of the Mon-Comm comparison, othership events were the only Comm events that were presented. The complications are addressed below.

Because of the relatively small number of events presented, data were pooled across the four different levels of tracking variables (high and equal priority, and easy and difficult), justified because there was no a priori reason why these differences would affect the *relative* switch choice preferences for the remaining tasks.

In Table 4.6, four rows are presented depicting four different conflict tasks pairs: Mon vs Comm (Other), Rman vs Mon, Comm (other) vs Rman, and Comm (own) vs Rman. In the second column the "dominance" of choice is presented (the dominant task, and how many choices favored it, the unfavored task and how many choices favored it). The disparity between the two numbers represents the overall strength of tendency toward the dominant task. In the next column ("why"), the reasons (including individual attribute ratings from Table 4.4) why the dominate task might have been dominant, and in the final column ("Despite") reasons why the unfavored task might have had increased attractiveness, ***but did not.*** In the "Why" and "Despite" columns, the 500ms head start or slow start reasons are given, as well as the salience attribute of STOM even though salience was not solicited as a rated attribute. The final column presents the total attractiveness as estimated by averaging the values from Table 4.3, across the two tracking priority conditions. The "?*" for the estimates for the Comm-other and Comm-own indicates the uncertainty as to how much more relevant (higher priority) the task event is assumed to be when it is *ownship* rather than othership. An arbitrary assumption was made of 4 rating points of difference between these.

Finally, it should be noted that one qualitative difference in salience existed between pairs containing the Comm tasks, even though this was not subjectively rated. The Comm task always is initiated with an auditory event, hence giving it higher salience.

Table 4.6. The four types of conflicting events are presented, one in each row.

Dominance	Why:	Despite:	Attraction
Mon > Comm other; (**34** >4)*	Easy, Priority (Com other= 0); 500ms head start	Lesser interest, lesser salience	**-4.5** vs -2?*
Rman > Mon (**77** > 20)*	Interesting, Priority	Harder, 500ms "slow start"	**-0.2** vs -4.5
Rman >Comm (other) (**67** > 20)*	Interesting, Priority (Com other=0)	Harder	**-0.2** vs -2?*
Comm (own) > Rman (**61** > 30)*	Priority, salience, easier	Lesser interest	**+2?*** vs -0.2

Note: The **Dominant** task is in bold versus the comparison task, and the relative number of switches to the 2 tasks made is presented below (* indicates statistical significance by chi-squared test of proportion). In the **Why** column, the evidence for why the dominant task dominated is presented, and in the **Despite** column, the factors that went against the dominant task are listed. In the final **Attraction** column the averaged overall subjective ratings for the two tasks are presented. "?*" refers to values that include an assumption of priority difference between the two different Comm tasks (own and othership).

Each of the comparisons in the dominance was entered separately into a chi-square test of proportions. All of the comparisons were significantly different than equal proportions; Mon was chosen more often than Comm other ($\chi^2(1) = 22.14$, $p < .001$), Rman more often than both Mon ($\chi^2(1) = 32.32$, $p < .001$) and Comm other ($\chi^2(1) = 24.32$, $p < .001$), and Comm ownship responses dominated Rman ($\chi^2(1) = 9.9$, $p < .01$).

One feature that rules the dominance data in the first column of Table 4.6 is that switching preference is consistently driven by higher priority, whether this was manifest in the subjective ratings, or in an implicitly assumed difference between the Comm other/ownship events. In the two middle rows, comparisons with Rman can be seen; of note, despite the more favored task (Rman) being harder and, in one case, having its onset penalized by a 500ms slow start (row 2) the Rman task is rather dominant. This relationship is presumably *because* of its substantially greater interest, and *despite* its greater difficulty. These findings may collectively suggest priority and interest should play a greater role in the STOM model (greater weightings) than task difficulty, predicting competition with other tasks for attention.

Experiment 1 Discussion

Returning to the three main hypotheses concerning switching, first, fewer switches in general were predicted to occur during difficult tracking conditions. Indeed, this was the case. Importantly this finding of "effort avoidance" agrees with the prior literature review, which incorporated studies that used the more basic voluntary task-switching paradigm (e.g., Wickens et al., 2013; Wickens et al., under review). It is a novel finding primarily because prior work has focused on dual-task situations where switching is a default choice to the "other" task, instead of the current multi-task paradigm where a switch could be to one of three other tasks. Additionally, the prior work has not clearly tested these hypotheses in the high demand environment that the STOM model predicts within.

Secondly, it was expected there would be fewer "to" tracking switches when it was difficult; this effect was also found. Hence the difficulty of an ongoing task exerted several of the predicted effects and represents the most validated and successful parameter of the STOM model (Fig. 3.1) thus far, reflecting both the overall cost of limited cognitive resources, and the effect of this cost on decision making, when the decision is assumed to be deliberative and conscious, aspects of ‘system 2’ thinking (e.g., Kahneman, 2011).

Third, a main effect was predicted with task priority, such that more switches to the tracking task was hypothesized to be found in the prioritized, than in the equal priority condition. No clear evidence was found to support this third hypothesis in the switching data itself; however, secondary relationships did emerge. When tracking time on task was analyzed, priority of tracking marginally interacted with task difficulty; participants spent more time tracking in the difficult compared to the easy tracking condition, and this difference was only evident in the tracking priority group. The survey data do suggest the priority manipulation was at least successful in influencing subjective prioritization, in a fairly dramatic fashion: priority ratings of the tracking task were much higher in the tracking priority, than the equal priority condition (and prior work has shown instructional manipulation of priority effects are effective, though in concurrent multitasking; Gopher et al., 1989). For future testing of priority effects, then, it appears critical to examine not just switching choices, but also their consequences (in this case, by looking at the time spent performing a task).

The analysis of performance data supported the influence of tracking task difficulty manipulations on tracking task performance, but also on resource management, in line with the literature on dual-task decrements (Welford, 1967; Pashler, 1994).

Experiment 1 attempted to populate the STOM model with additional data, and compare the relative strengths of existing parameters, and task attribute ratings played a useful role. They represent a novel contribution to literature on the MATB paradigm, but more specifically for the purposes of this work they confirmed the priority manipulation. In addition, ratings explain the results of task choice when two events are presented near-simultaneously to participants, and may help adjust the relative weights of each rated factor in future revisions of the STOM model. Using the conflict assessments, it appeared that in at least two cases, subjective interest and priority, were better predictors of a task choice during the conflict than difficulty.

CHAPTER 5: EXPERIMENT 2

What Triggers A Task Switch?

Do time on task predictions tell us about whether a task switch from an ongoing task is likely? What is the role of increased practice on general switch behavior? And, are individual differences in attentional control related to task switch frequency and choice?

Intuitively, task switch choice may be influenced by time-on-task factors. For example, spending a long time reading a lengthy document, you may either get bored (time to do something else), or instead, you may become more and more wrapped up in the content and so instead of falling asleep, choose to finish the final chapter because there are "just a few more pages!"

In Exp. 1, factors that influenced overall task switching behavior, as well as what task was chosen were explored. In Exp. 2, a further explanation of what could lead to a task switch was explored. Several potential explanations are given related to whether switch likelihood is related to how much time is spent performing a given task, which can be divided into those that make monotonic predictions (switch resistance increases or decreases monotonically with time on task) and those that make more periodic predictions (switch likelihood varies across time based on other factors, as related to fluctuations in tasks and cognitive demands). In the STOM model, these effects would operate generally on the switch versus stay decision ratio (it could be called switch resistance), but the model currently includes no time-based change. A general outline of the different potential effects involved, described in more detail below, is presented in Figure 5.1.

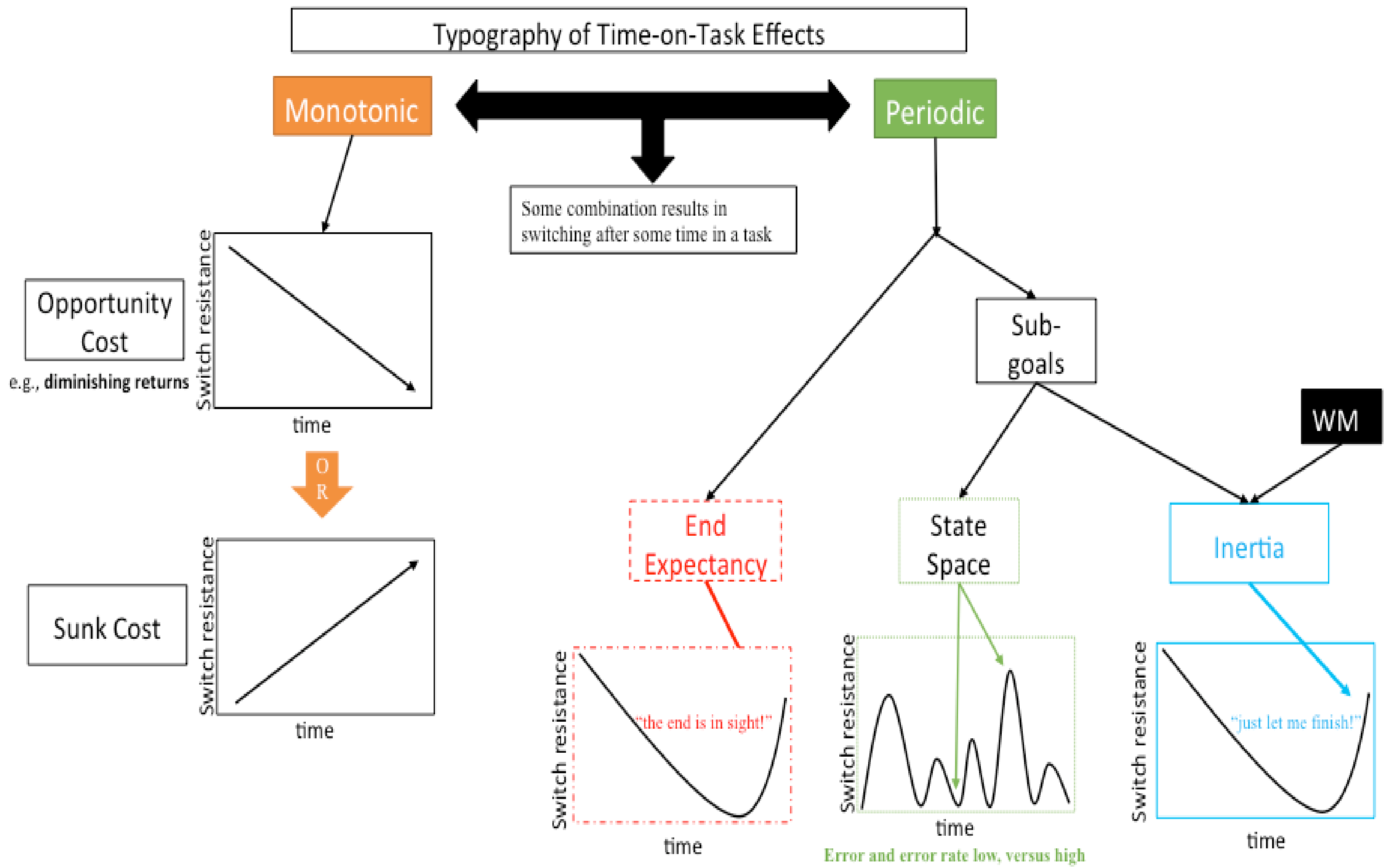


Figure 5.1. The main hypothesized factors that may play a role in determining switch choice over time. WM = working memory. Graphs represent the change in switch resistance over time in an epoch.

Monotonic Effect Predictions

Kurzban et al. (2013) suggest one reason to switch away from an ongoing task is the inherent opportunity cost of withholding limited attentional resources from other alternative tasks. Further they claim that this cost is manifested as a subjective feeling of effort, and when higher value ATs are present, this effort accumulates at a higher rate. Using this theory, an OT would be abandoned earlier or more often when ATs are of high value compared to low value. In either case, general switch resistance decreases in a monotonic fashion as a function of time on task until a threshold is reached and a switch occurs.

Another theory (Sheridan, 2007) suggests with increasing time spent in the OT, there is a mounting uncertainty in the alternative task states. For example, Sheridan (2007)'s attentional modeling posits the need to update the information about alternative tasks (creating an opportunity cost problem if not done, as elaborated by Kurzban et al., 2013), as well as switch decision making being complicit with the value of continuing to attend to an ongoing task. The cost of AT sampling is pitted against optimal sampling strategies – in other words the value of attending is a function of how recently information was sampled, along with the cost of a sample (i.e., the switch cost as discussed within this dissertation). With increasing time away from an alternative task, then, the value of sampling that task increases (Sheridan, 2007). While these two theories make general claims about information sampling, they do not specifically address task choice in terms of actual performance.

One explanation of diminishing switch resistance at odds with Kurzban et al. is addressed generally by Baumeister and colleagues (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Baumeister, Vohs, & Tice, 2007). Their theory posits that if an OT demands mental resources, the switch away is motivated by the need to switch to *different* demands - as related to different tasks, or to take a break (as in Jin & Dabbish, 2009) to “restore” effort. This effort-depletion account is counter to the Kurzban explanation, in that switch effects would be related solely to the ongoing task, and its resource supply, and not any value-based property of the AT. The presence of fatigue effects that occur *without* an alternative task present also bolster the effort depletion argument (Hagger, 2013). The explanation provided by this theory is that switch resistance should decrease linearly with time but depend on demands of the OT. In MATB, for example, increasing time spent under a difficult tracking condition would be predicted to decrease switch resistance at a faster rate, compared to increasing time in the easy condition.

A final perspective covered here is the idea of influencing switch choice through task-related diminishing returns. Similar to search tasks with multiple potential targets, an operator must determine which tasks are more valuable to perform over a given time. The search has even been likened to food foraging strategies with no replenishment, and begins with the assumption that operators choose tasks with higher value and potential for payoffs – the ‘low hanging fruit’ rather than those with less potential (Duggan et al., 2013; Payne et al., 2007). It also fits easily into a “special case” of opportunity cost (see Fig. 5.1), and so pairs nicely with the Kurzban et al. perspective. A more dynamic version of this is diminishing returns: over performance time in a task, initially high returns may begin to decline. The decline in return increases the chance operators will switch to a different task, but such an aspect of the theory (proposed in Kurzban et al.) has yet to be experimentally tested. To foreshadow, it is later tested within the multiple alternative task, high load workload context of MATB (Ch. 6).

One final potential monotonic effect could be a ‘sunk cost’ phenomenon. In other words, opposite of the theories of Kurzban, Baumeister, and Sheridan, perhaps over time the cost of interrupting or abandoning an ongoing task is increased. Other accounts of monotonic time-on-task effects may be possible, but, no current theory or data here speak to them.

Periodic Effect Predictions

Three other related factors are integral to this discussion, all of which are more periodic in their hypothesized effects. Each depends more on characteristics of the task, a specific phase of task performance an operator may be in, or a task/subtask boundary condition, rather than a general "time spent performing task" factor.

The first is **task inertia**. Task inertia suggests tasks will be prioritized ("stayed with") more often when they are close to completion because the difficulty or demand has built up over the task. This is especially the case when the performer can predict a clear endpoint. In theory this is similar to dealing with interruptions from a colleague at work – often, you can negotiate to delay the conversation a few seconds, finish writing a critical email and hit "send", instead of risking losing your place or forgetting to send the message altogether. Tasks that create inertia are typically ones in which WM demands associated with maintenance of information from the OT, or high costs in re-orienting to the task, both highly related to the maintenance of a goal state which drives behavior (Altmann & Trafton, 2002). A switch away from an OT before it is complete (like leaving the kitchen before the turkey in the oven is done cooking) generates the need for a prospective memory to order to return to the task, and complete it at a later time (Dismukes, 2010; Trafton, Altmann, Brock, & Mintz, 2003). Because prospective memory may fail to return the operator to the original task, the closer current performance gets to the final steps of a task, generally, the less likely people should be to switch away from it. Thus, inertia effects help an operator to improve performance, rather than risk a switch when a return becomes less guaranteed over time, or requires refreshing working memory on the task state that existed at the time the task was left.

The second is **task end expectancy.** Expectancy is defined here as the knowledge of the upcoming boundaries of a task; knowing that, approaching a red light, the driving task is temporarily on “pause”, so using a cell phone briefly is not disruptive. Or vice versa, that a green light means resumption of the driving task, and to put the phone down. The likelihood that the task continues to be performed and switches should be resisted in knowledge of an upcoming end is similar to task inertia, but does not rely specifically on memory. In addition, expectancy developed about another alternative task onset (the turkey is almost done and needs to be taken out of the oven) may prime the operator to leave a currently ongoing task – thus prioritizing its completion or at least priming a stopping point (e.g., looking for the sub-task boundary). Both are periodic effects that occur at the end or beginning of a task.

Two tasks in MATB that follow this logic are the Comm and Trk tasks. In the Comm task, clear working memory demands are present. Operators listen to information presented in sequence over a small time window, memorize it, and enter it in to a task window. The clear “end” point, relevant for inertia effects, is the “enter” button confirming the entered information. Switch resistance for Comm performance may increase because of the demand on WM as time on task increases (Salvucci & Bogunovich, 2010) specifically because the information must continue to be maintained until it is entered.

In the tracking task, the task scheduler signifies end points of the task. Operators are able to observe the end of the task period approaching, and may use this information to attempt to keep error in the task low until the task is complete. However there is an absence of WM demand in tracking, unlike that present in the Comm task, and so the inertia effect, reflected in the up curve on the right side of the functions in Fig. 5.1 may be reduced when tracking is an OT. Using these two tasks, potential task inertia (Comm) and expectancy effects (Tracking) can be examined under both conditions.

The existence of **task or subtask completion ‘boundaries’** is also likely to influence switching over time. Subtask boundaries are a known “trigger” for switches in many dual-task experiments, and completing a task or a portion of a task serves as a natural indicator that it may be optimal to switch tasks (as does a delay in a task; Katidioti & Taatgen, 2013). When dialing a phone number for example, which follows a set format of digits such as *XXX-XXX-XXXX*, operators are more likely to switch tasks at the break points within the number string (Brumby et al., 2007; Janssen, Brumby, & Garnett, 2010; Monk, Boehm-Davis, & Trafton, 2004; Salvucci, 2005; Trafton & Monk, 2007). But not all tasks have such clear boundaries, and, boundaries may be idiosyncratic to individual operators. For example within the tracking task of MATB, there are not clear subtask boundaries or goal state transitions in part because of its absence of WM demands. However there may be periods of tracking that are better candidates for switch than others. Given that tracking difficulty is also manipulated between easy and difficult, it is possible that more effortful conditions may increase the likelihood that these boundaries are utilized. In Back, Cox, and Brumby (2012), more difficult information access costs led to more subtask completions before switches – in other words difficulty in maintaining information led participants to switch more at the boundaries. A similar effect may exist in tracking.

During tracking, participant attempt to minimize the error distance between their reticle and the target: when this distance is large, error is large. There is another property to the tracking task outcome of some importance: the *rate* of error change. The rate of error should influence whether the tracking task can be switched away from. In fact an analogy to steering control (alluded to in Brumby et al., 2007) can be made. For example, drivers might be much more inclined to take their eyes off of the roadway (switch attention) if the car is in the center of the lane – low *magnitude* of lane deviation - and not drifting out of the lane (a low *rate* of lane deviation). The rate of decline in performance should cause a rational operator to return to the task more often. Error magnitude and rate then interact to form predicted "optimal" switching time periods.

Table 5.1. A representation of the state space for tracking.

Negligible Error	Improving even more	**Optimal switch**	Trk may need attention
Small Error Magnitude	**Improving**	**May ignore: switch candidate**	Trk may need attention
Large Error Magnitude	**Trk needs continued attention**	**Trk needs attention**	**Trk needs attention**
	-----	**0 value**	**+++++**
	Error Rate		

Note: On the vertical axis, the magnitude of error increases from neg. to larger error; on the horizontal the error rate first derivative of the error in terms of whether error was decreasing, staying the same, or increasing.

When the magnitude of error is low (reticle is close or on the target area) and the rate of change of error is low or zero (reticle is not moving away from the target at a high rate via the underlying drift function or user input), a switch away may be very low cost. In contrast it is non-optimal to initiate a switch when the magnitude of error is high, and the rate of error change indicates further increases in error are likely to continue occurring or at least error is not declining. Such a perspective in MATB always views the tracking task as the OT, and any other tasks as alternative tasks.

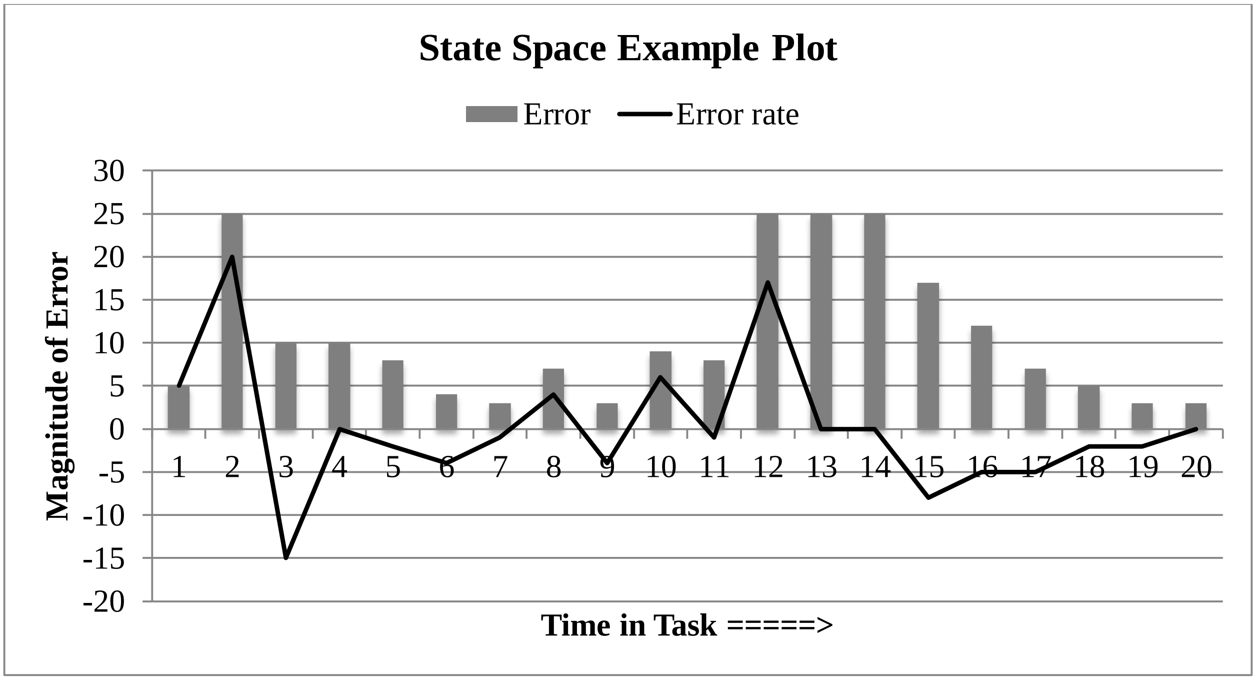


Figure 5.2. An example of state space data for a fictitious run of a tracking task over time 1 through time 20. Although error magnitude (bars) may be relevant for task switching, the *rate* of error (line) may interact to form the state space guidelines as alluded to in Table 5.1. Optimal switch periods in this graph would be times 6-9 when error is relatively low and the rate is relatively stable. Non-optimal periods would be similar to times 1 & 2, and 11 & 12, when error is high and increasing.

Within this “state space” the periodic time-dependent switching perspective is also clearly invoked. Using this, it is possible to illustrate and compare different theories of time on task dependent switches (Fig. 5.2, Fig. 5.3), which have never before been tested empirically. There are clear advantages to using this approach to characterize tracking task behaviors. For example, endogenously motivated switches, a large portion of real-world switching (Jin & Dabbish, 2009; see also Hoffman, 2013) can be tracked within the state space even when no external events may be influencing switching directly. Also the ability to test the state space predictions exists in concert with the ability to examine the other theories of time on task dependent switching resistance.

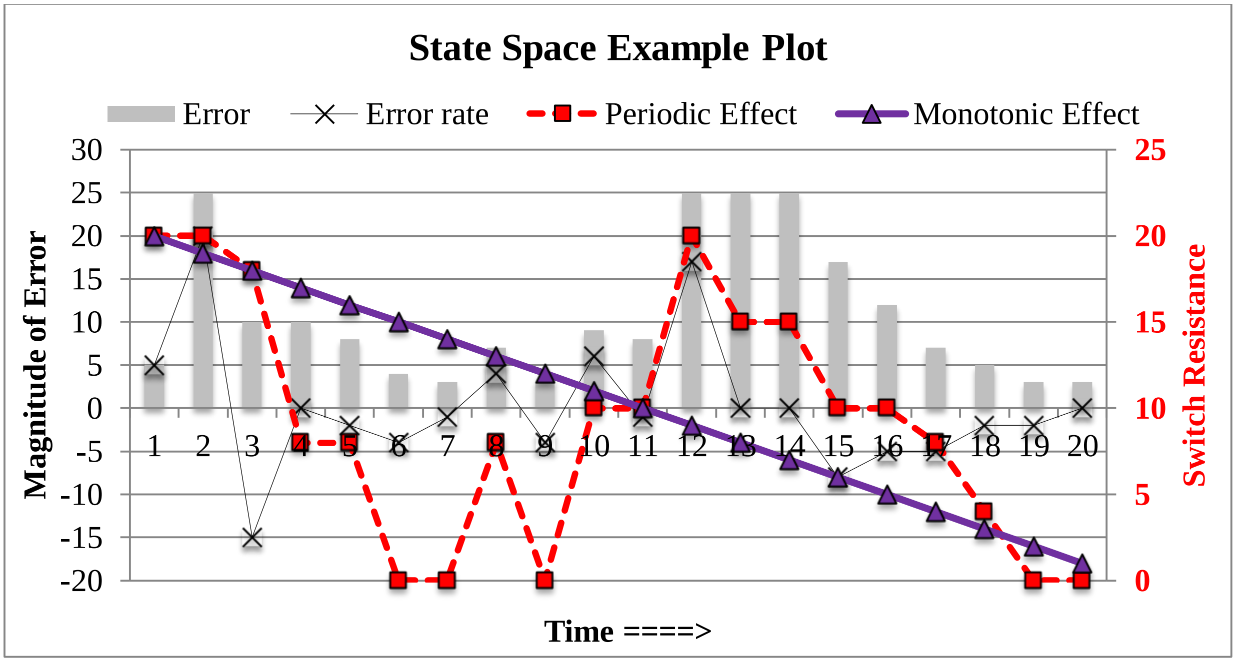


Figure 5.3. A fictitious plot of state space data, now including predictions of time-on-task effects. Error for the tracking task is plotted on the left vertical axis, and switch resistance is plotted on the right. The relative predictions of Kurzban et al. (monotonic decrease in switch resistance over time on task) plotted in solid purple, triangles. This is contrasted by the switch likelihood expected in tracking as calculated using the state space, a periodic effect (dotted red line, squares).

Figure 5.3 highlights the main claims synthesized from the earlier introduction sections, as related to the tracking task in MATB. The figure highlights the role of monotonic, and periodic influences that may be related to time on task. The monotonic effects can be extrapolated to any other task, while the periodic effects are probably task specific. For example, the Comm task differs because of the inherent WM load, clear sub-goal completion points and user-determined task endpoints. The next sections discuss elements of the Comm task and why it was chosen as the other task in MATB that could be examined for time-on-task dependent switching behaviors.

Communications Task

The communications task was chosen because unlike tracking, it represents a discrete task with distinct phases of early, middle and late performance. Responding to the Comm task (once it is determined it is relevant -- ownship) requires three steps, including choosing radio type, entering frequency, and clicking a final confirmation button. Based on the description and analysis of the Comm events, three periods were chosen for interruption (Fig. 5.4) based on the hypothesized WM load and resultant switch resistance.

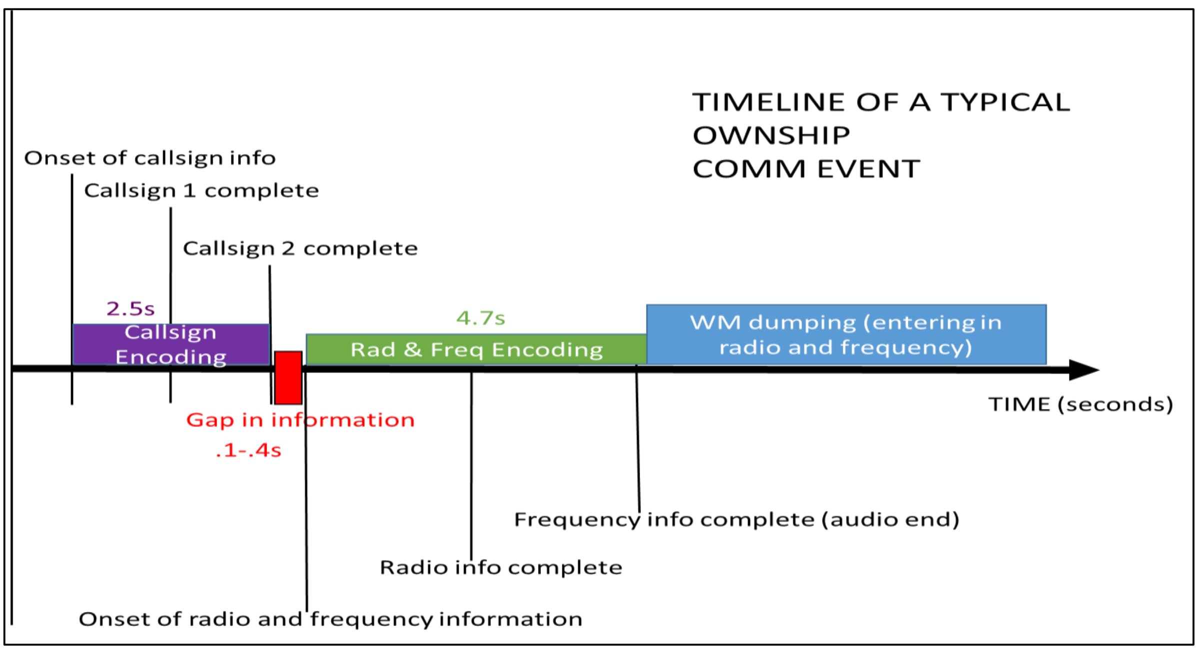


Figure 5.4. An example timeline of a typical ownship communications event. Phases targeted for interruption include the callsign encoding period, radio and frequency encoding period, and the time immediately following it when instructions are being carried out.

These three periods are slightly different for each communications event because the audio presentation is different across different event files; however, using the audio recording program *Audacity*, detailed times for information presentation phases were obtained by carefully determining the beginning and end of information presentations for callsign, radio, and frequency enunciation for each event (see Appendix D). Thus, a carefully timed interruption of phases of the Comm task was possible, as well as the characterization of potential task boundaries, as interruptions at boundary conditions are usually less severe (Bailey & Iqbal, 2008; McFarlane & Latorella, 2002; McFarlane, 2002).

In Figure 5.5, predicted WM load effects on switch resistance are described, and in Figure 5.6 alternative theories of switch resistance are overlaid.

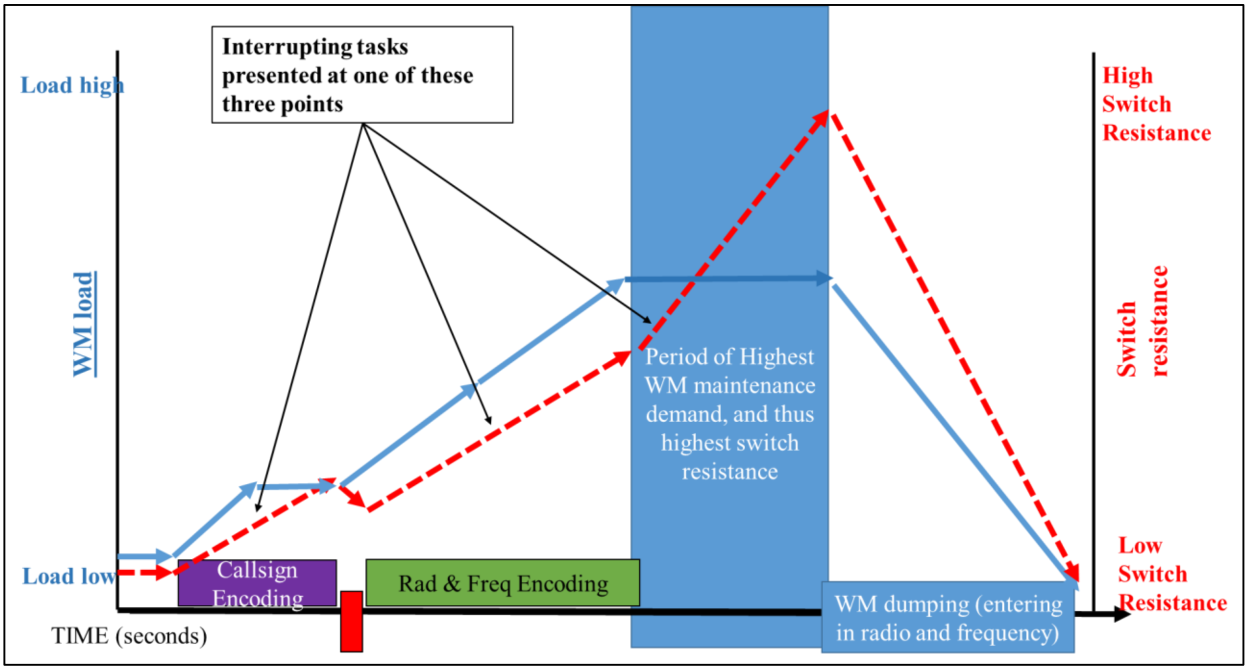


Figure 5.5. Showing the predicted load (in blue) and periodic switch resistance (in dotted red) across time in the Comm task. Using the results of the analysis, three points for interruption of the task were chosen; one during callsign encoding (early), one during the radio and frequency information presentation (middle), and one following the presentation of the information (again, during the critical WM maintenance period).

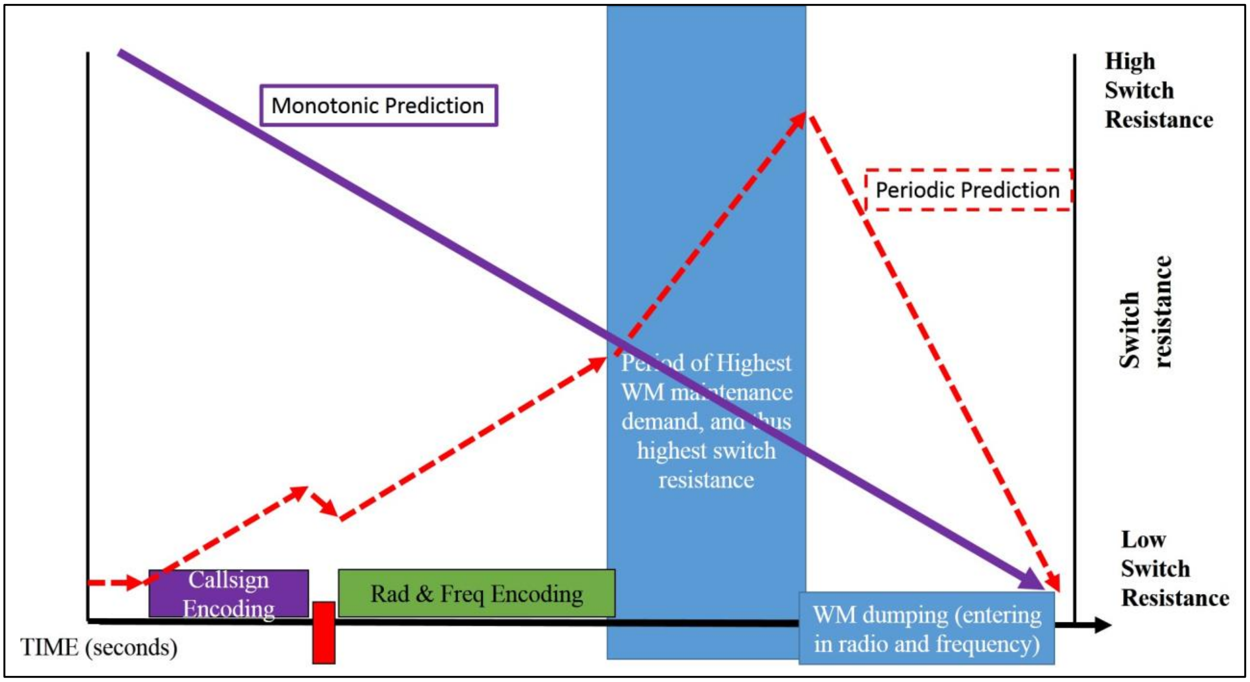


Figure 5.6. Showing periodic switch resistance (in red) across time in the Comm task. Additionally, now the monotonic predictions of Kurzban et al. and others in terms of switch resistance are present (purple solid line).

Interrupting Tasks And Justification

It was critical to consider that the Comm task may be inherently switch resistant because of its auditory nature (Latorella, 1996; Wickens & Colcombe, 2007). Therefore, events used to interrupt both the Trk, and the Comm task phases, were specifically chosen to be high value, salient events. The Mon event chosen was based on observations from Exp. 1, and a brief task analysis of the resource management task. Red-light onset events were responded to well over both easy and difficult tracking trials, in comparison to green light and scale events. Therefore the red light monitoring event was chosen as the interrupting Mon event.

Based on a task analysis of the resource management task (Appendix C), two pumps (1 and 3**)** were deemed the most critical in the maintenance/attainment of the resource management task goal states, because both had an input flow rate equal to the fuel rate consistently 'depleting' in the main tanks**.** Thus if either pump failed, participants have to respond with several other pumps in order to maintain the critical levels in tanks A and B.

In addition to the effect of time on task on switching, a secondary goal of Exp. 2 was to examine switching in MATB under more stable, and extended practice of the simulation. Potential practice related effects are discussed below.

Practice Effects

One realization moving forward from Exp. 1 was that participants may have developed different perceptions of the tasks as a function of mixing task difficulties. Therefore, practice on same-difficulty trials of MATB was increased in Exp. 2 and Exp. 3.

The effect of long periods of practice on task management and multiple task performance is not well known. One hypothesis is that over additional practice, participants may learn to effectively manage tasks as a type of skill (e.g., Damos & Wickens, 1980). Despite the understanding that such improvements occur, the circumstances differ substantially in the current MATB paradigm. Participants are limited to task switching and are also under significant workload. The learning is not expected to be the same, but could still result in improved task management behaviors. Some amount of practice or familiarity with a task may be useful to determine where break points exist, or to learn to resist interrupting events. There is indeed some evidence that with experience, operators learn what to prioritize (Raby & Wickens, 1994) and what can be ignored (even in high demand task management; Koh, Park, Wickens, Ong, & Chia, 2011). This is unlikely to be a function of a massive reduction in switch cost in the traditional sense (Rogers & Monsell, 1995). Little evidence has accumulated to address potential changes over practice.

Therefore, given the lack of a strong hypothesis to make, two main influences are examined for their possible role in switch behavior. The first is that with increased time practicing, task difficulty as a whole is expected to decrease, a function of gains toward automaticity in performance. According to the STOM model, and also the results of Exp. 1, this should result in more switches as practice continues and performance improves. However it could be that task improvement is the result of learning breakpoints – one critical place to examine this is in the tracking task, where more experienced operators should be able to reach optimal switch conditions according to the state space more often. So, over time, switches would be more likely to occur.

Individual Differences

Differences in executive control as they affect switch behavior were one finding in the review of the VTS literature (Chapter 2). Under concurrent WM load, task repetition was increased (Demanet et al., 2010; Weywadt & Butler, 2013). Given the close ties between WM and executive functioning, Butler et al. (2011) speculated higher WMC as related to more adherences to the Arrington and Logan (2004) instructions meant that higher WMC allows for more attentional control. Further, a significant relationship between the control measures of the attentional network task (ANT) suggested higher executive control ability lowered switching likelihood, with increased time pressure (e.g., RSI interval) enhancing this reduction (Arrington & Yates, 2009). One drawback in interpreting these results was that free choice was not allowed in their experiment. Thus the relationship between executive functioning and switch choice was limited to constrained switches. The current experiment used unconstrained, applied environments and represented a stronger test of the hypothesis that executive control may result in differences in switching frequency in more applied domains.

In summary, Exp. 2 sought to examine time-in-task dependent explanations of switching across tracking and communications. In the Trk task, a state space is utilized to characterize the non-monotonic predictions that may shed light on when a switch is optimal. In the Comm task, a task analysis helped reveal the potential for memory maintenance effects across task performance.

Experiment 2 Hypotheses

To test the theories, periods of time within task performance (early, middle and late) were targeted for interruptions. Specifically, examining switches to interruptions early in performance test the extent to which switch resistance is high or low, high resistance being predicted by all theories. Switches to interruptions in the middle of phases of Trk performance test whether the resistance to switching has declined; and whether resistance has increased in the Comm task as a function of WM buildup. Finally, switches to interruptions at the end of task performance for both tasks examines whether expectancy and task inertia increase switch resistance, or whether switches are more likely compared to the other two time periods (as Kurzban et al. would suggest). Within each task, multiple theories suggested that (*H1*) over time, switches to tasks (as measured by responses to interrupting tasks) should increase. Kurzban et al. suggests this may occur even more so for the interrupting task deemed more valuable, which based on Exp. 1 (and, to foreshadow, Exp. 2 survey data) would be Rman compared to Mon.

Within the tracking task, (*H2*) switches should be more likely to occur during tracking when error and error rate are low or negligible, compared to times when error is high and error is increasing, as revealed by the state-space analysis. During the performance of the Comm task, to the extent that operators are under a period of WM maintenance (*H3*) switch resistance will increase as WM load increases across time in the task (e.g., from middle to late time, a task inertia effect). An additional hypothesis was generated for the role of task end expectancy, in that (*H4*) switches away from tracking should decrease over time in task. If found, this would not be an inertia effect because there is no WM demand.

One hypothesis was tested regarding task practice in MATB. First, it was expected as tasks become easier to perform with time, there should be (*H5*) more switching in general. This should be the case for both tracking difficulty levels, though a priori it may be the case that as the difficult tracking condition presumes more room for improvement, the effect may be more pronounced under that condition. And finally, one hypothesis was made about the results of the individual differences task (the ANT task). Specifically, (*H6*) participants with higher results on the executive control outcome will switch less in general, leading to a negative correlation as found in (Arrington & Yates, 2009).

METHOD

Participants

Seventy participants participated in this experiment. Some were students enrolled at Colorado State University who participated in return for optional, partial course credit. Other participants were recruited from the general student body through announcements and postings, and were paid an hourly rate of $10 for participation. Paid participants did not receive incentives for any in-task performance.

Materials

The same computer interface setups were used as in Exp. 1 and a newer version (programmed in house) of the MATB simulation was used. The version used was altered to track joystick action, which provided additional data on tracking task activity. The PowerPoint training materials for MATB were changed. First, the importance of specific interrupting events in the

Comm and Trk tasks were highlighted within the training PowerPoint. Second, the tracking instructions were changed; participants were told to keep the dot (present in the middle of the reticle) positioned on top of the crosshair of the overlay as much as possible. The purpose of such a change was to make the tracking task difficulty manipulations stronger than they were previously (see revised instructions in Appendix E).

Procedure

After informed consent, participants began the experiment by viewing a training PowerPoint document, which they had ten minutes to read and review. Then participants completed the brief training trial as used in Exp. 1. Importantly, during initial training and all following trials in MATB, participants were not allowed to make responses other than with their single dominant hand (e.g., no dual-handed performance, or hand switching, was ever allowed). Thus, when choosing to make responses with the joystick, participants had to have their dominant hand on the stick, and when switching to make responses with the mouse (all other tasks), the same hand had to be moved. Participants were asked before beginning the training trial to make sure that this physical movement was easy for them to make and allowed to move the placement of the mouse and joystick until they were comfortable.

Before transitioning to six testing trials, participants were randomly assigned to one of two tracking task difficulty conditions. The same tracking task difficulty was used for all test trials in MATB for each group. Difficulty of the tracking task was manipulated using bandwidth and update rates as were used in Exp. 1.

During each of the six 10 minute test trials, participants encountered six *phases* of Trk task, and Comm task operation. These phases were approximately 1 minute in duration for tracking, and 30s for communications task. Phases referred to the indications present in the task scheduler display, on which the beginning and end of a task was displayed with a bright green bar which occupied the duration time on the timeline on the vertical axis of the scheduler (see Fig. 5.7).

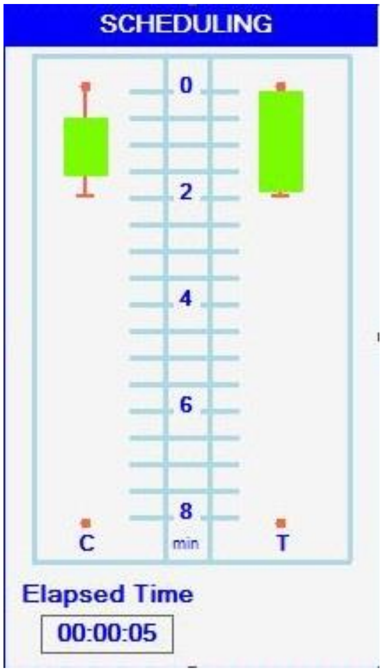


Figure 5.7. The task scheduler in MATB. Time advances from bottom to top. On the left side, in column "C" an upcoming Comm task period is shown (green block) when a Comm event could occur. On the right side, column "T", a tracking period is also shown. Red lines around both green events represent the surrounding total time. Based on this example, the tracking task will start before the communications task, the Comm task will end before the tracking task, and the trial has about 2 minutes left until it is over.

Participants could always view in real time the upcoming tasks and durations. Tracking and communication tasks were alternated, and did not overlap in these trials.

During each of the six tracking phases in each trial, participants were interrupted early (within 10s of beginning the task), midway (around the 30s mark, about halfway) or late (within 10s of the end of the tracking phase). For the state space analysis, tracking error was scored as low, medium or high by using the lower, mid, and upper tertiles for each trial of MATB. The rate of change was scored from measurement to measurement of the tracking task, occurring on a one second timer. An increase in error was scored as positive, and decreasing error scored as negative; in the event there was no change in error

In the communications task, participants were presented with a Comm task that included an interruption based on testing hypothesized switch resistance. For each Comm task phase, an interruption (either a Mon red light onset, Rman pump 1, or Rman pump 3 failure) occurred during the encoding of the callsign (early), encoding of the radio and frequency information (middle), or during the following period of expected WM rehearsal (late). These three periods were slightly different in terms of timing for each Comm event, because the Comm audio information presentation rates varied between events. Interrupting event timings and types were pseudo randomized within, and across trials. All pumps were fixed 30s later to avoid overlapping pump failures in the Rman task. As in Exp. 1 (the equal priority condition) participants were instructed to attempt to complete all tasks quickly and accurately, and no specific priority instructions were given.

Subjective paired task ratings as in Exp. 1 were collected after the first test trial performance, and then again after the final test trial (see Appendix F for the revised survey format used in Exp. 2 and Exp. 3). Following the completion of the final survey, participants completed a measure of individual differences in attention (the attentional network task, ANT; Fan et al., 2005). The experiment took approximately two hours to complete.

RESULTS

Programming errors resulted in the loss of data for four participants. Additionally, three participants had tracking error >3 standard deviations (*case 1* = 72.37; Group mean = 17.45; *case 2* = 63.43, *case 3* = 63.43; Group mean = 32.57), and their data was also discarded. All analyses were on the remaining 63 participants unless otherwise noted. Where sphericity has been violated based on Mauchly's *W* reaching significant at $p < .05$, Greenhouse-Geisser corrections were used.

Task Performance

As in Exp. 1, performance for the tracking task was examined to determine whether difficulty manipulations were successful, as well as the general observation (see Table 5.2) that participants were at least attempting to perform all of the tasks and whether any difficulty effects that were present.

Table 5.2. General performance outcomes for each task in Experiment 2.

	Trk Error	Rman Error	Comm Accuracy	Mon Accuracy
Easy	17.45 (4.3)	267.49 (171)	0.66 (.04)	0.95 (.26)
Difficult	32.57 (5.5)	320.52 (296)	0.60 (0.20)	0.92 (.34)

Note: Standard deviation is shown in parentheses.

Error in the tracking task was assessed in a 2 (difficulty) x 6 (trials) repeated measures ANOVA. There was a significant main effect of task difficulty, such that the easy tracking condition had less error (M = 17.45) than the difficult task (M = 32.57) overall ($F(1, 60) = 135.59$, $p < .001$, $\eta_p^2 = .70$). There was also a main effect of practice such that error decreased over time ($F(3.362, 201.707) = 32.83$, $p < .001$, $\eta_p^2 = .35$). A significant interaction (Fig. 5.8) suggested that the benefit of practice was greater for the difficult condition than for the easy tracking condition ($F(3.362, 201.707) = 6.60$, $p <.001$, $\eta_p^2 = .10$), and this trend is visibly confirmed in the first two trials.

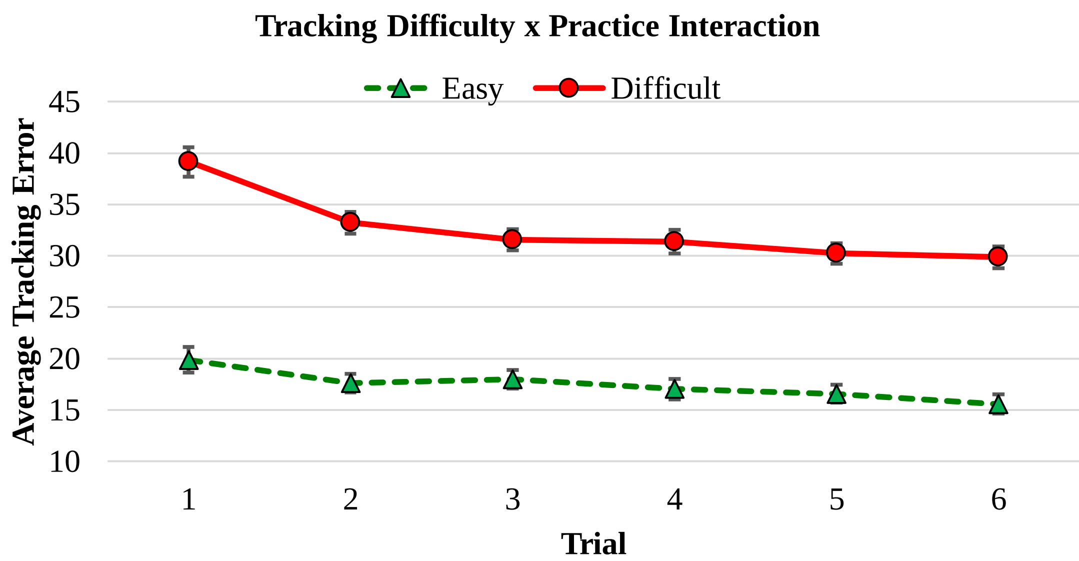


Figure 5.8. The average tracking RMSD for each condition (easy and difficult), across each of the practice trials. Error bars are standard error of the mean.

Therefore, in the tracking task the effect of difficulty and practice are both evident, and suggest the difficulty manipulation was effective.

Task Switching

Switches were tracked as in Exp. 1, with improved fidelity now offered from the newer version of the MATB paradigm. The primary hypotheses were whether the time on task effect (a general increase in switches away from an ongoing task) was present, as examined with various interruptions to the tracking and communications tasks. Further, the role of task difficulty should change the effects between easy and hard tracking as a function of time on task.

To answer these questions, a 2 (ongoing task; tracking or communications) x 2 (interrupting task; Rman or Mon) x 3 (time of interruption; early, middle or late) x 2 (difficulty) repeated measures ANOVA was conducted on the average sum of switches to interruptions across all test trials. Thus, despite the longer amount of time spent tracking, the same opportunities for a switch to an interrupting task were present (a total of 6 on average). From this initial analysis, the role of task difficulty did not emerge as significant in any capacity, and the analysis was re-run without it as a factor.

A significant main effect of ongoing task was found, with Tracking being more likely to be interrupted (M = 4.17) compared to Comm (M = 2.73; $F(1,59) = 108.75$, $p < .001$, $\eta_p^2 = .65$). A main effect of the type of interrupting task showed that Mon events were more powerful interrupters (M = 4.81) compared to Rman events (M = 2.08; $F(1, 59) = 561.51$, $p < .001$, $\eta_p^2 = .91$). Significantly fewer switches were made over time from early (M = 3.95) to middle (M = 3.45) to late (M = 2.94), regardless of OT or IT, a main effect for time of interruption ($F(1.855, 109.419) = 35.46$, $p < .001$, $\eta_p^2 = .38$).

Interactions between time and ongoing task ($F(1.694, 99.948) = 36.11$, $p < .001$, $\eta_p^2 = .38$), time and IT ($F(1.91, 112.70) = 35.90$, $p < .001$, $\eta_p^2 = .38$), and the interaction between OT and IT ($F(1,59) = 6.35$, $p = .014$, $\eta_p^2 = .10$) were all significant but superseded by a significant three way interaction between time of interruption, OT and IT type ($F(1.821, 107.414) = 37.50$, $p < .001$, $\eta_p^2 = .39$).

Perhaps the best way to showcase the interaction is to separate the interactions into those that occurred for the Rman interrupting task, and those that occurred for the Mon interrupting task (Fig. 5.9).

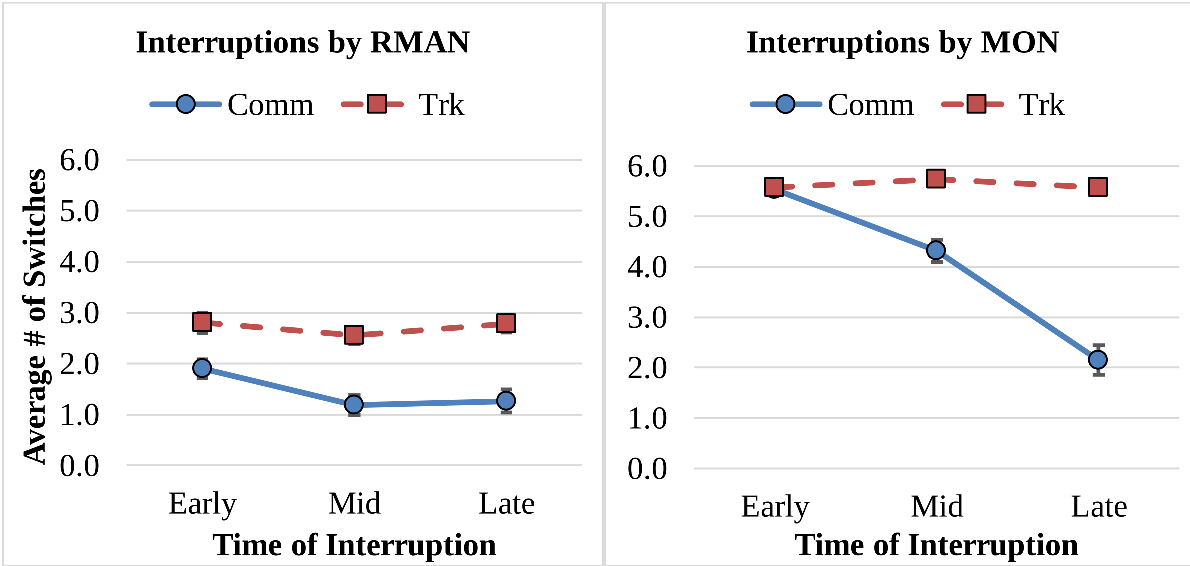


Figure 5.9. Average number of switches made to interrupting Rman or Mon events across time in the ongoing tasks, Comm or Trk. Error bars are standard error of the mean.

In general Trk was interrupted more than Comm, and the Mon event was more *interrupting* than the Rman event. While Mon events were almost always successful interrupting an ongoing Trk task (regardless of time on task), the same interruptions on the ongoing task of Comm showed a *decrease over time* in addition to being less interupting in general. This provides evidence for a rise in switch resistance over time in the Comm task, a contradiction to the Kurzban et al. perspective (rejecting *H1*), but in line with the WM demand / task inertia effects (confirming *H3* of an inertia effect in Comms, and rejecting *H4* of an expectancy effect in Trk).

A similar pattern was shown for the interrupting Rman tasks. Rman was more likely to interrupt an ongoing Trk task compared to Comm; and the likelihood for interruption appeared to decline slightly for the Comm task over time, but not for the Trk task. Once again evidence that switch resistance increases (decreasing the chance for a switch) over time due to an inertia effect (*H3* confirmation) for the Comm task was found. The fact that switches from tracking do not show an increase toward the end of a trial indicates that there is no end-expectancy influence, rejecting *H4.*

Tracking State Space Calculations. A novel component to this experiment was an examination of a tracking state space. Optimal switch times may exist in the tracking task, as related to error rate and error magnitude. To test this idea, the data were constrained to time periods in which participants were actively engaged in the tracking task (tracking blocks). Time spent *away* from the tracking task was not used in these analyses. From each of the test trials tracking error magnitude was categorized as low, medium or high based on tertiles, calculated for each participant on each trial. The rate of error change (increasing, decreasing, or not changing) was scored by comparing adjacent error magnitudes.

For the current assessment, only the first three trials of the difficult tracking condition data were examined. This was justified for two reasons, (a) the difficult category represents the best chance to find operators using the hypothesized points in the tracking task, compared to the easy condition; and (b) this should be readily apparent early in practice.

For the analyses, data were binned into two categories based on Figure 5.2 and Figure 5.3, in which switching would be most likely to occur. These were "optimal" (error is low and rate is decreasing or 0) or periods when error is high and either zero or increasing, and therefore switching is predicted to be less likely, "resistant". A paired two tailed t test showed that there were more switches away from the tracking task made during the optimal conditions (M = 7.85) compared to the resistant (M = 6.0) conditions; $t(38) = 2.38$, $p = .02$. The utility of the state space data was supported in predicting a task switch in this specific grouping, in line with a periodic effect of time on task, and confirming *H2*.

Practice Effects

Having assessed time-in-task effects, the notion of practice changing switch behaviors was explored. Practice-based switch frequency effects were assessed in a 2 (difficulty) x 6 (practice trial) repeated measures ANOVA on total switch counts (see Fig. 5.10). No effect of task difficulty was found ($F(1,59) = 1.50$, $p = .23$, $\eta_p^2 = .03$). A marginally significant effect of increasing switches over time was shown ($F(4.034, 237.988) = 2.03$, $p = .09$, $\eta_p^2 = .03$). No interaction was found between the two factors ($F<1$).

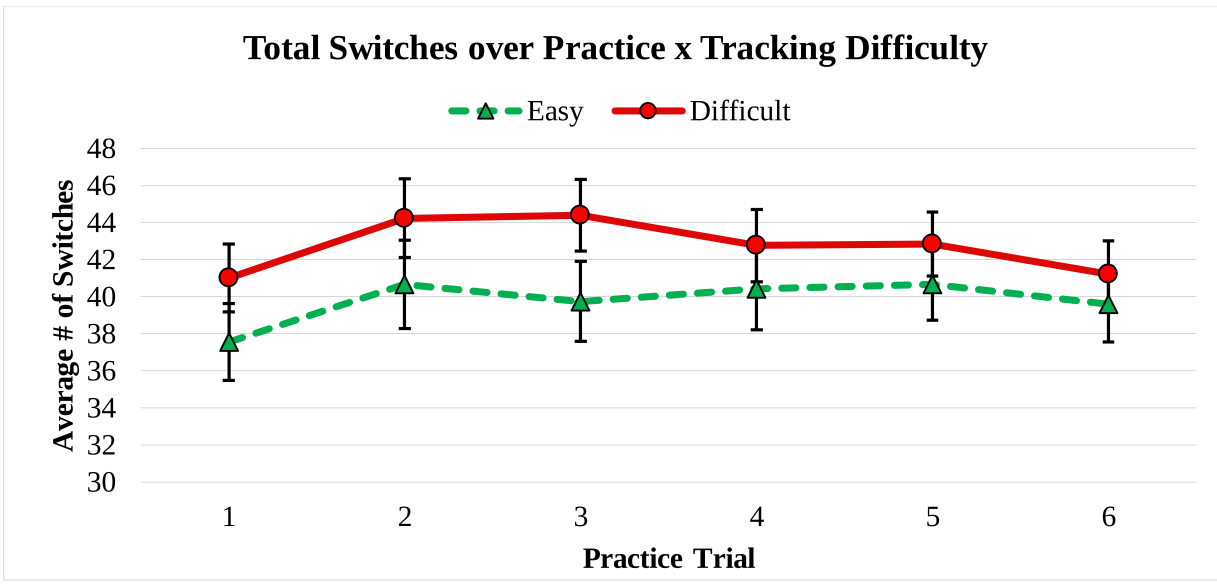


Figure 5.10. Number of switches plotted across practice trials by tracking difficulty condition. Error bars represent standard error of the mean.

Therefore, addressing *H5* (more switches over task practice), the results show that there was no effect of practice, nor did this interact with task difficulty. Interestingly, these results contradict those found in Experiment 1 in that no difference in switches was found between the two difficulty conditions and the trend was in the opposite direction; this may be a result of the between subjects manipulation.

Survey Data

Of substantial interest were what subjective properties were aligned with the switching patterns observed. Survey data from 13 participants were missing and not included here.

An overall score of task attractiveness was calculated combined across early and late surveys (see Table 5.3), as examination of change between ratings indicated relative stability (i.e., no large changes across time). As in Exp. 1, the interest was in the predictive utility of this table in determining the switch likelihoods for one task versus another, and characterizing this within the scope of the interruption effects.

Table 5.3. The results of the surveys in Exp. 2 pooled across time.

	PRIORITY	INTEREST	DIFFICULTY
MON	-0.87	-1.68	-1.48
RMAN	0.68	1.14	1.51
TRACK	-0.07	0.45	0.08
COMM	0.27	-0.79	-0.11

Note: More positive rankings = higher task attractiveness for Priority and Interest, but *lower* attractiveness for Difficulty.

Table 5.4 shows interrupting task (Mon and Rman) ratings that make them attractive (in purple) and ongoing task ratings (Trk and Comm) that make staying in the OT more attractive (in green). In general then, it appears that Rman interruptions would be attractive because of their priority and interest, though the difficulty might inhibit switches. In contrast, the Mon interruptions would be attractive *because of* their ease, and despite their lower priority and lower interest ratings.

Table 5.4. The outcome of the surveys in terms of task attractiveness.

		INTERRUPTING TASK	
		Rman	**Mon**
ONGOING TASK	**Comm**	Rman > C (less priority and interest); BUT Comm is easier	Comm > M (higher priority and interest); BUT Mon is easier
	Trk	Rman > T (less priority and interest); BUT Trk is easier	Trk > M (higher priority and interest); BUT Mon is easier

Note: Ongoing task attractions are represented in green. Interrupting task attractions are in purple.

A summarized interpretation of the interruption data can be found in Table 5.4. In the second column ("Dominance"), the dominantly chosen task is highlighted based on the results presented earlier in analyses of switching. In the next column ("Why"), the reasons (including individual attribute ratings from Table 5.3) why the dominant task might have been dominant are given. And in the final column, ("Despite"), reasons are given for why the unfavored task *might* have had increased attractiveness, but did not. The final column presents total attractiveness comparisons between the two tasks as estimated by collapsing across the values from Table 5.3.

Table 5.5. Dominance based on the interruption data in Exp. 2, along with reasons both for and against the findings, and each task's total attractiveness rating in the attraction column (bolded for the survey-based most attractive task).

Ongoing Task	Dominance	Why:	Despite:	Total attraction
COMM	**Mon** > Rman	Easier	Higher Interest, Priority	Comm: -.13, Mon: -.35, **Rman**: .10
TRK	**Mon** > Rman	Easier	Higher Interest, Priority	**Trk**: .15, Mon: -.35, Rman: .10

Note: Salience is not represented here for the interrupting tasks because it was assumed Mon and Rman were equally salient (visual, red light onset events).

Table 5.5 indicates the strong influence of task difficulty: in spite of the lower perceived task interest and priority, the Mon task consistently received more switches.

Individual Differences

The automated ANT program scored participant performance in the ANT test designed to measure executive control. Three measures comprise the outcome as related to three attentional networks: alerting, orienting and conflict effects. Of interest was the conflict outcome. Analyses were limited to testing whether a relationship between switching frequency and the measure of conflict existed outside of the confined experimental procedure typically used in switching studies (see Chapter 2; i.e., using rules instead of true tasks and restricting true task choice). I predicted (*H6*) fewer switches for participants with higher scores on the control measure.

Pearson correlation coefficients were calculated between executive control and average switch frequency for 49 participants, due to missing ANT data. One outlier that exceeded 3 SD from the mean was also discarded. Scores on the executive conflict measure overall ($M = 139$, $SD = 42$) appear to be in the range of other values reported in Fan et al. (2002), Fan et al. (2005) and Arrington and Yates (2009). No significant correlation was shown between the control measure of ANT and switching overall ($r = -.22$, $p = .13$).

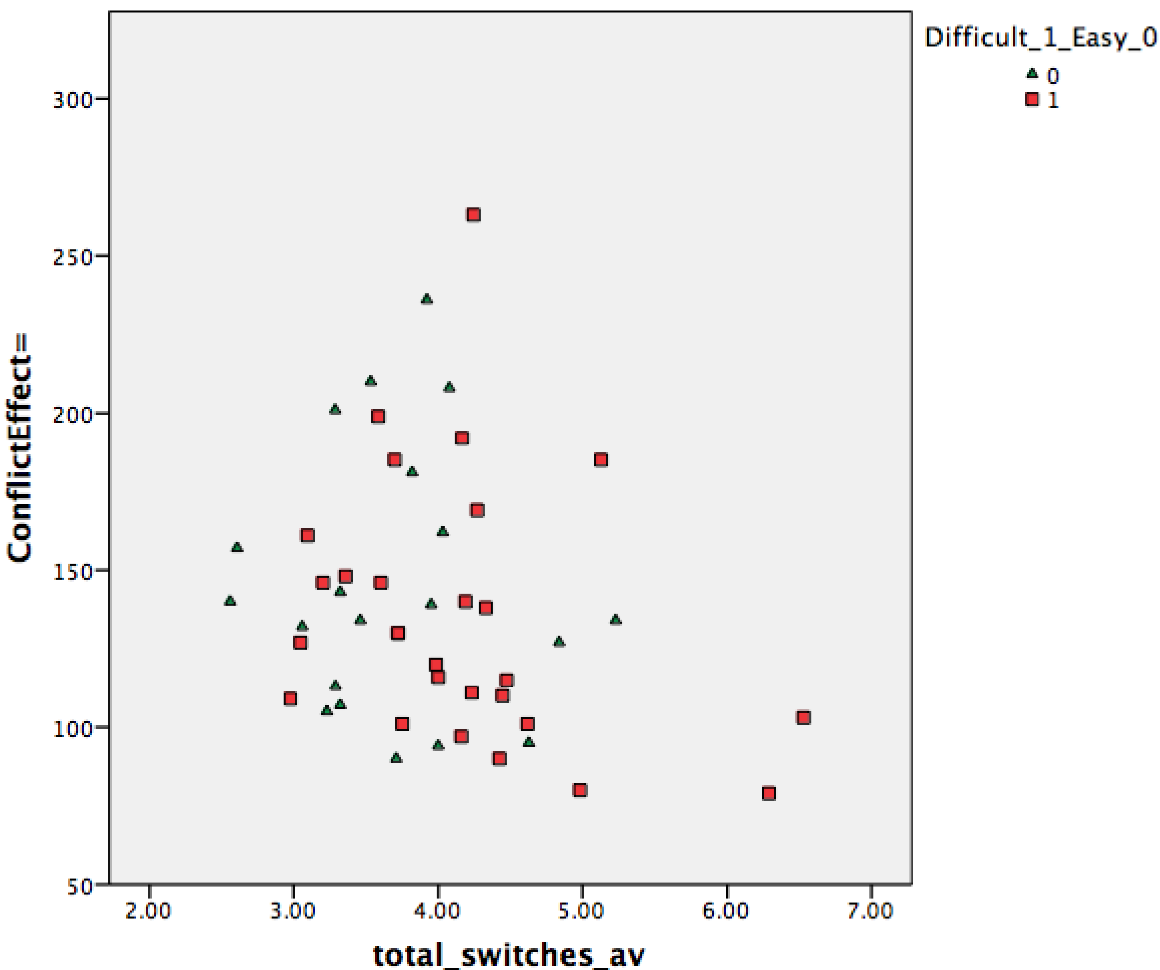


Figure 5.11. Scatterplot between total average number of switches across all trials and the executive control "conflict" outcome from the ANT task. Green triangles are easy tracking trials, and red squares are difficult tracking trials.

However, the relationship could differ by group condition (see Fig. 5.11): tracking difficulty could bring out the effect versus easy (as backed up by prior examples in the literature). Therefore the correlation was recalculated on the same data but split for difficulty condition. In the easy condition (r = -.06, p = .82) and the difficult condition (r = -.29, p = .14) no significant relationship was shown. Given the moderate power of these examinations, a strong conclusion is not advised, however, the current data do not support the relationship, and the size of the current effect suggests at most the relationship is weak between the executive control measures and switching frequency (e.g., Arrington & Yates, 2009). Certainly this remains a topic for future examination, but as it appears, the relationship does not hold for more applied task situations when switching rules are not invoked.

Experiment 2 Discussion

In general, the main purpose of Exp. 2 was to test and examine time-on-task effects on switching. The main hypothesis (*H1*) suggested that switches should in general increase over time (Kurzban et al., 2013). This increase in switching was suggested to be even greater in the presence of high-value alternative tasks. In the frame of STOM and the results of the surveys, high value most closely relates to high priority, and so between the two interrupting tasks, resource management was perceived as more valuable in comparison to monitoring. Not only was this main prediction *not supported*, but the secondary component of the theory failed as well. In fact, more switching was found in general to the less valuable (but *easier*) task of monitoring. And furthermore in the communications task, a decrease of switches to monitoring events was observed toward the endpoint of the task due to inertia (from WM loading, confirming *H3*). A side benefit to the subjective rating data was that it was possible to confirm that participants viewed the two ongoing tasks as approximately the same difficulty level (because prior work shows interruptions during difficult tasks are worse for goal maintenance than during easier tasks; Altmann & Trafton, 2002). Therefore these results are at least *not* confounded by OT task difficulty.

In the state space data (a unique contribution of this experiment to task switching research), switching from tracking was indeed observed more frequently if the tracking error rate was decreasing or stable, and error was low, confirming *H2*. As in other work (Brumby et al., 2007; Janssen, Brumby, & Garnett, 2012; Janssen et al., 2010) on task switching choices, some amount of optimality may be inherent in these types of decisions, especially at subtask boundaries. As the analysis focused on a “best-case” subset of the total data, further analysis will be undertaken to determine whether the same relationship exists when tracking is less difficult. However the potential utility of the state space notion may extend to other realms of multi-tasking under high demands; specifically in controlling unmanned robots during search and rescue, a task often performed with low automation and still using joystick based control (e.g., Burke, Murphy, Coovert, & Riddle, 2004; Murphy, 2004), interrupting the operator should only be done at ‘opportune’ times to avoid crashing the robot.

Finally, it was not confirmed that expectancy effects operated on the tracking task (H4). There was no increase in switch resistance toward the end of tracking tasks, despite a salient visual indicator for the task endpoint in the scheduling display. It may be the case that such expectancy effects only manifest for cognitive tasks, as related to increasing cognitive investment that may be “lost” when a switch occurs. This was not the case for the tracking epochs.

Although one a priori assumption was that switching may generally increase over task practice, such an effect was not found under either easy, or difficult tracking. It is perhaps not surprising that switching on average does not radically change over time. If switching frequency is influenced by task switch costs, and task switch costs do not decline significantly or disappear with even extensive practice (e.g., Rogers & Monsell, 1995; Stoet & Snyder, 2007) this should be expected. However it was tested in part because the prior work, as discussed in Chapters 1 and 2, have not examined task switch choice. It should be noted that the current examination look at switching across *all* tasks; it is possible that switching to specific tasks could vary across time and should be addressed in future analyses.

The results of the individual differences examination suggested that the power of an executive control component of attention does not greatly influence task management behavior in the current experiment. This lends support to the claims by Arrington and others that one reason such a relationship existed in basic, voluntary task switching experiments is due to instruction-level constriction of free switch choice. More evidence should be collected to further solidify this initial finding. Specifically, within the current paradigm, instructions for switching frequency (subject should attempt random switching between tasks) and overall switch makeup (switches should be even across the tasks) would have to be instructed as in Arrington and Yates, in comparison to a version of this experiment without restrictions on switching.

Combined, evidence from this experiment suggests that some aspects of the current theoretical approach to task switch choice may not address the complexities of the outcomes, and that the relationship between time on task and switch likelihood may be monotonic (generally decreasing with time), however, some periodic effects also exert control (as shown by the use of a state space for tracking). It is clear that the Kurzban et al. (2013) theory does not garner support from these results. One categorically related hypothesis that has some merit but little direct support, which is diminishing returns, was discussed at the beginning of Chapter 5. Diminishing returns falls into the category of monotonic effects of time on task. The main prediction is a decrease in switch resistance, but specifically as a result of a reduction in the rate of return on attention investment on a task over time. Existing experimental evidence for static rate of return on switching is addressed, and the explanation is experimentally examined in Chapter 6 in order to further determine time-based effects on switch choice.

CHAPTER 6: EXPERIMENT 3

In experiment 3, further examination about a prediction of switch resistance as related to the time on a task was undertaken. Specifically, a set of hypotheses were derived from the notion that in some circumstances diminished returns over time of performance a task may influence the likelihood to switch away from a task.

In Exp. 2, the main focus was on determining whether theories of time-on-task dependent switching likelihoods held out in both the communications and tracking task. However, one of the explanations briefly covered in Chapter 5, the notion of diminishing returns for performing a task driving an operator to switch away, was not tested in Exp. 2. Anecdotally, rate of returns has been tied to foraging behaviors. One might imagine going apple picking, and the amount of expected return in apples may dictate which area of a forest you choose to visit initially. When a sufficient number of apples have been picked to *decrease* the rate of return – all of the ‘low hanging fruit’ is gone - one might be expected to move on to another, richer area to pick apples.

All fruit aside, Experiment 3 was intended to psychologically operationalize diminishing returns, and then test the theory as it relates to a time-dependent switch explanation.

Returns On Performance

Two specific hypotheses, as derived from Payne et al. (2007) and Duggan et al. (2013) suggest people switch away from ongoing tasks for one of two reasons. First, a lack of return on investment in the task can drive operators to abandon a task. Second, a completion of a goal or sub goal in the task can signal a relevant switching place, as in the “optimal” switch points in the tracking task (Ch. 5).

Addressing this first explanation, the value of performing a task, from an economic perspective, may influence the decision to stay or to switch. While the explanation of subtask boundaries has been addressed (Janssen, Brumby, & Garnett, 2012; Janssen & Brumby, 2010; Salvucci, Taatgen, & Kushleyeva, 2006) , little if any existing literature addresses *diminishing* returns over time (as in foraging for a non-replenishing commodity) as an explanation for time-dependent switch effects. Payne et al. and Duggan et al. have *only* shown that a difference in static rate of return influences task choice. They do not address the idea of diminished value as a time-dependent factor related to switching.

Kurzban et al. (2013) suggested that subjective feeling of increasing effort over time is an internal signal, an indicator of the increasing expected value of performing an alternative task (and hence switching from an ongoing one). Hence, theirs is an expected value account that parallels the concept of diminishing returns over time. This was offered as a contrary explanation to an “effort depletion” view of Baumeister and colleagues. However, Kurzban et al. do not explicitly manipulate this factor.

Thus there are no experimental data to support diminishing returns as an explanation of switch behavior, especially in the unique case of multi-task management under load, where the STOM model operates. In Exp. 3 this paucity is rectified by decreasing the rate of return of performing an ongoing task by changing (reducing) its event frequency (monitoring), or increasing its difficulty (tracking), as a task episode progresses.

Experiment 3 Hypothesis

It was predicted that switch resistance would decline in proportion to the decreasing rate of return in a task. Thus, (*H1*) switch resistance should be high initially when payoffs are high, and there should be fewer switches away during this period, compared to later as participants receive diminishing returns for allocating attention to an ongoing diminishing returns task. Additionally, this general effect of switch resistance decreasing over time could result in differences between task switches away from the ongoing tasks in the constant returns trials, as was shown in Exp. 2.

METHOD

Participants

Seventy-one students at Colorado State University participated for either optional, partial course credit, or were recruited as paid participants and received $10 per hour. Paid participants performance was not incentivized.

Materials

The same computer setups as in Exp. 1 and Exp. 2 were used in Exp. 3. The MATB simulation version from Exp. 2 was used. Instructions on the MATB task were presented via a PowerPoint (as in Exp. 2). Two tasks (monitoring and tracking) were chosen for manipulation based on creating the clearest cases for subjectively recognizable decreasing value of spending time in a task, in addition to other properties described below.

T**he tracking task** is also continuous, and could be altered in its difficulty (tracking difficulty served as a proxy for diminishing returns) on a task. In other words, as tracking gets harder over time in a trial, the rate of return (amount of effect on a task's performance) for the same level of resource investment *decreases* over that time (returns are based on keeping error low, a goal for the task). Therefore, more switches should be seen as returns diminish. The diminishing returns trials are contrast to a constant rate of return trial, in which difficulty of the task does not change over time, and return on resource investment does not decrease over time in the trial.

The **monitoring task** was chosen because it was both discrete, in contrast to the continuous tracking task, and its event frequency could be altered easily and clearly across time in the task. Event frequency in monitoring served as a proxy for diminishing returns – when the frequency of monitoring events is initially very high, resources focused on the task provide a high rate of return. However, if the frequency rate *declines over time*, (as progressively more "apples are picked"), then there was a diminished return for the same level of resource investment in the task over time. Thus, with declining rewards/rate of return on attentional investment, switch resistance would be expected to decrease – and this may represent part of the time on task effect.

In experiment 3, in a within-subject design, each participant was exposed to two variants of the MATB task within subjects. One variant had a constant rate of return over time for both of the target tasks (thus, constant difficulty in tracking periods, and constant rate of monitoring periods). The other was a variant in which these rates decline significantly over time (described below). These declining rates trials help operationalize diminishing returns.

Constant Returns Trials. Trials during this block contained the same number of overall events as the diminishing returns trials; however, monitoring events were spread out approximately evenly across the 5 minutes per trial (mean event occurrence = 1 event /15.43 s, *SD* = 1 event / 3.5 s). Tracking difficulty was set to the "difficult" level such that the rate of return would be constant across the trials.

Diminishing Returns Trials. Two types of diminishing returns (DR) trials were constructed. For the **monitoring task, t**he rate of potential for return was manipulated from extreme high, to extreme low within the allotted 5 minutes of a trial (see Fig. 6.1). A very high frequency of presentation was used early with less than or equal to 4s of time gap between events

(*M* = 2.49, *SD* = .80). This frequency declined over time to low frequency (event 1: *M* = 51.67, *SD* = 5.84; event 2: *M* = 68.83, *SD* = 8.08) starting after one minute of trial performance. In the last time period of the trials, frequency was reduced to the lowest rate, and only a single monitoring event occurring within the last two minutes of the trial (*M* = 118.60, *SD* = 6.59).

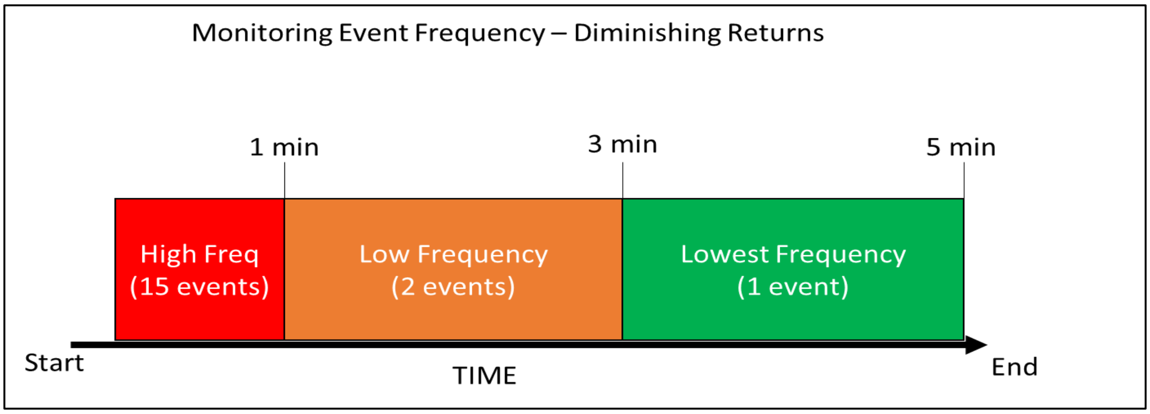


Figure 6.1. Monitoring event frequencies for the diminishing returns trials **(DR-Mon**) in which monitoring frequency creates degradation in returns over time between high, low, and the lowest frequency.

In contrast to Exp. 2 but similar to Exp. 1, Mon events were any of the scale or light events, randomly alternated but in equal numbers overall on each trial. Events in the Mon task were 'jittered' slightly in their temporal location trial by trial to avoid participant strategy development related to knowing the sequence of events that might occur.

Difficulty in the tracking task was varied trial to trial in the diminishing returns for tracking trials (DR-Trk), by blocking tracking into 30s periods across the 5 minutes available in a trial (Fig 6.2).

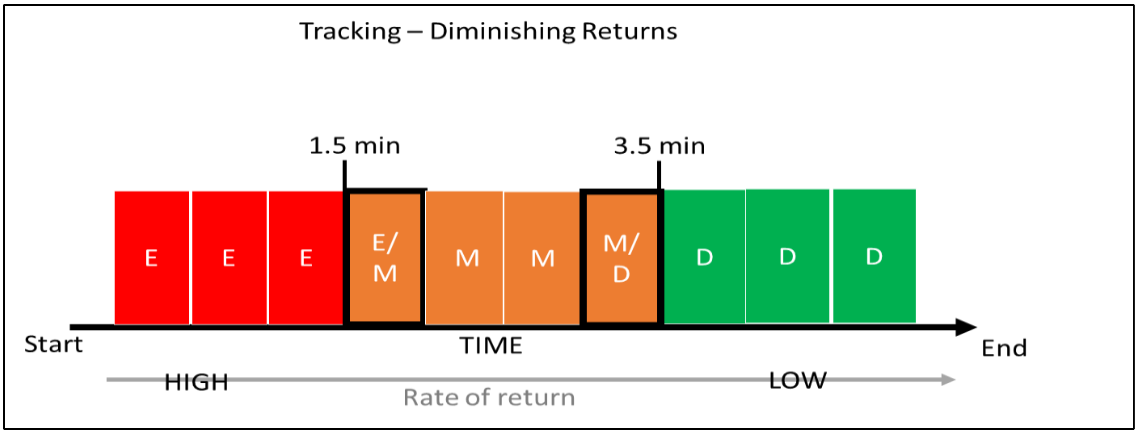


Figure 6.2. The use of increasing tracking difficulty (red to green) to create diminishing returns in the tracking task diminishing returns trials (**DR-Trk**). Under easy condition periods (individual blocks marked E), rate of return should be high, and under medium (M) and difficult (D), rate of return should decrease and become low. E/M and M/D represent time periods that varied between easy or difficulty in the scripting of the total DR-Trk trials.

The first three periods were always easy tracking difficulty. Period four varied between trials (see bold outlines) between easy or medium difficulty, and period seven varied between medium and high difficulty (based on whether period 4 was easy or medium difficulty, respectively) in order to create approximately even periods across the difficulty levels. Periods five and six were always medium difficulty, and the remaining periods of tracking (8-10) were always high tracking difficulty.

Events in other alternative tasks (Rman and Comm) varied trial by trial to avoid the learning of a predictive component for switch choice as participants perform several trials in a row, however, their occurrence was approximately evenly distributed across each trial, and the same amount of each event type was used in each. Six Comm ownship events occurred in each trial for both constant and diminishing returns trials. Six Rman events also occurred and reflected pump failures of two of the high priority pumps (pump 1 and 3, as in Exp. 2). Pump failures were always repaired within 30s.

Procedure

As in Exp. 2, participants were instructed via a PowerPoint on aspects of the MATB simulation (the same instructions used in Exp. 2) following informed consent procedures. Participants then received the same instructions about restricting responses to a single, dominant hand, and were allowed to position the joystick and mouse in comfortable positions. No dual tasking was allowed at any point, and participants were not allowed to switch hands.

Participants then completed the training trial in MATB (same as Exp. 1 and 2). Following completion of the training trial, participants were split into two groups based on random assignment to either tracking, or monitoring task for diminishing returns trials. Both groups completed two main trial blocks, one for *constant returns* and one for *diminishing returns*.

The *constant returns* block contained MATB trials with an evenly distributed set of monitoring events, and tracking task difficulty was set to difficult and did not vary. Participants completed six trials (5mins each), followed by a subjective paired task survey at the end of the constant returns block. In the *diminishing returns* block, participants completed twelve trials (5mins each) in which <u>either</u> the tracking or the monitoring task – depending on randomly assigned condition - exhibited diminished returns. Following the diminishing returns block, participants completed a subjective rating survey. Blocks for constant and diminished returns trials were counterbalanced. The experiment took approximately 2 hours to complete.

RESULTS

Due to programming errors and missing data, only 41 participants' data were included for these analyses unless otherwise noted. Where sphericity has been violated in ANOVAs, based on Mauchly's W reaching significance at $p < .05$, Greenhouse-Geisser corrections were used.

Task Performance

As in Exp. 2, the main reason to examine task performance was to determine whether tasks were indeed performed. As practice in the trials is not the focus, such analyses are not performed here. Table 6.1 provides descriptive statistics on the key performance outcomes for the two key conditions of diminished returns – tracking, or monitoring.

Table 6.1. Means of the key performance metrics in the MATB simulation, across the two conditions (**DR-Mon** = Diminishing return in the Mon task; **DR-Trk** = Diminishing returns in the Trk task).

Condition	Trial Type	**Trk Error**	**Rman Error**	**Comm Acc**	**Mon Acc**
DR-Monitoring	Constant	37.4 (7.5)	246.0 (132.2)	0.94 (.08)	0.61 (.16)
	Diminishing	35.6 (6.0)	224.3 (122.1)	0.85 (.20)	0.59 (.13)
DR-Tracking	Constant	35.4 (8.0)	292.6 (316.3)	0.89 (.06)	0.61 (.21)
	Diminishing	26.1 (5.5)	212.0 (170.4)	0.92 (.04)	0.66 (.18)

Note: Both conditions completed the Constant return trials. Standard deviations are provided in parentheses.

From Table 6.1 it is clear that performance was not neglected toward any task in particular in either block for either condition. Therefore it is safe to assume that patterns of switching found are not due to complete neglect of any task.

Task Switching

The main hypothesis of Exp. 3 was that in the diminishing returns trials, switches away from an ongoing task should increase over time. To assess this, switches were counted within bins of time. These time bins differed between the diminishing Trk and Mon trials, a function of tracking difficulty period jitter for Trk diminished returns trial. In the constant trials, the average time spent in easy, medium, and difficult tracking under the diminishing returns trials was used to bin time. In the DR-Mon trials, time was binned into three time periods. Within each time block, the number of switches for both constant and diminishing returns was divided by the number of events present for the trial, respectively, and used in calculation of the rate of switching.

A 3 (time period; early, middle or late) x 2 (returns; constant, or diminished) repeated measures ANOVAs was run on the data for switching from the tracking task Trk trials. Figure 6.3 presents the data from the DR-tracking group.

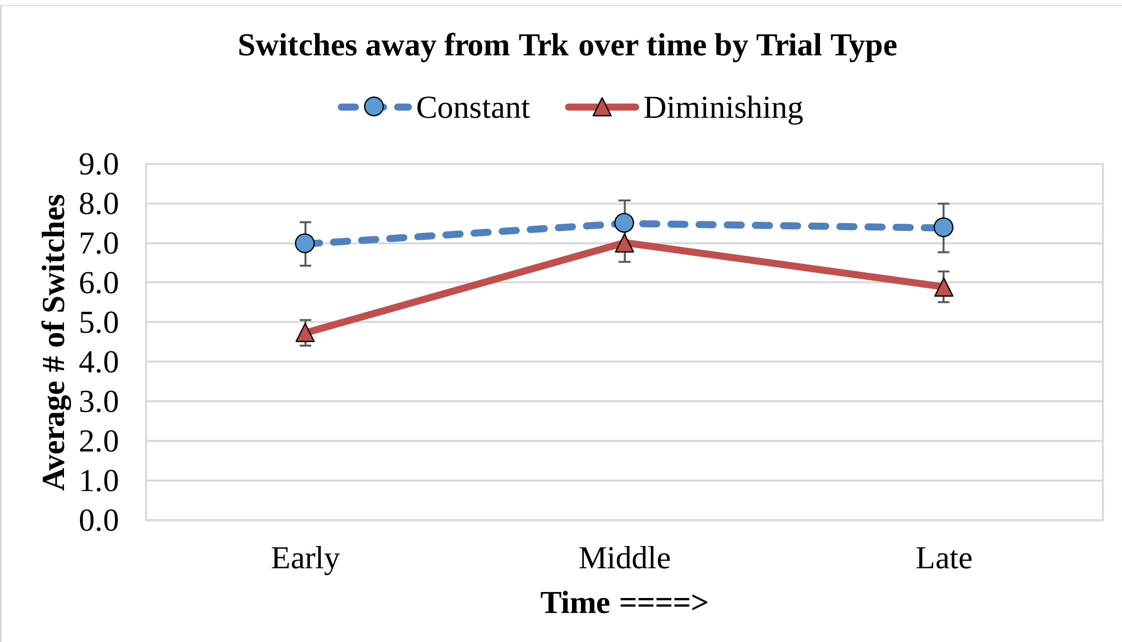


Figure 6.3. The average number of switches away from the tracking task between constant, and diminishing returns (DR-Trk) trials across time. Error bars represent one standard error of the mean.

The data presented in Figure 6.3 yielded a main effect of trial type, with fewer switches on diminishing trials (M= 5.88) than on constant trials (M= 7.29: $F(1,18)$= 17.78, p< .01, η_p^2= .50). They also revealed a main effect of changing switch frequency over time ($F(1.539, 27.711)$ = 23.62, p< .001, η_p^2= .57). Importantly, the analysis revealed a significant interaction between trial type and time ($F(1.656, 29.804)$ = 12.12, p< .001, η_p^2= .40). However, this interaction was not monotonic in form as predicted by the hypothesis, with a monotonic increase in switch frequency over time for the diminishing returns group. Instead, though the pattern is shown from early to middle, and early to late, the switch frequency actually appears to decrease from middle to late.

In order to understand this trend in the DR condition, two comparisons were evaluated: between early and late, and between middle and late. The first comparison did indeed reveal a significant increase, suggesting that with or without the middle data point, switching increased after the first period ($t(18) = -3.31$, $p < .01$). The second comparison was also significant ($t(18) = 5.97$, $p < .001$), indicating that decreased, a possible task inertia or end-expectancy effect.

An additional analysis of simple main effects confirmed that there was indeed a small, but significant over-time switch increase in switching in the constant returns trials ($F(2, 36) = 3.97$, $p = .03$, $\eta_p^2 = .18$). As above, an early-late comparison revealed more switches later in performance (M=7.39) compared to early (M=6.98; $t(18) = -2.16$, $p = .04$), as compared to the more pronounced effect of about a 1.5 switch increase in the diminishing return trials. The middle-late comparison was not significant, however and no differences were present between middle (M= 7.5) and late (M= 7.39).

In a second analysis, switches away from the Mon task were assessed in a 3 (time period) x 2 (constant, or diminishing returns) repeated measures ANOVA. These data are presented in Figure 6.4.

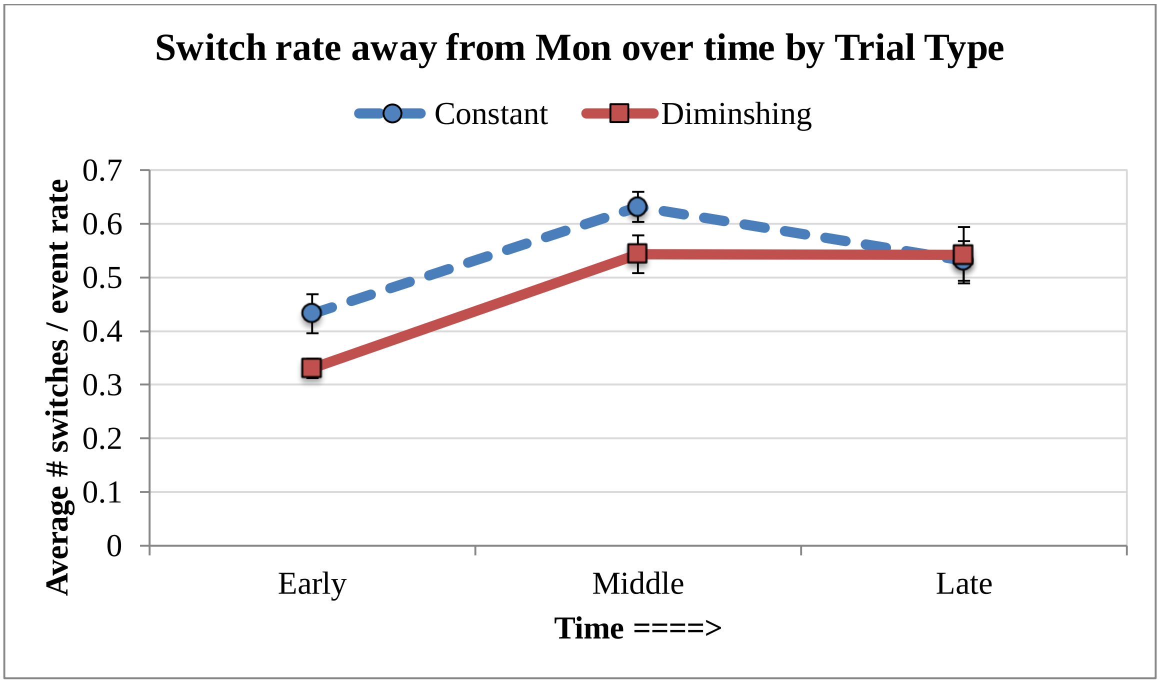


Figure 6.4. The average number of switches away from the Mon task, corrected for event rate, per period. Constant and diminishing returns trials for the DR-Mon condition are across time. Error bars represent one standard error of the mean.

Again a main effect for trial type indicated fewer switches away from Mon in the diminishing returns trials (*M rate*= 0.49) than the constant trials (*M rate*= 0.56; $F(1,21) = 7.69$, $p = .01$, $\eta_p^2 = .27$). Over time, switches away from monitoring increased ($F(2,42) = 49.22$, $p < .001$, $\eta_p^2 = .70$) as shown in Figure 6.4. The interaction between group and time only approached significance ($F(2,42) = 2.93$, $p = .07$, $\eta_p^2 = .12$).

Two assessments were made on diminishing returns trials to determine the locus of the incline effect. A comparison of early to late times showed that there was indeed more switching late ($M = .54$) compared to early ($M = .33$; $t(21) = -4.14$, $p <.001$). A comparison between middle and late times showed no difference between middle ($M = .544$) and late ($M = .542$) times ($t(21) = .05$, $p = .96$). Two further assessments were made to determine the locus of an increase effect in the constant returns trials ($F(2,42) = 39.05$, $p < .001$, $\eta_p^2 = .65$). A comparison of early to late showed an increase in switching between early ($M = .43$) and late periods ($M = .53$; $t(21) = -$

5.49, $p < .001$). A comparison of middle to late times, however, revealed a significant decrease in switching over time from middle ($M = .63$) to late ($M = .53$; $t(21) = 4.21$, $p < .001$). The results suggested that early in performance, switches proportionally increased with similar trends for both conditions. Diminished returns trials 'leveled off', but constant reliability trials appeared to have increased switch resistance (decreased switch frequency) during this later period. The non-significant interaction is thus caused by the slightly different trends on this middle-late transition period.

Survey Data

Surveys were administered following practice in the constant returns trials, and after practice in the diminishing returns trials. Two additional participants data were removed for failure to complete surveys.

As in Exp. 2, for the communications tasks, participants were asked to rate attributes only for ownship events. This was designed both to provide still more data for task attributes as well as to establish if the diminishing vs constant aspects of the tasks had any substantial effect on the attribute ratings, as shown in Table 6.2.

Table 6.2. The survey results for the main factors of expected change, collapsed across measurement time for each block.

	Priority		Interest	
Trial Block	**Mon**	**Trk**	**Mon**	**Trk**
Constant	-1.01	0.09	-1.57	0.70
DR-Mon	-0.92	-0.03	-1.49	0.53
DR-Trk	-1.16	0.21	-1.64	0.84

Note: **DR-Mon** = diminishing Mon task returns; **DR-Trk** = diminishing Trk task returns. Greater (positive) values indicate higher task attractiveness.

The survey data do not suggest any differences between priority and interest attributes, and so the data were collapsed into total attractiveness for each task under each trial condition (Table 6.3).

Table 6.3. The total task attractiveness for each of the tasks by the performing condition, collapsed across time of survey.

Trial Block	**Mon**	**Trk**	**Rman**	**Com**
Constant	-0.47	0.07	0.14	**0.27**
DR-Mon	-0.44	-0.04	**0.29**	0.19
DR-Trk	-0.52	0.16	-0.03	**0.40**

Note: Higher (more positive) values indicate greater task attractiveness.

From the above table, it is clear that perceptions of task attractiveness also do not differ as a whole between the conditions of performance examined in Exp. 3.

Experiment 3 Discussion

In Exp. 3, the effect of diminishing returns for performing a task over time and its effect on switching behavior was empirically tested. In the tracking task, switching increased between early and late periods, support for the role of diminishing returns in time-on-task switching, an important effect that suggests dynamic economic value (in terms of return on attention allocation) can drive task switching in applied multi-task management situations. Such a result adds a novel finding to the current understanding of static rate of return on investment effects for task switching (Duggan et al., 2013; Payne et al., 2007).

However, switching decreased significantly from middle to late periods for the diminishing returns tracking trials. It could be that this reflects a task inertia effect, but this is unlikely for two reasons. One, in Exp. 2 the end-expectancy effect was not found for the tracking task. And two, unlike in Exp. 2 where this explanation was potential, given that participants could use the task scheduler to predict the relative ending of tracking tasks, in Exp. 3, the only salient end of tracking was the end of the experimental trials. Furthermore, any indication of the end of a tracking trial (e.g., its learned approximate duration, after several trails) was the same for the constant returns trials (although to the extent it was *learnable*, more trials were completed for the DR trials). Another clear difference between DR and CR trials in the tracking diminishing returns condition was the increasing difficulty. In Exp. 1 (Ch. 4), more time was spent in tracking when tracking was difficult, so perhaps participants are simply allotting more time to it over the course of the trials and, as a result, switch less as the difficulty increases (also supported in Exp. 1 overall).

The results for diminished returns in the Monitoring task show a somewhat different pattern. Switch rate increases across time for diminishing returns trials, in line with the predicted diminishing returns effect. However, the *constant* returns trials, intended to be the control condition for the diminishing returns manipulation, *also* increased in the first time epoch, and then significantly declined in switch frequency between middle and late epochs. The non-significant interaction between time and trial type hint that increases in switching is more pronounced in the diminishing returns than in the constant returns trials, consistent with the general hypothesis. This set of effects may be viewed as three time-dependent influences on switching.

First, for both diminishing and constant trials, increased switching over time may reflect increasing boredom with the monitoring task (which rates low on task interest). The same pattern would be observed with the alternative explanations of resource-depletion, or with opportunity cost build up. It appears that at least the general phenomena described by Kurzban et al. (2013) were observed, independently of the diminishing return manipulation.

Secondly, this general trend does not continue further than the middle of the time periods used, and while the trend from middle to late is abolished in the diminishing trials, it actually reverses for the constant trials. It could be that this reflects some global task inertia effect (preserving or increasing switch resistance), which combined with diminishing returns results in no change for the diminishing returns trials, while in the constant trials may override the general monotonic increase in switch likelihood.

While the monitoring task itself is unlikely to build inertia as defined here (Fig. 5.1), there is a general inertia that may build as the end of the trial approaches – this is possible because participants are able to view a trial timer and, having learned the average trial length may also come to expect the end of the trial period at approximately 5 minutes. The lack of change in the second to third segment for diminishing returns trials, offsetting the possible inertia effect, thus reflects the diminishing payoff structure and reduced economic value that was intended by the manipulation.

The role of task feedback may also help explain why the results differ between the monitoring and tracking tasks despite manipulation of the same underlying, diminishing returns concept. It may be that the difference between groups during the periods of lowest rate of return reflects the observability of feedback on performance (essentially, see that the number of low-hanging apples is decreasing, and it might be time to move on). In the tracking task, the feedback is clear because the reticle would slowly become harder to keep on target over time. In the monitoring task, the feedback is the same no matter the particular diminishing returns condition (responding to a light turns it off or on, and responding to a scale resets it). Feedback is present in both tasks, but to slightly different degrees.

The observed effects, in sum, appear to present the multiple actions of several of the theoretical influences on time-on-task switching, with clear evidence for diminishing returns increasing switching over time, and some evidence for task inertia and expectancy effects mediating this relationship at the end of trial performance. One of the difficulties in parsing out which are occurring is that it is unclear for monotonic effects (opportunity costs, diminishing returns) what time frame they require to be shown; whereas for periodic effects (inertia, expectancy, and task stabilization from Exp. 2) all have a definable time period associated with them. This issue is revisited in Chapter 7 as it applies to the next steps in research.

CHAPTER 7: GENERAL DISCUSSION, IMPLICATIONS AND FUTURE WORK

The topic of task switching in general is not a new research domain, with roots traceable to the early 20th century (Jersild, 1929). Though the existing research on task switching is extensive (e.g., see Kiesel et al., 2010; Monsell, 2003), it was not easily accessible and useful when attempting to scale to the "real world". While the discovery that switches result in significant costs is important (e.g., Monsell & Driver, 2000; Monsell, Sumner, & Waters, 2003; Pashler, 1994; 2000; Rubinstein, Meyer, & Evans, 2001), they were generally made in absence of participant choice, and may not be particularly important in all applied domains. Contrary to a lab environment where choice may be constricted as a way to reduce extraneous variance and maximize the number of switches that occur, real world tasks tend to involve a great degree of operator choice freedom. Whether switching between email and word processing in the office, or between navigation and aviation in the cockpit, operators are usually given the ability to freely choose which task they perform (usually more than two!), in part because designs are (and should be) human-centric. As the role of technology and its influx complicates and rapidly changes the work environment and the cockpit, the role of task management becomes more essential. Without understanding how operators make decisions on a level that includes abandoning an ongoing task for an alternative one, it could be difficult to successfully design and incorporate displays and controls, and further to ensure they are utilized properly.

A wealth of literature has begun to investigate the critical choice involved in task switching, from basic (Arrington & Yates, 2009; Arrington & Logan, 2005; Arrington & Rhodes, 2010; Arrington, 2008; Demanet, Liefooghe, & Verbruggen, 2011; Demanet & Liefooghe, 2013; Lien & Ruthruff, 2008; Vandamme, Szmalec, Liefooghe, & Vandierendonck, 2010; Weywadt & Butler, 2013) to more applied contexts (Brumby, Rosario, & Janssen, 2010; Duggan, Johnson, & Sørli, 2013; Janssen, Brumby, & Garnett, 2012; Janssen & Brumby, 2010; Payne, Duggan, & Neth, 2007; Spink, Park, & Koshman, 2006; Wickens, Santamaria, & Sebok, 2013). Despite the amount of prior work, it was a surprising conclusion from the literature review (Ch. 2) that the applied relevance and scalability of studies using the Voluntary Task Switching paradigm was limited. Since part of the goal for understanding task switching choice was to help build and inform a model of task switching under high cognitive load (Wickens et al., 2013; Wickens et al., under review; also see Ch. 3), determining what task attributes and switching behaviors are exhibited in real-world tasks was a priority.

Benefits of a Literature Review

The literature revealed two strong predictions, based on the combined voluntary task switching and applied findings. First, in general, there is a switch avoidance effect, and in the context of the model this results in operators being less likely at any given moment to switch away from a task, than to stick with it (a ratio of about 60:40; Wickens et al., under review; Ch. 2 & 3). Second, and most related to switch choice itself, is that when a switch is chosen, the task most likely to be chosen to will be easy, rather than difficult (a ratio of about 63:37). Combined, these findings represent progress in understanding and modeling what an operator might choose to do. However, other factors play a role in task switching choices. Intuitively, and at times based on prior work, several factors of task attributes were deemed important for further study, including task priority, interest, and the general role of task salience. These types of factors, while shown in isolation throughout a limited literature, have never been combined to current knowledge, into one predictive model for task choice under load. This is not to say that the role of these additional attributes has been ignored (see Freed, 2000), and a few have been studied one or two at a time (e.g., Gopher, Brickner, & Navon, 1982; Gopher, Weil, & Siegel, 1989; Liao & Moray, 1993). The construction and continued development of the Strategic Task Overload Model (STOM) promotes the study of multiple *concurrent task factors*, as each has a predictive weighting associated with it.

Informing a Model of Task Switching Choice

As the current work strives to inform both future theory and practice, informing a model sits nicely at the intersection between these scientific pursuits. Therefore, many of the results can be couched in terms of the model, in addition to the broader implications. The STOM model (Wickens et al., 2013) had several basic predictions, which provides a framework for reviewing some of the findings from the current dissertation while also addressing the potential updates recommended to the model. STOM specifically predicted or assumed:

I. Overall switch resistance (expressed as a stay preference), is expected to be higher under higher workload

II. The model posits three distinct features of an ongoing task; task difficulty, interest, and priority, assumed to be causally related to general overall switch resistance. One additional feature of the alternative task, salience, does not exist for the ongoing task.

III. These same attributes from (2) are useful for describing and measuring alternative or interrupting task attractiveness. In this case, they represent weightings that influence what task is switched to, once a switch is initiated in the model.

IV. Features described in (2) and (3) may be weighted differentially, and their net combination in a task affects overall switch propensity (features of the ongoing task) or task attractiveness (features of the alternative task).

V. Critically, one additional feature of the ongoing task is the time-on-task (TOT) factors. By definition, this does not exist for the alternative or interrupting tasks, and while not part of the current STOM model, its role is central in the present experiments.

Each of the above points is further discussed below, in full view of the current work in this dissertation and in context of prior work concerning task switching. Most of these core findings are breaking new ground in research and theory, providing key experimental results, and validating a new model of task switching choice.

A) Overall Switch Resistance Was Generally Found (Especially For Difficult Compared To Easy Conditions)

Based on the literature review, people stay more often than they switch. This general effect was found in multiple studies and was always about the same ratio (.60:.40), leading to a very stable estimation for the model. There is, further, an assumed association of this resistance with the cognitive mechanisms of working memory/executive control associated with the demands of switching tasks. Therefore in general, less switching is the result when resources are scarcer (e.g., when one or more of the tasks in a multi-task scenario is more difficult).

Switch cost effects were observed under a difficult tracking load. In Exp. 1, tracking difficulty was manipulated and indeed, under more demanding conditions switching frequency was reduced. Further evidence in Exp. 1 supported the general difficulty effect when switching occurred between two tasks – resource management and tracking - that heavily competed for the same mental resources within the multiple resource framework (Wickens, 2002), tracking and resource management. Not only did these two tasks garner the majority of the time spent by participants, despite higher ratings of task difficulty, they were switched to and from more often (see Figs. 4.6 & 4.7). The number of switches between these two tasks was greater in the easy tracking condition, an expected result given the shared resources between the two tasks and the role of difficulty.

B) **Attributes Of A Task Explained Switch Choice**

Several attributes of tasks were hypothesized and built into the STOM model that should influence task-switching choice (see Fig 7.2 below).

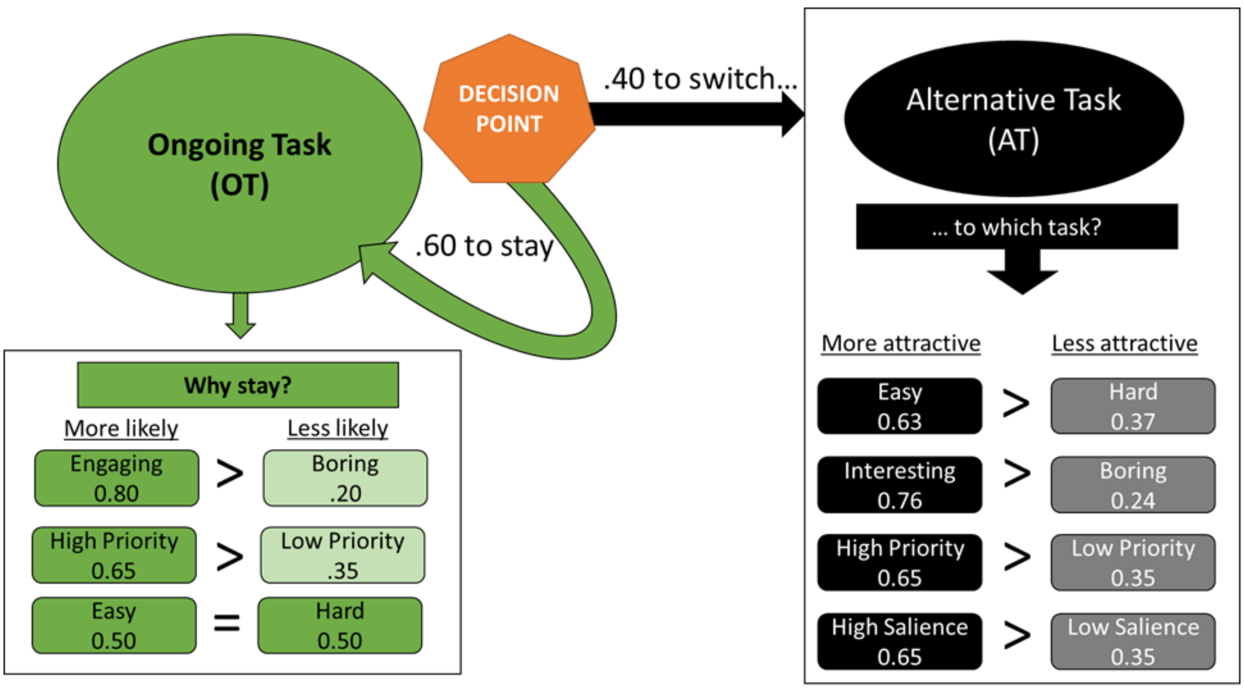


Figure 7.2. The STOM model (updated from Wickens, Santamaria & Sebok, 2013), as shown in Chapter 3.

Task difficulty, as previously described, does indeed appear to cause a general increased switch resistance, and when this difficulty was low (as in the case of the monitoring task), the task was switched to frequently (Exp. 2). Priority of a task, however, does not appear to have the same quantity or quality of support in these experiments. From the manipulation of tracking task priority in Exp. 1, it is clear that at most, priority marginally affects how much time participants spend in a task with this effect interacting with task difficulty. No effects were found for tracking priority on any switching measures.

The survey results from Exp. 1 showed that the instructional manipulation of a task priority was reflected in subjective assessments. And of course, prior work has shown manipulations of this type are successful in changing attention allocation (Gopher et al., 1989), although in timesharing (e.g., concurrent task performance), not task switching (sequential task performance) experiments. As described earlier, however, relationships between each of the task attributes can take multiple forms and do not have to be additive (though this simplifying assumption was made in the current research, and in the STOM model) or interactive; they can simply exist in a hierarchy where one attribute exerts dominance on the others.

In Exp. 2, the survey provided further evidence that priority may play a part in choices for staying with an ongoing task. Of the two tasks being interrupted in the experiment, tracking (rated lower priority, but more *interesting* than communications) was switched away from more often than the communications task. This effect is difficult to disentangle from the jointly different rating of task interest. Given that prior work identified interest as a key attribute that motivates switching in human-computer interactions (Spink, Park, & Koshman, 2006) it would not be surprising if this was indeed an influence, even if it was not explicitly manipulated in these experiments.

Finally, the role of salience in the choice of an alternative task was seen to exert strong effects when examining the conflict event results of Exp. 1, specifically when examining the responses to ownship Communication events (which were assumed high salience as auditory tasks). The results collected in Exp. 1 provided a more direct test of task attractiveness - two task events occurred around the same time, forcing an operator choice. In the strongest test (not confounded with shortcomings of the platform or design), the communications task (ownship) was chosen twice as often as the resource management task due in combination to its priority, salience, and rated difficulty (see Fig. 4.6). In one other conflict event, the resource management task was chosen almost four times more often than a monitoring task, *despite* the higher difficulty and “slow start” in onset behind the monitoring event. The subjective ratings of task attractiveness as a whole did not predict this, due to the equal weighting of attributes. Individually, the interest and priority rating of the resource task explain this outcome (even without any ratings for salience), suggesting that at least some decisions incorporate the influence of task priority. In contrast, in Exp. 2, the factors of priority and interest were consistently overridden by task difficulty ratings – more switches were consistently made to the easier monitoring interruption than to the difficult resource management interruption. The difference between the two findings is intriguing, and suggests that there are times (perhaps during interruption of difficult ongoing task performance) that task difficulty is weighed more than priority and interest.

In sum, task attributes theorized initially to influence both the stay preference (difficulty, priority, interest) and the switch choice (difficulty, priority, interest and salience) all exerted influence on task switching behaviors and choices. Increasing difficulty led to decreasing switches across a variety of conditions (Exp. 1, Exp. 2) and may provide some explanation for the effects in Exp. 3. Subjective difficulty ratings further help predict switch choices in Exp. 2, in which participants switch more often to easier tasks. While priority was also manipulated, its effects on behavior were less evident in traditional switch metrics of frequency, but were present in measures of switch choice in the conflicting task events (Exp. 1) where operators chose to switch to higher priority tasks. Although interest was not manipulated here, higher subjective interest ratings resulted in more switching when operators were presented with two events at the same time (as did higher task salience, another non-manipulated attribute). The remaining challenge is determining the weights in the STOM model, a task for future work.

C) **Time-on-task Effects Were Influential To Switch Choice**

A large portion of this dissertation concerned the **time on the ongoing task**, and its effects on switching behavior (Fig 5.1, reshown below), which was not incorporated into the initial version of the STOM model.

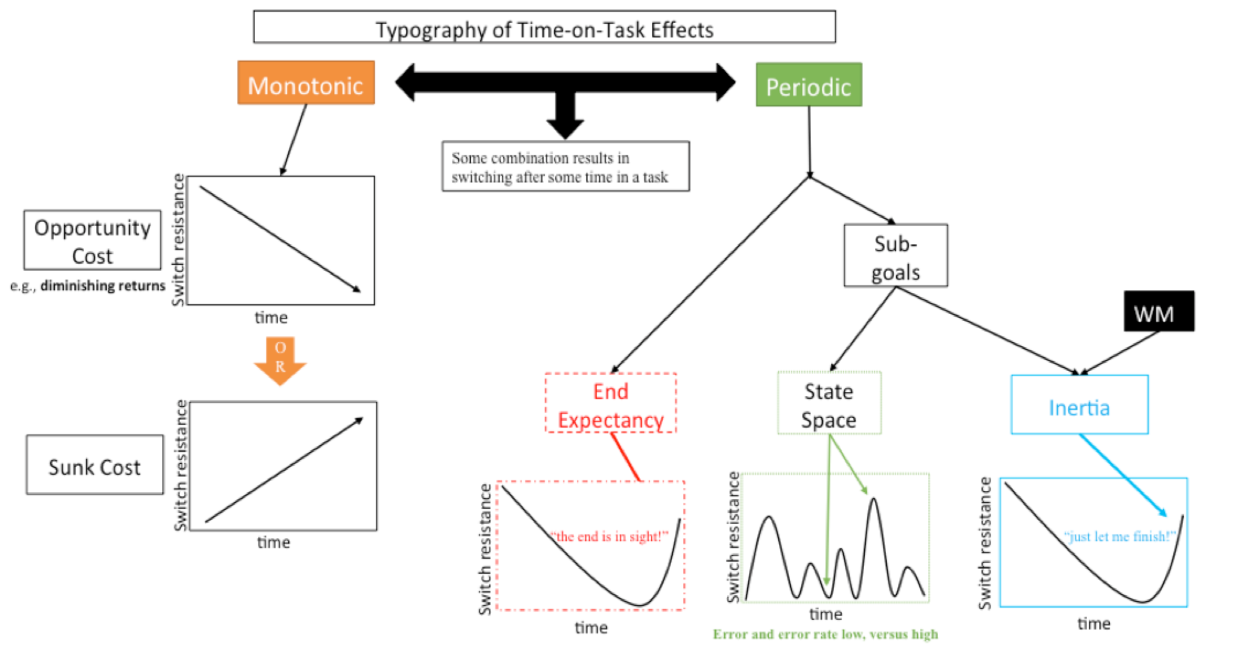


Figure 5.1. The main hypothesized factors that may play a role in determining switch choice over time. WM = working memory. Graphs represent the change in switch resistance over time in an epoch.

The time on task feature was only marginally understood, based on a number of different theories (Baumeister et al., 1998; Duggan et al., 2013; Kurzban, Duckworth, Kable, & Myers, 2013; Payne et al., 2007; Sheridan, 2007) and anecdotal or modeling suggestions for possible outcomes, particularly in the context of the various potential source of time-on-task effects. Each of these theories has merit, but little if any strong experimental support for their assertions, or provide for different influences postulated by different theories may combine in overall influence on switch resistance or likelihood. By addressing this gap with two experiments, additional new ground has been broken.

Two general classes of time-on-task effects were postulated; monotonic (which exert their effects over the task epoch) and periodic (which exert effects based on switching opportunities that develop during the course of task performance). The former monotonic effects have been most recently attributed to declines in switch resistance that result either from resource depletion (Baumeister et al., 1998; Baumeister et al., 2007), or building opportunity cost for an alternative task (Kurzban et al., 2013). However, it was speculated in this dissertation that in addition to other explanations for the same trend (e.g., diminishing returns), other trends (periodicities), such as task inertia, sunk costs or end-expectancy might counter, offset or dominate general decline patterns. The increase in resistance could occur in a continuous fashion across an epoch, as in the observed pattern for task inertia; or only exert influences at specific points in epochs, as in expectancy effects toward the endpoints, or the role of subtask or sub goal boundaries throughout the task epoch.

Periodic Effects. Periodic effects of time-on-task were the most evident in this dissertation. In Exp. 2, the role of working memory load in the communications task decreased switching over time, an excellent example of the task inertia effect (Ch. 5), and at odds with Kurzban's opportunity cost and Baumeister's resource depletion arguments. This effect was even more pronounced when the interrupting task was boring and low priority (monitoring task). Further, the inertia argument is supported because the trends in Comm, assumed to reflect working memory, were absent for the tracking task, which did not increase or decrease switch resistance over time. While these types of effects are sometimes present in task switching literature (Salvucci & Bogunovich, 2010), the current findings expand the generality of the effect to situations where operators have *multiple* competing tasks that they can switch to, and complete freedom in so doing.

State Spaces can illustrate how operators choose periods for switching within a non-working memory task. The state space used in Exp. 2 began as a notional concept for evaluating tracking (see Wickens, Hollands, Banbury & Parasuraman, 2013). Here it was used for the first time to help determine when an 'optimal' period for switching away from a tracking task may occur. It was clear from the analysis that operators switched more often when the task reached an optimal, low rate and magnitude of error condition. The results are importantly in line with the prior work examining switching with task and goal state boundaries (Back, Cox, & Brumby, 2012; Bogunovich & Salvucci, 2008; Janssen et al., 2012; Janssen, Brumby, & Garnett, 2010; Janssen & Brumby, 2010; Salvucci, Taatgen, & Kushleyeva, 2006). The consistent usage of the optimal points in tracking could be a function of workload, with increased load increasing adherence to task boundaries (e.g., Back, Cox, & Brumby, 2012b; Salvucci et al., 2006). This would suggest that periodic effects related to task boundaries vary in strength depending on the demands in the task environment, though the state space must be examined under the easy conditions to make this claim.

Monotonic Effects. A general over-time increase of task switching away from the ongoing task (decrease in switch resistance) was hypothesized to exist on the basis of the prior literature summarized by Kurzban et al. (2013). However little evidence for this in its general form (as due to resource depletion or growing opportunity cost) was present in Exp. 2 or Exp. 3. In fact, in Exp. 2, switches <u>away</u> from the communications task decreased over time. In Exp. 3, some evidence suggested the operation of a general monotonic effect; in both tasks there was a clear decrease in switch resistance from beginning to end, and for both tasks this decrease was greater for diminishing returns trials than constant returns trials (although in monitoring switches this was only supported by a non-significant trend). However, these effects appeared to be added to a global inertia, or end-expectancy effect during the second half of the trial, as the trends were mitigated (monitoring trials) or even reversed from the middle to end segment (tracking trials, the one task repeated as an ongoing task across Exp. 2 and 3).

The source is puzzling, since the effect was not predicted to differ between the constant and diminishing trials, nor was there the kind of look-ahead time available in the task scheduler display as there was in Exp. 2 A plausible end-expectancy effect explanation, in which participants learned the end of the trial was upcoming over multiple practice trials, still does not explain why this pattern was shown for diminishing returns trials, or shown more so in the constant trials. A take home point is that the STOM model will incorporate periodic as well as monotonic effects.

In addition to the main hypotheses examined, two others were explored. The evidence for them, and further work that remains is covered in the sections below.

Practice Does Not Clearly Change Overall Switch Rates

Though in theory task switching may decrease over increased practice time this was not found. Although increased time on task should reduce the demands of performing tasks, in Exp. 1, operators had relatively little task practice. In Exp. 2, no increase in switching was found when practice was extended enough and resulted in a clear performance change.

A practice effect on switching may require a different set of manipulations to reveal. This is the case even in light of a performance improvement in the tracking task over time.

Specifically, given the basis of the general switch resistance effect in resource demands, a task that would show reduced demand over practice time would need to be combined with a series of other tasks in a multi-task management scenario, and be directly compared to a task that did not show these same demand reductions. Alternatively, a variant of a part-task training procedure could be used to train one task extensively, for later performance of it in conjunction with the overall multi-task platform. Of course, invoking such a setup also invokes the up and downsides of part-task training (Gutzwiller, Clegg, & Blitch, 2013; Naylor & Briggs, 1963; So, Proctor, Dunston, & Wang, 2012; Wickens, Hutchins, Carolan, & Cumming, 2012; Wightman & Lintern, 1985; Wightman & Sistrunk, 1987), but the exhibitions of switching choice over multi-task practice is essentially an unknown.

Executive Control Did Not Clearly Affect Switch Rates

The hypothesized effect was that greater individual differences in executive functioning may lower switch likelihood (as shown in Arrington & Yates, 2009), an aspect not covered in the STOM model. However, the evidence for this was weak at best, and suggested the relationship may exist only to the extent that task instructions impose constraints on switching. One possibility that supports that idea is that executive functioning tends to be associated with a 'planning' or 'strategic' component of cognitive processing; in the voluntary task clearly this strategy component is tapped repeatedly under the standard instructions; whereas in the current experiment the strategy component may take several different forms and cloud any relationship. A follow-up experiment, manipulating task instructions restricting free choice in the current platform would be needed in order to determine whether such a relationship exists.

However, though individual differences in executive control (as measured with the ANT task) were not present in the current experiment (Exp 2., Ch. 5), other individual differences may exert effects on task switching choice and behavior. Several examples can be found for individual differences that influence multitasking, in which large effects of polychronicity (whether there is a preference to multitask) account for whether operators multitask, in addition to impulsivity (e.g, König, Oberacher, & Kleinmann, 2010). Further polychronicity may enable those *with* good multitasking ability to be more effective in the workplace (Sanderson, Bruk-lee, Viswesvaran, Gutierrez, & Kantrowitz, 2013). Individual differences are not constrained to an abilities perspective, but may also arise from extensive practice. A recent study showed frequent heavy media multitasking (e.g., consuming multiple media contents) may lead to reduced cognitive control on tasks with distracting components (Ophir, Nass, & Wagner, 2009). This suggests that while no clear training effects on switching were found here, perhaps the timeframe did not induce enough task practice.

Limitations And Future Directions

One obvious limitation to the findings in this dissertation is that they can only inform the model to the extent that they capture all potential task switches. However, it is quite possible that some switches occur that do not manifest as behaviors traceable in the platform. For example, it is possible that operators "switched" to a task that was not able to be recorded – such as scanning for information in MATB, a visual task that would have competed with each of the other tasks present at any given time. Thus a question arises whether the current data informs a model of task switching, or a model of task executions.

As difficulty was manipulated, it could be that this 'scan' task occurred more frequently, helping participants to more effectively task manage than during difficult conditions. As another example, within the time on task experiments, any participant potentially engaging in a "scanning" for visual information task could be attending to different tasks more frequently than others based on the diminishing rate of return. One potential solution to these problems would be to add eye-tracking assessment or, to restrict the visual presentation of information within the platform itself to only the ongoing task.

Additionally, the current experiments involve novices performing tasks and gaining some amount of task proficiency. However, this is likely to result in a much lower proficiency than most operators in cockpits or offices posses, and prior evidence suggests experience develops some hierarchies of task attributes like priority (Raby & Wickens, 1992). Therefore the current experiments do a good, but not perfect, job of speaking to the model.

A follow-on limitation needs to be expressed for the STOM model itself. The model is *not* intended as a complete model of task switching behavior. In fact, it is fairly restricted to contexts specifically when concurrent multitasking is not possible, and instead, sequential (more in line with task management) multitasking behavior occurs. One reason all of the experiments attempted to heavily tax the operator was that high cognitive load represents one of the characteristics of single-channel processing conditions (see Chapter 1) along with the restriction of responses, and the presentation of near-simultaneous events. A separate but related model does exist that addresses quite well the *concurrent* multitasking realm in terms of multiple resources (Wickens, 2002; 2008).

Time On Task Effects Need Quantification Of Time. Although the results for Exp. 2 and Exp. 3 partially support a general category of switch over time effects, they are not conclusive. It is possible that both monotonic and periodic effects can also be found, but that it is difficult to separate them out. The main problem in doing so seems to be that timeframes for effects are not established. For example, by their nature, periodic effects will not appear uniformly across any time period of task performance. But monotonic effects as described here may require some amount of time in a task, in order to manifest. Salvucci and Taatgen (2011) clearly describe the limitations of different time scales for multitasking, one of which is that as the time frames enlarge, multitasking may become more sequential. In the current scope, we strove to keep the experiments in the sequential, single-task management, task-switching domain, rather than the concurrent multitasking, parallel processing, multiple resource domains. But it is possible that the theories collected, with their broad goals of explaining human behavior, do not account for these types of differences.

However, one clear way to address such an issue is to examine different literal time courses of ongoing task performance for the presence of different effects, instead of using overall multi-task trial time epochs as the timeframe for effects as was done in this dissertation. Timeframes could be done in one of two ways, either by using tasks with varying durations, or, by allowing operators to perform tasks naturally which take them varying amounts of time, and then using a filter to categorize shorter to longer time periods taken, examining the switch frequency in each. The secondary approach has more merit – the result of less experimental control means results may be more easily scaled from the lab to the real world where these choices are most likely to play a role in human performance. This is demanded to maintain rigor as these theories develop, and to make more specific conclusions about the theories.

Conclusions

From the current literature review, modeling efforts, and three experiments, a few take-home points can be described briefly. First, while task switching is a basic science concept in cognitive psychology, applications are beginning to encourage experimentation using more real-world tasks. Thus increasing the need to examine free choices helps to move this work into domains that can make use of the findings shown here. For example, in driving multiple other tasks may impinge on a drivers attention toward the road, and simply telling people that driving is the highest priority task may not actually change behavior – perhaps this priority needs to be incorporated with an increased interest component to induce longer periods of sustained attention. Although priority may be a focus of driver training, it has to combat the role of task difficulty and interest of competing tasks, and these competing tasks (especially cell phone use) is increasing in ease and interest. These same issues are at the heart of technological integration in terms of display use in aviation; and in fact the overwhelming nature of interest and engagement can lead some displays to be dangerous precisely *because* attention is less likely to be switched away, even to critical high priority events in the environment.

A similar perspective can be taken for operators interacting with high-complexity command and control systems. While high priority tasks for drone operators may emerge during a mission, their likelihood to be switched to (assuming all of the constraints of the given model used here) might be further enhanced if the task is also easier or at least more interesting and more salient than another task. When considering military domains, priorities can sometimes be shown in hierarchies or determined by automated systems, but operators may not always agree with such optimal conceptualizations in terms of behavior, and may not always weight priority over difficulty, interest, and saliency. Thus a more optimal system may actually exploit more than priority and attempt to combine task attributes rather than focus solely on one at a time. While these facets have not yet been explored, the current model does allow for some predictions, and represents a clear step forward in our thinking.

Second, these real-world contexts still present a scenario in which switching has a cost, and thus switching will generally be less preferred than staying with a task. Once a decision to switch has been made, though, tasks with high priority in combination with other attributes may garner attention (if they directly compete with other tasks). Otherwise, the consistent role of task difficulty will come to bear and operators appeared to be most likely to choose the easiest alternative task when faced with interruptions. Over time, switch likelihood may increase (when tasks are diminishing in returns) or decrease (when task inertia and the costs of leaving are high), and may vary with periodic fluctuations in task and goal boundaries (and influenced by overall task load). These types of effects need to be examined more closely for command and control tasks. A good example may be a search and reconnaissance type of task, where detection and tracking of potential targets may initially have a high rate of return, but over time the ‘low hanging fruit’ of more obvious targets is picked, and operators may become more likely to then switch to a different task.

As a final concluding note, it seems JFK’s famous decision to put a man on the moon indeed makes sense within the current framework. In the face of the competing interests of the country, the priority *and* interest attributes dominated the ease of the task at hand; and the decision was made to follow through “because it is hard".

REFERENCES

Adler, R. F., & Benbunan-Fich, R. (2013). Self-interruptions in discretionary multitasking. *Computers in Human Behavior*, *29*(4), 1441–1449. doi:10.1016/j.chb.2013.01.040

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model. *Cognitive Science*, *26*(1), 39–83. doi:10.1207/s15516709cog2601_2

Arrington, C. M. (2008). The effect of stimulus availability on task choice in voluntary task switching. *Memory & Cognition*, *36*(5), 991–997. doi:10.3758/MC.36.5.991

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science*, *15*(9), 610–615. doi:10.1111/j.0956-7976.2004.00728.x

Arrington, C. M., & Logan, G. D. (2005). Voluntary task switching: chasing the elusive homunculus. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*(4), 683–702. doi:10.1037/0278-7393.31.4.683

Arrington, C. M., & Rhodes, K. M. (2010). Perceptual asymmetries influence task choice: the effect of lateralised presentation of hierarchical stimuli. *Laterality*, *15*(5), 501–513. doi:10.1080/13576500902984695

Arrington, C. M., Weaver, S. M., & Pauker, R. L. (2010). Stimulus-based priming of task choice during voluntary task switching. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*(4), 1060–1067. doi:10.1037/a0019646

Arrington, C. M., & Yates, M. M. (2009). The role of attentional networks in voluntary task switching. *Psychonomic Bulletin & Review*, *16*(4), 660–665. doi:10.3758/PBR.16.4.660

Back, J., Cox, A., & Brumby, D. (2012a). Choosing to interleave: Human error and information access cost. *... the 2012 ACM Annual Conference on Human …*, 1651–1654. Retrieved from http://dl.acm.org/citation.cfm?id=2208289

Back, J., Cox, A., & Brumby, D. (2012b). Choosing to interleave: Human error and information access cost. *... the 2012 ACM Annual Conference on Human …*, 1651–1654.

Bailey, B. P., & Iqbal, S. T. (2008). Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transactions on Computer-Human Interaction*, *14*(4), 1–28. doi:10.1145/1314683.1314689

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9599441

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355. doi:10.1111/j.1467-8721.2007.00534.x

Bogunovich, P., & Salvucci, D. D. (2008). Inferring Multitasking Breakpoints from Single-Task Data. *Proceedings of the 32nd Annual Meeting of the Cognitive SCience Society*, 1732–1737.

Brumby, D., Rosario, N. del, & Janssen, C. (2010). When to switch? Understanding how performance tradeoffs shape dual-task strategy. *Proc. ICCM 2010*, 19–24. Retrieved from http://csjarchive.cogsci.rpi.edu/proceedings/2010/papers/0438/paper0438.pdf

Brumby, D., Salvucci, D., & Howes, A. (2007). Dialing while driving? A bounded rational analysis of concurrent multi-task behavior. In *Proceedings of the 8th International Conference on Cognitive Modeling*. Hove, UK: Psychology Press. Retrieved from http://web4.cs.ucl.ac.uk/uclic/people/d.brumby/publications/Brumby.Salvucci.Howes.2007.ICCM.pdf

Burke, J. L., Murphy, R., Coovert, M. D., & Riddle, D. L. (2004). Moonlight in Miami: Field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction*, *19*, 85–116. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/07370024.2004.9667341

Butler, K. M., Arrington, C. M., & Weywadt, C. (2011). Working memory capacity modulates task performance but has little influence on task choice. *Memory & Cognition*, *39*(4), 708–724. doi:10.3758/s13421-010-0055-y

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident; Analysis and Prevention*, *40*(4), 1282–93. doi:10.1016/j.aap.2008.01.009

Chou, C., Madhavan, D., & Funk, K. (1996). Studies of cockpit task management errors. *The International Journal of Aviation Psychology*, *6*(4), 307–320. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15327108ijap0604_1

Comstock, J. R., & Arnegard, R. J. (1992). *The Multi-Attribute Task Battery for human operator workload and strategic behavior research (Tech Memoradum 104174). (Tech. Memorandum 104174).* Hampton, VA: NASA Langley Research Center.

Cook, M. B., & Smallman, H. S. (2013). Human-Centered Command and Control of Future Autonomous Systems. *18 Th ICCRTS : C2 in an Underdeveloped and / or Denied Environment*.

Damos, D., & Wickens, C. (1980). The identification and transfer of timesharing skills. *Acta Psychologica*, *46*, 15–39. Retrieved from http://www.sciencedirect.com/science/article/pii/0001691880900578

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology. Human Perception and Performance, 19*(5), 965–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8228846

Demanet, J., Baene, W. De, Arrington, C. M., & Brass, M. (2013). Biasing free choices : The role of the rostral cingulate zone in intentional control ☆. *NeuroImage, 72*, 207–213.

Demanet, J., & Liefooghe, B. (2013). Component processes in voluntary task switching. *Quarterly Journal of Experimental Psychology (2006)*, 1–18. doi:10.1080/17470218.2013.836232

Demanet, J., Liefooghe, B., & Verbruggen, F. (2011). Valence, arousal, and cognitive control: a voluntary task-switching study. *Frontiers in Psychology*, *2*(November), A336, 1–9. doi:10.3389/fpsyg.2011.00336

Demanet, J., Verbruggen, F., Liefooghe, B., & Vandierendonck, A. (2010). Voluntary task switching under load: contribution of top-down and bottom-up factors in goal-directed behavior. *Psychonomic Bulletin & Review, 17*(3), 387–393. doi:10.3758/PBR.17.3.387

Dismukes, R. (2010). Remembrance of things future: prospective memory in laboratory, workplace, and everyday settings. *Reviews of Human Factors and Ergonomics*, 1–86. Retrieved from http://rev.sagepub.com/content/6/1/79.short

Dismukes, R., & Berman, B. (2010). *Checklists and monitoring in the cockpit: Why crucial defenses sometimes fail.*

Dismukes, R., & Nowinski, J. (2007). Prospective memory, concurrent task management, and pilot error. *Attention: From Theory to Practice*, 225–236. Retrieved from http://books.google.com/books?hl=en&lr=&id=wEUi-bEaqWcC&oi=fnd&pg=PA225&dq=Prospective+Memory+,+Concurrent+Task+Management+,+and+Pilot+Error&ots=51Jn_6FL7p&sig=necsEjx_RHujHaD9qONd1IAbyKg

Drews, F., & Strayer, D. (2008). Cellular Phones and Driver Distraction. In *Driver distraction: Theory, effects, and mitigation* (pp. 169–190). Retrieved from http://books.google.com/books?hl=en&lr=&id=o7--7AS38tYC&oi=fnd&pg=PA169&dq=Cellular+phones+and+driver+distraction&ots=gJgXcA1r2_&sig=xeBFBFG2o0LSTKSO1eVIbMdnBWA

Duggan, G., Johnson, H., & Sørli, P. (2013). Interleaving tasks to improve performance: Users maximise the marginal rate of return. *International Journal of Human-Computer Studies, 71*(5), 533–550. Retrieved from http://www.sciencedirect.com/science/article/pii/S1071581913000050

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26*(2), 471–9. doi:10.1016/j.neuroimage.2005.02.004

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. doi:10.1162/089892902317361886

Fisher, D. L., Pollatsek, A. P., & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention*, *12(Suppl 1*, i25–i29. doi:10.1136/ip.2006.012021

Funk, K. (1991). Cockpit Task Management : Preliminary Definitions, Nonnative Theory, Error Taxonomy, and Design Recommendations. *International Journal of Aviation Psychology*, *1*(4), 271–285.

Gollan, T., & Ferreira, V. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *35*(3), 640–665. doi:10.1037/a0014981.Should

Gopher, D., Brickner, M., & Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: evidence for multiple resources. *Journal of Experimental Psychology. Human Perception and Performance*, *8*(1), 146–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6460080

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, *71*, 147–177.

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, *3*(1), 42–66. doi:10.1037//1076-898X.3.1.42

Gutzwiller, R. S., Clegg, B. A., & Blitch, J. G. (2013). Part-Task Training in the Context of Automation : Current and Future Directions. *American Journal of Psychology*, *126*(4), 417–432.

Gutzwiller, R., Wickens, C., & Clegg, B. (2014). Workload overload modeling: An experiment with MATB II to inform a computational model of task management. *Proceedings of the 2014 Human Factors Conference*.

Hameed, S., Ferris, T., Jayaraman, S., & Sarter, N. (2009). Using informative peripheral visual and tactile cues to support task and interruption management. *Human Factors*, *51*(2), 126–135. doi:10.1177/0018720809336434.

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, *12*(2), 67–78. doi:10.1037/1076-898X.12.2.67

Janssen, C., Brumby, D. P., & Garnett, R. (2012). Natural Break Points The Influence of Priorities and Cognitive and Motor Cues on Dual-Task Interleaving. *Journal of Cognitive Engineering and Decision Making*, *6*(1), 5–29. doi:10.1177/1555343411432339.

Janssen, C. P., & Brumby, D. P. (2010). Strategic adaptation to performance objectives in a dual-task setting. *Cognitive Science*, *34*(8), 1548–1560. doi:10.1111/j.1551-6709.2010.01124.x

Janssen, C. P., Brumby, D. P., Dowell, J., Chater, N., & Howes, A. (2011). Identifying Optimum Performance Trade-Offs Using a Cognitively Bounded Rational Analysis Model of Discretionary Task Interleaving. *Topics in Cognitive Science*, *3*(1), 123–139. doi:10.1111/j.1756-8765.2010.01125.x

Janssen, C. P., Brumby, D. P., & Garnett, R. (2010). Natural break points: Utilizing motor cues when multitasking. *Proceedings of the 54th Annual Human Factors and Ergonomics Society*, *54*(4), 482–486. doi:10.1037/e578652012-044

Jersild, A. (1927). Mental set and shift. *Archives of Psychology*, *Whole No.* .

Jin, J., & Dabbish, L. (2009). Self-interruption on the computer: a typology of discretionary task interleaving. *Proceedings of the SIGCHI Conference on Human …*, 1799–1808. Retrieved from http://dl.acm.org/citation.cfm?id=1518979

John, M. F. S., & King, M. A. (2010). The Four-Second Supervisor: Multi-Tasking Supervision and Its Support. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(4), 468–472. doi:10.1177/154193121005400441

Katidioti, I., & Taatgen, N. a. (2013). Choice in Multitasking: How Delays in the Primary Task Turn a Rational Into an Irrational Multitasker. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi:10.1177/0018720813504216

Kessler, Y., Shencar, Y., & Meiran, N. (2009). Choosing to switch: spontaneous task switching despite associated behavioral costs. *Acta Psychologica*, *131*(2), 120–128. doi:10.1016/j.actpsy.2009.03.005

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching- A review. *Psychological Bulletin*, *136*(5), 849–874. doi:10.1037/a0019842

Koh, R. Y. I., Park, T., Wickens, C. D., Ong, L. T., & Chia, S. N. (2011). Differences in attentional strategies by novice and experienced operating theatre scrub nurses. *Journal of Experimental Psychology: Applied*, *17*(3), 233–246. doi:10.1037/a0025171

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

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *The Behavioral and Brain Sciences*, *36*(6), 661–79. doi:10.1017/S0140525X12003196

Kushleyeva, Y., Salvucci, D. D., & Lee, F. J. (2005). Deciding when to switch tasks in time-critical multitasking. *Cognitive Systems Research*, *6*(1), 41–49. doi:10.1016/j.cogsys.2004.09.005

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, *38*(3), 322–332. doi:10.3758/MC.38.3.322

Liao, J., & Moray, N. (1993). A simulation study of human performance deterioration and mental workload. *Le Travail Humain*, *56*(4), 321–344. Retrieved from http://www.jstor.org/stable/10.2307/40659831

Liefooghe, B., Demanet, J., & Vandierendonck, A. (2009). Is advance reconfiguration in voluntary task switching affected by the design employed? *Quarterly Journal of Experimental Psychology*, *62*(5), 850–857. doi:10.1080/17470210802570994

Liefooghe, B., Demanet, J., & Vandierendonck, A. (2010). Persisting activation in voluntary task switching: it all depends on the instructions. *Psychonomic Bulletin & Review*, *17*(3), 381–386. doi:10.3758/PBR.17.3.381

Lien, M.-C., & Ruthruff, E. (2008). Inhibition of task set: converging evidence from task choice in the voluntary task-switching paradigm. *Psychonomic Bulletin & Review*, *15*(6), 1111–1116. doi:10.3758/PBR.15.6.1111

Loukopoulos, L., Dismukes, R., & Barshi, I. (2009). *The multitasking myth*. Farnham: Ashgate Publishing, Ltd.

Lu, S. A., Wickens, C. D., Prinet, J. C., Hutchins, S. D., Sarter, N., & Sebok, A. (2013). Supporting Interruption Management and Multimodal Interface Design: Three Meta-Analyses of Task Performance as a Function of Interrupting Task Modality. *Human Factors*. doi:10.1177/0018720813476298

Mayr, U., & Bell, T. (2006). On how to be unpredictable: evidence from the voluntary task-switching paradigm. *Psychological Science*, *17*(9), 774–780. doi:10.1111/j.1467-9280.2006.01781.x

McFarlane, D. (2002). Comparison of Four Primary Methods for Coordinating the Interruption of People in Human-Computer Interaction. *Human-Computer Interaction*, *17*(1), 63–139. doi:10.1207/S15327051HCI1701_2

McFarlane, D., & Latorella, K. (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction*, *17*(1), 1–61. doi:10.1207/S15327051HCI1701_1

McVay, J., & Kane, M. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*, 188–197. Retrieved from http://psycnet.apa.org/journals/bul/136/2/188/

Monk, C., Boehm-Davis, D., & Trafton, J. (2004). Recovering from interruptions: implications for driver distraction research. *Human Factors*, *46*(4), 650–663. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15709327

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. doi:10.1016/S1364-6613(03)00028-7

Monsell, S., & Driver, J. (2000). Banishing the Control Homunculus. *Control of Cognitive Processes*, (Chapter 1), 5–32. Retrieved from http://diyhpl.us/~bryan/papers2/neuro/Control of Cognitive Processes.pdf#page=5

Monsell, S., Sumner, P., & Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Memory & Cognition*, *31*(3), 327–342. doi:10.3758/BF03194391

Montgomery, H., Sharafi, P., & Hedman, L. R. (2004). Engaging in activities involving information technology: Dimensions, modes and flow. *Human Factors*, *46*(2), 334–348. doi:10.1518/hfes.46.2.334.37345

Murphy, R. R. (2004). Human–robot interaction in rescue robots. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, *34*(2), 138–153. doi:10.1109/TSMCC.2004.826267

Naylor, J. C., & Briggs, G. E. (1963). Effects of task complexity and task organizatoin on the relative efficiency of part and whole training techniques. *Journal of Experimental Psychology*, *65*(3), 217–224.

Norman, D., & Bobrow, D. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *4*. Retrieved from http://www.sciencedirect.com/science/article/pii/0010028575900043

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15583–15587. doi:10.1073/pnas.0903620106

Panepinto, M. P. (2010). Voluntary versus forced task switching. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(4), 453–457. doi:10.1177/154193121005400438

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, *116*(2), 220–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7972591

Pashler, H. (2000). 12 Task Switching and Multitask Performance. *Control of Cognitive Processes*. Retrieved from http://diyhpl.us/~bryan/papers2/neuro/Control of Cognitive Processes.pdf#page=267

Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology. General*, *136*(3), 370–88. doi:10.1037/0096-3445.136.3.370

Raby, M., & Wickens, C. (1994). Strategic workload management and decision biases in aviation. *The International Journal of Aviation …*, *4*(3), 211–240. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15327108ijap0403_2

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology-General*, *124*(2), 207. Retrieved from http://www.quartetfest.ca/documents/37871/Rogers_Monsell_1995.pdf

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 763–797. doi:10.1037//0096-1523.27.4.763

Ruthruff, E., Hazeltine, E., & Remington, R. W. (2006). What causes residual dual-task interference after practice? *Psychological Research*, *70*(6), 494–503. doi:10.1007/s00426-005-0012-8

Salvucci, D. D. (2005). A multitasking general executive for compound continuous tasks. *Cognitive Science*, *29*(3), 457–492. doi:10.1207/s15516709cog0000_19

Salvucci, D. D., & Bogunovich, P. (2010). Multitasking and Monotasking : The Effects of Mental Workload on Deferred Task Interruptions. *CHI 2010: Multitasking, Altlanta, Georgia*, 85–88.

Salvucci, D., & Taatgen, N. (2011). *The multitasking mind*. New York, NY: Oxford University Press, Inc. Retrieved from http://books.google.com/books?hl=en&lr=&id=YE-anWg_nNsC&oi=fnd&pg=PR11&dq=The+multitasking+mind&ots=R2u_ozW3h4&sig=JuvJ2gicY0vlGJyml6a5yvAm2_8

Salvucci, D., Taatgen, N., & Kushleyeva, Y. (2006). Learning when to switch tasks in a dynamic multitasking environment. *Proceedings of the Seventh …*, 1–6. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.72.7292&rep=rep1&type=pdf

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS One*, *8*(1), e54402. doi:10.1371/journal.pone.0054402

Sanderson, K. R., Bruk-lee, V., Viswesvaran, C., Gutierrez, S., & Kantrowitz, T. (2013). Multitasking: Do preference and ability interact to predict performance at work? *Journal of Occupational and Organizational Psychology*, *84*(4), 556–563.

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

Sarter, N. (2013). Multimodal Support for Interruption Management: Models, Empirical Findings, and Design Recommendations. *Proceedings of the IEEE*, 1–8. doi:10.1109/JPROC.2013.2245852

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: uncorking the central cognitive bottleneck. *Psychological Science*, *12*(2), 101–108. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11340917

So, J. C. Y., Proctor, R. W., Dunston, P. S., & Wang, X. (2012). Better retention of skill operating a simulated hydraulic excavator after part-task than after whole-task training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi:10.1177/0018720812454292

Spink, A., Park, M., & Koshman, S. (2006). Factors affecting assigned information problem ordering during Web search: An exploratory study. *Information Processing & Management*, *42*, 1366–1378. Retrieved from http://www.sciencedirect.com/science/article/pii/S0306457306000112

Stoet, G., & Snyder, L. H. (2007). Extensive practice does not eliminate human switch costs. *Cognitive, Affective & Behavioral Neuroscience*, *7*(3), 192–197. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17993205

Strobach, T., Liepelt, R., Schubert, T., & Kiesel, A. (2012). Task switching: effects of practice on switch and mixing costs. *Psychological Research*, *76*(1), 74–83. doi:10.1007/s00426-011-0323-x

Trafton, J. G., Altmann, E., & Brock, D. P. (2005). Huh, what was I doing? How people use environmental cues after an interruption. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 49*, *49*(3), 468–472. doi:10.1037/e577392012-053

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, *58*(5), 583–603. doi:10.1016/S1071-5819(03)00023-5

Trafton, J., & Monk, C. (2007). Task interruptions. *Reviews of Human Factors and Ergonomics*, 111–126. Retrieved from http://rev.sagepub.com/content/3/1/111.short

Vandamme, K., Szmalec, A., Liefooghe, B., & Vandierendonck, A. (2010). Are voluntary switches corrected repetitions? *Psychophysiology*, *47*(6), 1176–1181. doi:10.1111/j.1469-8986.2010.01032.x

Weaver, S. M., & Arrington, C. M. (2010). What's on your mind: the influence of the contents of working memory on choice. *Quarterly Journal of Experimental Psychology*, *63*(4), 726–737. doi:10.1080/17470210903137180

Welford, A. T. (1967). Single-channel operation in the brain. *Acta Psychologica*, *27*, 5–19.

Weywadt, C. R. B., & Butler, K. M. (2013). The role of verbal short-term memory in task selection: how articulatory suppression influences task choice in voluntary task switching. *Psychonomic Bulletin & Review*, *20*(2), 334–340. doi:10.3758/s13423-012-0349-0

Wickens, C. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors …*, *50*(3), 449–455. doi:10.1518/001872008X288394.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/14639220210123806

Wickens, C. D., Hutchins, S., Carolan, T., & Cumming, J. (2012). Effectiveness of Part-Task Training and Increasing-Difficulty Training Strategies: A Meta-Analysis Approach. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi:10.1177/0018720812451994

Wickens, C. D., & McCarley, J. S. (2008). *Applied attention theory*. Boca Raton, FL: CRC Press.

Wickens, C. D., Santamaria, A., & Sebok, A. (2013). A Computational Model of Task Overload Management and Task Switching. *Proceedings of the Human …*, *57*(1), 763–767. doi:10.1177/1541931213571167

Wickens, C., Gutzwiller, R., & Santamaria, A. (n.d.). Discrete task switching in overload: Two meta-analysis and a model. *International Journal of Human-Computer Studies*.

Wickens, C., Hollands, J., Banbury, S., & Parasuraman, R. (2013). *Engineering psychology and human performance* (4th ed.). Upper Saddle River, NJ: Pearson.

Wightman, D. C., & Lintern, G. (1985). Part-task training for tracking and manual control. *Human Factors*, *27*(3), 267–283.

Wightman, D. C., & Sistrunk, F. (1987). Part-task training strategies in simulated carrier landing final-approach training. *Human Factors*, *29*(3), 245–254.

Yeung, N. (2010). Bottom-up influences on voluntary task switching: the elusive homunculus escapes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*(2), 348–362. doi:10.1037/a0017894

APPENDIX A: MATB INSTRUCTIONS USED FOR EXPERIMENT 1

Learning how to perform MATB II

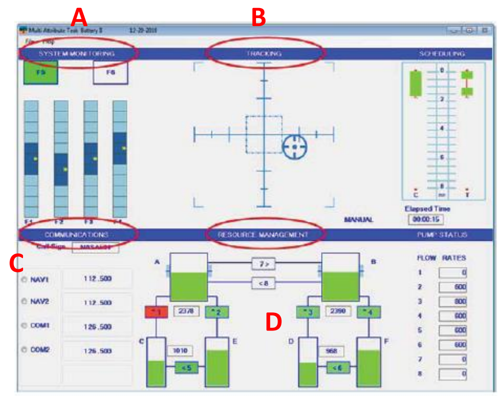
Welcome

- Please take 10-15 minutes to review the following materials. You can return to previous material to review as often as you'd like after you've reviewed the entire document once, up until time is called.
- At the end of the time, you'll be asked to actually perform these tasks in the simulation – so pay attention! No one can help you after the training period is over.

This is an overview of the main display that you will see during the experiment.

The main display shows the inner workings of several tasks (shown here marked A, B, C, and D).

Performance of these tasks is done simultaneously, but you will learn about each system one by one.



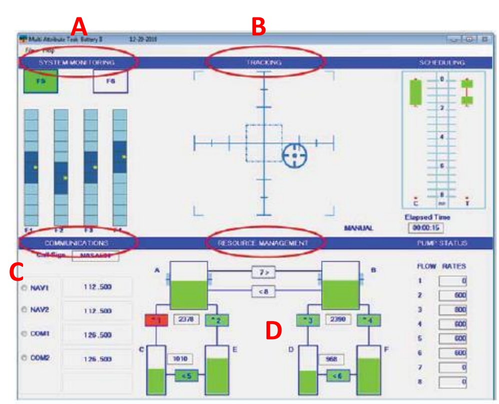
Four tasks simulate the kinds of tasks that REAL pilots perform during flight of a commercial jetliner.

These tasks are *critical* to master, as you will be asked to perform all of them.

These tasks include:

a) monitoring,
b) tracking,
c) communicating,
d) and resource management

Each of these tasks has been labeled – take a moment to find them on the main display



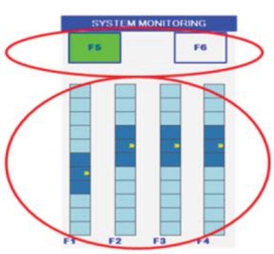
System Monitoring Task (task A)

The monitoring task appears in the upper left of the main display. There are 2 main components: At the top, two squares are shown. The rest of the task shows 4 sliding scales, which will be explained shortly.

The left square at the top should appear green; however, if may appear white if it is turned off. *You should respond by click it to turn the button back on.*

The right square at the top should appear white; however, it may appear red indicating a failure. *You should respond by clicking the button to turn it off.*

System Monitoring Task (task A)

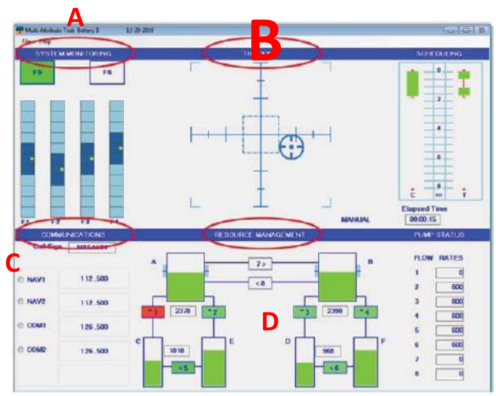


The rest of the task shows 4 sliding scales, each of which has an indicator.

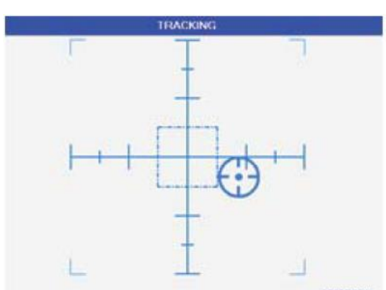
Each indicator can go “out of range” on each sliding scale, either too high or too low. Your job is to detect these deviations, and click on the scale with the deviation when it occurs.

Remember, you are shooting to quickly, AND ACCURATELY, detect these events.

Task B = Tracking Task



Tracking Task (Task B)

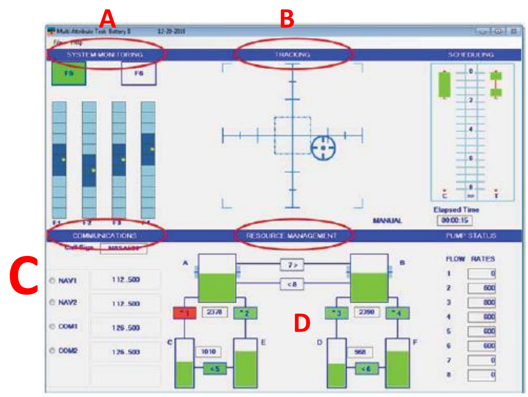


For this task, you will use the joystick to control the small, circular target reticle. Your job is to keep the circular target WITHIN THE BOUNDARIES of the smaller square and as much on the cross as possible.

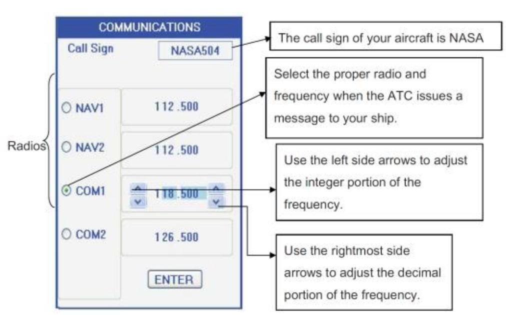
You can think of this like piloting a plane under turbulence; weather will alter the aircraft's path – your job is to keep on course.

If you find the circular reticle drifting out of the center, quickly use the joystick to re-center it.

Task C = Communications Task



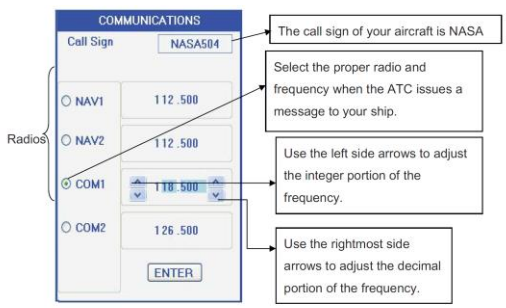
Communication Task (Task C)



You will be hearing messages through your headphones that call a particular aircraft by a "call sign" assigned to it. These will be presented twice.

When you receive a message it will contain two parts; the first part is the call sign. The second part is a command requiring you to tune one of the 4 radios to a particular frequency.

Communication Task (Task C)

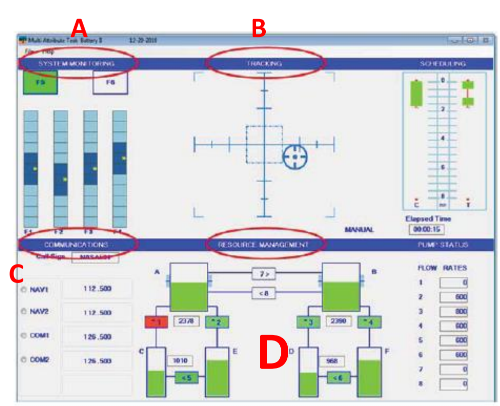


The overall goals are:

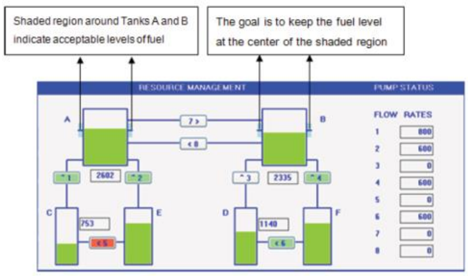
1) Determine whether the message is intended for you: does it match your call sign?

2) If so, quickly and accurately select the radio, and frequency and adjust it accordingly.

Task D: Resource Management Task



Resource Management Task (Task D)

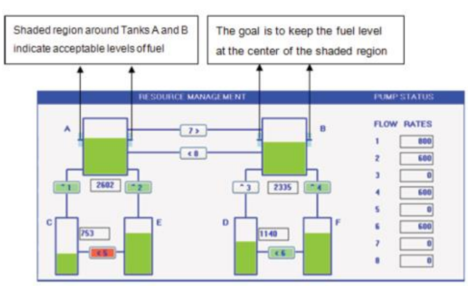


The resource management task represents the necessity to manage fuel aboard an aircraft (shown in green containers).

Each container is represented with a letter (A through F). Lines connect certain containers with other containers, and the flow of fuel between them is controlled by a pump.

Each PUMP is numbered (1 through 8). Pumps are used to transfer fuel from supply tanks to the main tanks.

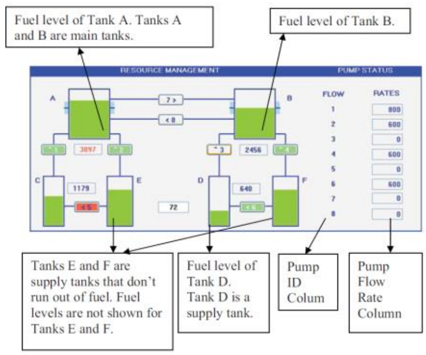
Resource Management Task (Task D)



Pumps are color-coded for their individual state of operation:

- green: pump is ON and flowing in the direction of it's arrow; can be turned off by clicking on it
- White: pump is OFF, but can be turned on if needed by clicking on it
- Red: pump has FAILED, BUT you MUST continue to keep the center fuel levels at the center of the shaded region.

Resource Management Task (Task D)

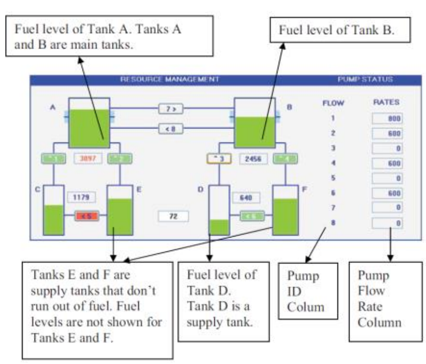


The numbers underneath each container represent the amount of fuel in the container. These numbers can decrease and increase, corresponding to flow into, or out of, the tank.

Tanks C and D are supply tanks, and contain a maximum of 2000 units. Tanks E and F are unlimited.

You must therefore transfer fuel to tanks A and B, via transfers from C and D. This is because tanks A and B continually use fuel to keep the engine going and the plane propelled through the air.

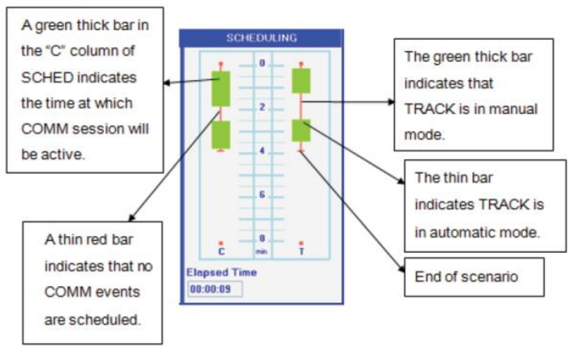
Resource Management Task (Task D)



You cannot let containers A, B, C, or D hit empty. They will become unable to accept transfers of fuel at that point.

Again, the goal of the task is to keep the containers A and B filled to within the range indicated by the shaded regions (around 2500 units). There are multiple ways to achieve this goal; you may use the method that works best for you.

Task Scheduler Display (upper right corner)



Task Scheduler Display (upper right corner)

- As shown, the task scheduler may help you allocate your attention to upcoming tasks (but it only shows the tracking task, and the comms task).
- It does not show the resource management events, or the monitoring task events from before.

Final Performance:

Manipulated between groups

- Now that you have reviewed all of the component tasks, some general goals for performing them.

- You will be performing all of the tasks simultaneously. Your goal is to be as fast and as accurate as possible in ALL OF THE TASKS. Spread your attention evenly to accomplish this as effectively as possible, UNLESS otherwise instructed.

- You will complete a basic training trial that will not be scored before beginning your test trials.

- Feel free to return to earlier slides if the time has not run out.

APPENDIX B: PAIRED TASK COMPARISON SURVEY (USED IN EXP 1)

Appendix B: Paired Task Comparison Survey

Based on your experience in MATB**, you will compare several tasks to each other by rating which task represents the category more**. For example, you might be asked to compare two tasks (A and B) on which was a higher priority to perform.

Task A ←--*higher task priority--lower -- 0 (equal) -- lower --higher task priority*→Task B

So, if you thought Task A was *higher priority* than Task B, then you would circle the "3" closest to Task A; if however, you thought they were equal, you would circle "0".

COMPARE TASKS FOR **PRIORITY – WHICH TASK WAS HIGHER PRIORITY?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

COMPARE TASKS FOR **INTEREST – WHICH TASK WAS MORE INTERESTING?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

COMPARE TASKS FOR **DIFFICULTY – WHICH TASK WAS MORE DIFFICULT?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

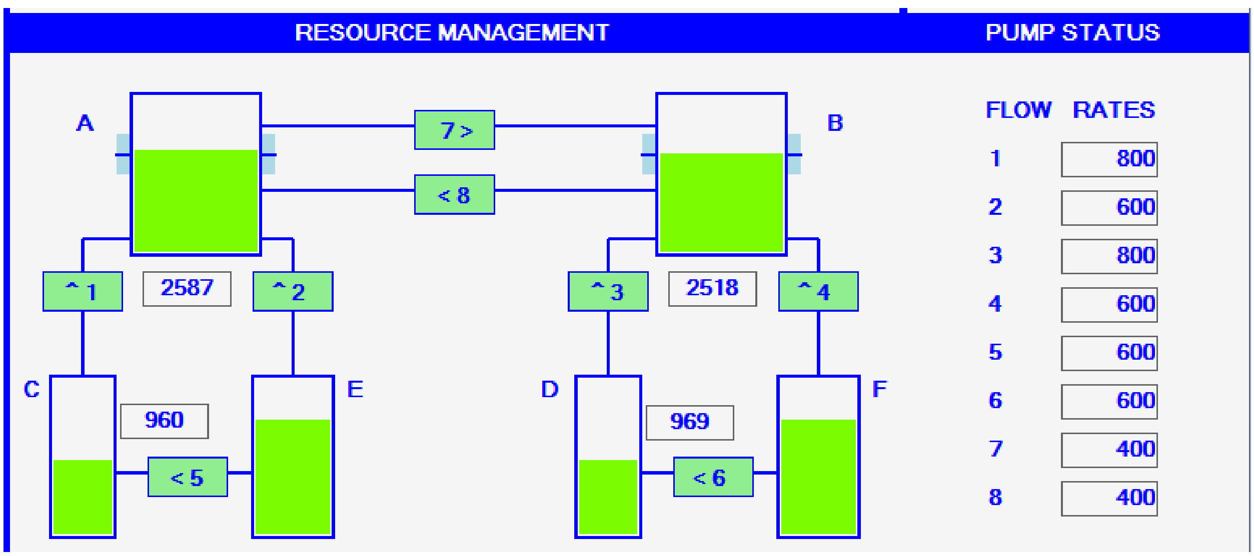
APPENDIX C: RESOURCE MANAGEMENT TASK INFORMATION AND TASK ANALYSIS

APPENDIX C: Resource Management Task Information and Task Analysis

All strategies assume a start point of a stable system (made instable by the continual depletion of tanks A and B, which must be kept in a target level of 2500 and visually on the line between shaded regions).

- **Pump rates used:**
- **1 = 800**
- **2 = 600**
- **3 = 800**
- **4 = 600**
- **5 = 600**
- **6 = 600**
- **7 = 400**
- **8 = 400**

Tank A and B both deplete at a rate of 800 each.



Maintain stability with no pump failure:

- Turn on pump 1 and pump 3 – generally stable, with occasional needs to turn on pumps 1 and 4 to keep it close to 2500 because tanks C and D will deplete too fast over a 10 minute experiment because pumps 5 and 6 replenish at 600, not 800.
- An over-full tank is a much less critical issue than an underfilled one; pumps can always be turned off, and pump failures will have less of an effect for their duration if they are not needed
- The most common state, however, is tank underfill

Reacting to pump failures

#1 pump fails (affects tank A)

- Make sure pump 2 is on to flow from tank E to A

- May need to turn on pump 8 and increase flow to tank B (pumps 3 and 4) to compensate if A is too low

#2 pump fails (affects tank A) – less critical than pump 1 failure in terms of return from LOW amounts

- Make sure pump 1 is on to flow from tank C to A (IF TANK NORMAL OR LOW)
- May need to turn on pump 8 and increase flow to tank B (from pumps 3 and 4) to compensate if A is too low
- Must make sure tank C does not go to 0 by turning on pump 5

#3 pump fails (affects tank B)

- Make sure pump 4 is on to flow from tank F to B (IF TANK NORMAL OR LOW)
- May need to turn on pump 7 and increase flow from tank A (IF TANK NORMAL OR LOW), and therefore need to turn on pump 1 or 2 or both depending on rate of pump 7

#4 pump fails (affects tank B) – less critical than pump 3 failure in terms of return from LOW amounts

- Make sure pump 3 is on to flow from tank D to B (IF TANK NORMAL OR LOW)
- May need to turn on pump 7 and increase flow from tank A (IF TANK NORMAL OR LOW), and therefore need to turn on pump 1 or 2 or both depending on rate of pump 7
- Watch tank D, make sure does not hit 0 by turning on pump 6

#5 pump fails (affects tank C and A)

- Make sure that the level in C is fairly high
- Make sure pump 2 is on, and only use pump 1 if necessary (and after compensating with pump 8 and adding fuel to tank B for a transfer

#6 pump fails (affects tank D and B)

- Make sure that the level in D is fairly high
- Make sure pump 4 is on, and only use pump 3 if necessary (and after compensating with pump 7 and adding fuel to tank A for a transfer)

#7 pump fails (affects tanks A and B)

- If B is below levels, and already pumps 3 and 4 are on, there is no other way to move fuel and rate of improvement is constrained

#8 pump fails (affects tanks A and B)

- If A is below levels, and already pumps 1 and 2 are on, there is no other way to move fuel and rate of improvement is constrained

APPENDIX D: COMMUNICATIONS TASK AUDIO ANALYSIS INFORMATION

APPENDIX D: COMMUNICATIONS TASK AUDIO ANALYSIS

Exp2 Trial	Order in Trial	Interrupted during phase	Interruption time (add to start time of int. event)	Com ID	Freq ID	Total clicks required	Time until audio start	Call1 time stamp	Call2 time stamp	Used	Radio start time stamp	Freq start time stamp	Audio end time stamp
1	1	Early	3	Com1	124.58	5	1.1	2.55	4.3	x	4.75	6.85	9.75
1	2	Middle	6	Nav1	114.45	3	1.91	3.33	5	x	5.48	7.47	10.35
1	3	Late	9	Nav2	113.55	2	0.35	1.45	2.7	x	2.98	5	7.26
1	4	Late	9	Com2	126.55	2	0.45	1.5	2.63	x	2.64	4.46	6.43
1	5	Early	1	Com1	125.55	3	0.54	1.72	2.9	x	3	4.88	7.52
1	6	Middle	5	Nav1	113.6	3	0.71	1.84	3.09	X	3.3	4.9	7.12
2	1	Middle	5	Com2	125.58	4	0.72	2.23	3.77	x	4.22	6.19	8.85
2	2	Early	2	Com1	127.55	3	0.49	1.53	2.68	x	2.76	4.4	6.57
2	3	Late	9	Nav2	110.5	2	0.42	1.51	2.88	x	3	4.8	7.05
2	4	Early	2	Nav1	115.4	5	0.28	1.36	2.43	x	2.48	4.11	6.07
2	5	Late	9	Com2	127.5	1	0.68	1.79	3	x	3.16	4.77	6.66
2	6	Middle	5	Nav2	113.5	1	0.81	2.27	3.83	x	4.42	6.38	9.5
3	1	Late	8	Nav1	111.5	1	0.48	1.6	2.68	x	2.75	4.44	6.46
3	2	Middle	5	Com2	128.55	4	0.33	1.43	2.52	x	2.57	4.4	6.53
3	3	Early	3	Com1	126.53	1	1	2.56	4.25	x	4.74	6.84	9.74
3	4	Middle	4	Nav2	111.65	4	0.45	1.55	2.82	x	3.06	5.01	7.37
3	5	Late	12	Com1	128.58	5	0.64	2.16	4	x	4.44	6.42	9.25
3	6	Early	1	Nav2	112.45	1	0.53	1.55	2.59	x	2.64	4.26	6.24
4	1	Middle	6	Nav2	111.4	3	1.75	3.15	4.74	x	5.2	7.79	10.62
4	2	Late	9	Com1	127.5	1	0.52	1.63	2.98	x	3.21	5.14	7.38
4	3	Early	1	Com2	124.45	4	0.47	1.58	2.66	x	2.67	4.39	6.71
4	4	Late	9	Nav1	114.65	5	0.5	1.69	2.91	x	3.1	5.02	7.25
4	5	Middle	4	Com2	129.58	6	0.53	1.72	2.87	x	3.22	4.89	6.68
4	6	Early	2	Nav1	113.5	1	0.38	1.4	2.56	x	2.62	4.17	6.19
5	1	Early	2	Com2	126.48	1	0.57	1.72	3.05	x	3.39	5.46	7.59
5	2	Late	9	Nav2	115.65	6	0.21	1.27	2.66	x	2.88	4.72	6.84
5	3	Middle	7	Nav1	112.45	1	1.59	3.01	4.58	x	4.99	6.89	9.58
5	4	Middle	6	Com1	126.45	2	0.88	2.16	3.41	x	3.75	5.25	7.3
5	5	Early	2	Com2	128.53	3	0.47	1.6	2.82	x	3.01	5	6.95
5	6	Late	9	Com1	128.48	3	0.93	2.15	3.22	x	3.29	4.89	6.72
6	1	Late	9	Com1	125.5	1	0.44	1.61	2.9	x	3.2	5.2	7.25
6	2	Early	1	Nav1	112.55	1	0.5	1.63	2.9	x	3.14	5.18	7.47
6	3	Middle	5	Com1	129.45	5	0.44	1.55	2.8	x	3.09	4.81	7.17
6	4	Early	1	Com2	127.53	2	0.89	2.31	3.97	x	4.45	6.47	9.35
6	5	Middle	8	Nav1	110.65	4	1.76	3.22	4.71	x	5.22	7.2	10.13
6	6	Late	8	Nav2	114.45	3	0.58	1.66	2.73	x	2.78	4.44	6.48
		Middle	5	Com2	125.5	1	0.59	1.69	2.87	no	3.22	4.77	6.81
		Late	9	Nav2	112.6	2	0.77	1.87	3.11	no	3.33	5	6.97
		Early	1	Nav2	114.6	4	0.77	1.96	3.22	no	3.36	5.18	7.14
		Late	9	Nav1	111.6	3	0.33	1.43	2.74	no	2.89	4.65	6.87
		Middle	5	Nav2	114.45	3	0.58	1.66	2.73	no	2.78	4.44	6.48

APPENDIX E: MATB INSTRUCTIONS (USED IN EXP 2 and EXP 3)

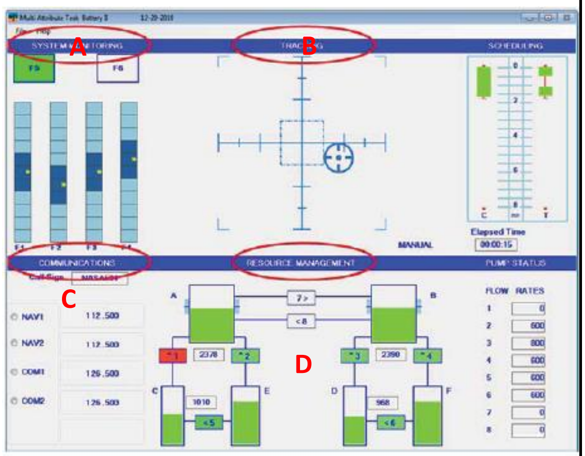
Learning how to perform MATB II

(APPENDIX E: Experiment 2 Instructions)

Welcome

- Review the following materials.
- At the end of the time, you’ll be asked to actually perform these tasks in the simulation – so pay attention! No one can help you after the training period is over.

This is an overview of the main display that you will see during the experiment.

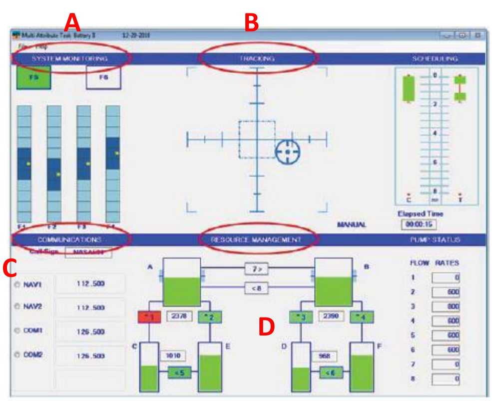


These four tasks simulate tasks that REAL pilots perform during flights of commercial jetliners. These tasks are *critical* to master, as you will be asked to perform all of them.

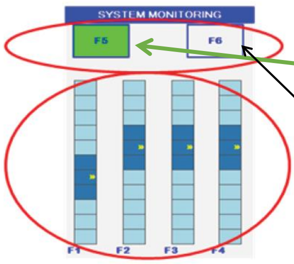
These tasks include:

a) monitoring,
b) tracking,
c) communicating,
d) and resource management

Each of these tasks has been labeled – take a moment to find them on the main display



System Monitoring Task (task A)



The monitoring task appears in the upper left of the main display. At the top, two squares are shown.

The left square at the top will normally appear green; however, it may appear white if it fails. **You should respond quickly by clicking the square to turn the system back on.**

The right square at the top will normally appear white; however occasionally it may turn red indicating a failure. **You should respond by clicking the square to acknowledge the failure.**

Thought all monitoring events are important, the RED on is especially critical to detect.

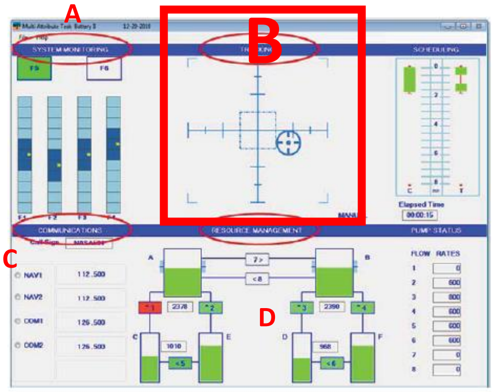
You will only have a brief time to attempt to respond to these events.

System Monitoring Task (task A)

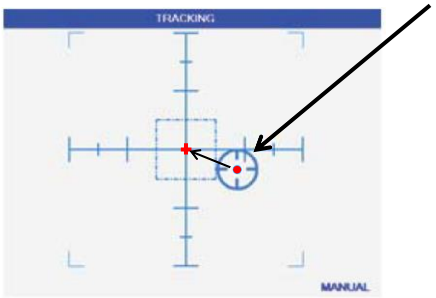
The rest of the display shows 4 sliding scales, each of which has an indicator in blue. Each indicator can go "out of range" on each sliding scale and stick at the top, or the bottom for a few seconds.

Your job is to detect these deviations by clicking on the correct scale when a deviation occurs. Remember, you are to *quickly*, and *accurately*, respond to these events, which may or may not occur.

Task B = Tracking Task



Tracking Task (Task B)

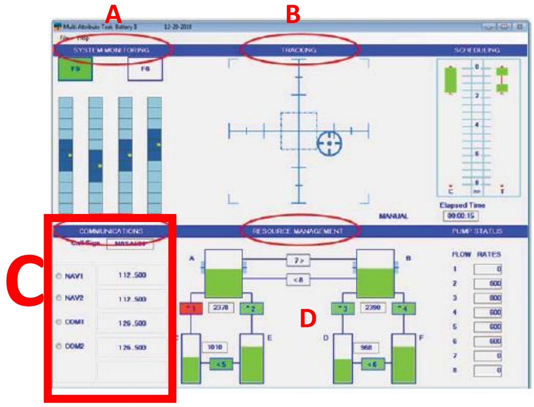


For this task, you will use the joystick to control a small, circular target reticle.

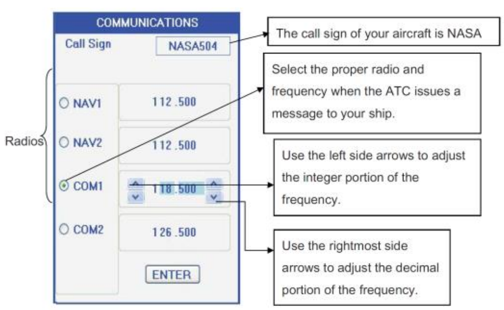
Your job is to use the joystick to keep the dot (shown in red here) on the cross point in the center (shown in red) as much as possible.

The reticle will move erratically, so use the joystick continuously to move it back on target. You can think of this like piloting a plane under turbulence; weather will alter the aircraft's path – your job is to keep on course.

Task C = Communications Task



Communication Task (Task C)



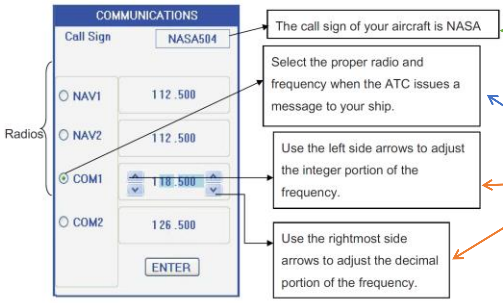
You will be listening for messages through your headphones.

When you hear a message it will contain two parts; the first part is the callsign, the "name" of the aircraft.

The call sign will be repeated – listen for "NASA504" – this is your callsign.

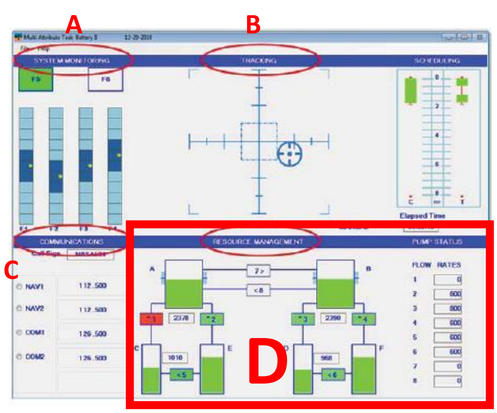
The second part is a command requiring you to tune one of the 4 radios to a particular frequency.

Communication Task (Task C)

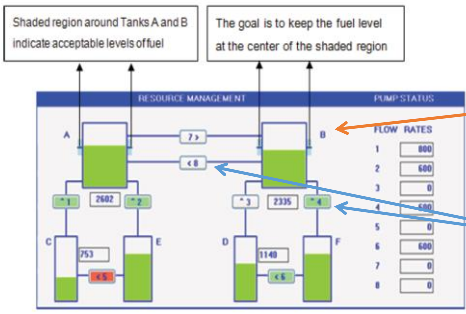


1) Determine whether the message is intended for you: does it match your call sign (NASA504)?
2) If so, quickly and accurately select the radio, and adjust it to the frequency you were instructed to, using the mouse.
3) Click "ENTER"

Task D: Resource Management Task



Resource Management Task (Task D)

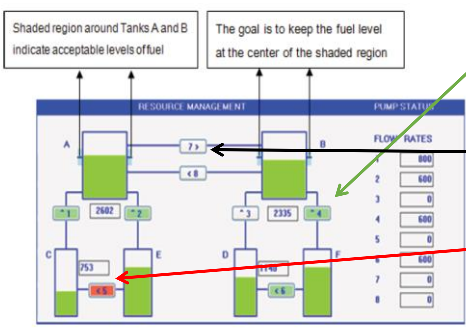


The resource management task represents managing fuel aboard an aircraft (shown in green).

Each fuel container is represented with a letter (A through F). Lines connect certain containers with other containers, and the flow of fuel between them is controlled by a pump.

Each PUMP is numbered (1 through 8).

Resource Management Task (Task D)



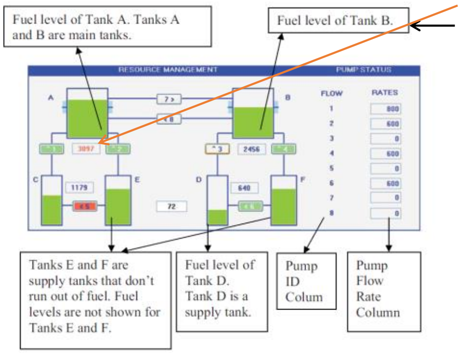
Pumps are color-coded for their state:

\\ green: pump is ON and flowing in the direction of it's arrow; it can also be turned off by clicking on it

\\ White: pump is OFF, but can be turned on if needed by clicking on it

\\ Red: pump has FAILED, BUT you MUST continue to keep the center fuel levels (in A and B) at the center of the shaded region.

Resource Management Task (Task D)

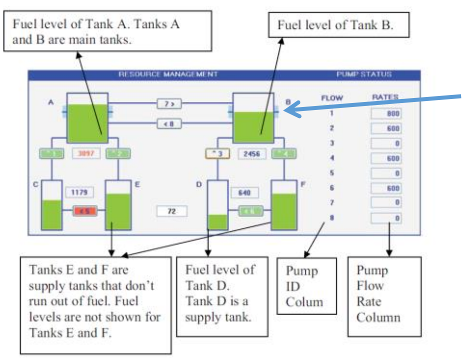


Numbers underneath each container represent the amount of fuel in the container. These numbers decrease and increase, corresponding to the fill level.

THE TASK:

- Tanks A and B continually use fuel to keep the engine going, so fuel is always depleting
- You must *transfer fuel to tanks A and B*
  - via transfers from C, D, E & F.
  - Tanks E and F are unlimited
  - Tanks C and D are supply tanks, and contain a maximum of 2000 units, and can also go to 0 making them useless.

Resource Management Task (Task D)

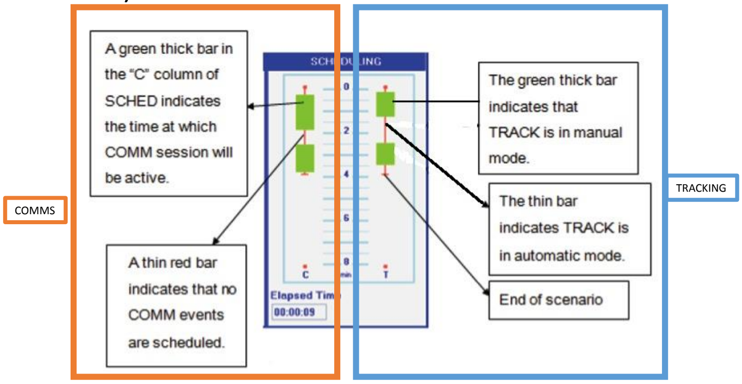


The **goal of the task** is to keep containers A and B filled within the range indicated by the shaded regions (2500 units).

There are multiple strategies to achieve this goal; you may use whatever method works best.

A pump failure of 1 or 3 is especially critical to respond to immediately.

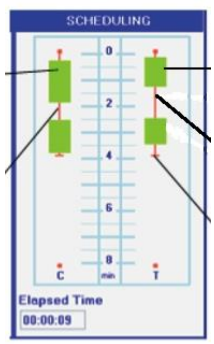
Task Scheduler Display (upper right corner)



Task Scheduler Display (upper right corner)

- As shown, the task scheduler may help you allocate your attention to upcoming tasks (it only shows the tracking task, and the comms task).

- It does not show upcoming resource management events, or upcoming monitoring task events.



Putting it all together

- You will be performing all of the tasks simultaneously during the upcoming trials.
- Your goal is to be as fast and as accurate as possible in ALL OF THE TASKS. Spread and balance your attention between them to accomplish this as effectively as possible.
- You will complete a basic training trial next, before beginning test trials.
- Feel free to return to any earlier slides if the time has not run out.

APPENDIX F: PAIRED TASK COMPARISON SURVEY (USED IN EXP 2 AND EXP 3)

APPENDIX F: SURVEY FOR EXP. 2 and EXP. 3.

Participant Number = ___________

Date and Time = _____________________

Computer Number used = ___________

DO NOT OPEN UNTIL INSTRUCTED TO DO SO

Paired Task Survey

Based on the trial you just completed , you should compare several tasks to each other by rating which task represents the category more. For example, you might be asked to compare two tasks (A and B) on which was a higher priority to perform.

Task A ←*--higher task priority--lower -- 0 (equal) -- lower --higher task priority*→Task B

So, if you thought Task A was *higher priority* than Task B, then you would circle the "3" closest to Task A; if however, you thought they were equal, you would circle "0".

COMPARE TASKS FOR **PRIORITY – WHICH TASK WAS HIGHER PRIORITY?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

COMPARE TASKS FOR **INTEREST – WHICH TASK WAS MORE INTERESTING?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

COMPARE TASKS FOR **DIFFICULTY – WHICH TASK WAS MORE DIFFICULT?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

STOP HERE FOR NOW UNTIL INSTRUCTED TO CONTINUE

Paired Task Survey

Based on the trial you just completed , you should compare several tasks to each other by rating which task represents the category more. For example, you might be asked to compare two tasks (A and B) on which was a higher priority to perform.

Task A ←*--higher task priority--lower -- 0 (equal) -- lower --higher task priority*→Task B

So, if you thought Task A was *higher priority* than Task B, then you would circle the "3" closest to Task A; if however, you thought they were equal, you would circle "0".

COMPARE TASKS FOR **PRIORITY – WHICH TASK WAS HIGHER PRIORITY?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

COMPARE TASKS FOR **INTEREST – WHICH TASK WAS MORE INTERESTING?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

COMPARE TASKS FOR **DIFFICULTY – WHICH TASK WAS MORE DIFFICULT?**

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Tracking Task

Monitoring Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Communications Task

Tracking Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

Communication Task ---- 3 -------- 2 -------- 1 ----0---- 1 -------- 2 -------- 3 ---- Resource Management

A few more questions:

1) Do you have any experience with flight simulators? If so, please estimate the number of hours per week (and total hours spent)

2) Do you have a pilot's license, or any recorded hours flying an aircraft of any kind? If so, please describe (and list the number of hours you have flown)

3) Do you spend any time playing real-time strategy games, like Starcraft, Command and Conquer, and/or Civilization? If so, please list which games and about how many hours a week (and over your lifetime) you would estimate you've played.

You're done with the survey! We have one more task for you to complete. Please await instructions….

APPENDIX G: COUNTERBALANCING AND EXTENDED INTERPRETATION (FROM EXPERIMENT 1)

APPENDIX G: COUNTERBALANCING / INTERPRETATION FROM EXP 1

Analysis 1: Tracking task performance

A 2 (priority condition: EQP or TRKP) x 2 (tracking difficulty: easy or difficult) x 2 (counterbalance: easy or difficult trial performed first) repeated measures ANOVA was run on overall averaged error in the tracking task. A significant interaction between tracking difficulty and counterbalancing condition also emerged ($F(1,75) = 7.97$, $p < .01$, $\eta_p^2 = .10$), such that easy-first performance resulted in less error on the difficult tracking trial ($M = 40.15$) compared to difficult-first ($M = 47.90$). Error on easy trials did not differ between counterbalance conditions (Easy-first = 21.76, Difficult first = 21.26). No other significant interactions or main effects were found. The magnitude of the difficulty effect depended on which trial was performed first, suggesting task practice may exert an influence on tracking performance, but only when tracking was difficult.

Analysis 2: Task performance

An exploratory 2 (priority) x 2 (difficulty) x 3 (task type) x 2 (counterbalance) repeated measures MANOVA was conducted using resource error, and accuracy in the monitoring and communications tasks. Only one interaction between tracking difficulty condition and counterbalancing condition was significant in the multivariate analysis (Wilks' $\lambda = .675$, $F(3,73)=11.70$, $p < .001$). With sphericity assumptions met, univariate analyses showed the interaction was only significant for the resource management task ($F(1,75) = 29.02$, $p < .001$, $\eta_p^2 = .28$), but was not significant for communications ($F<1$) or monitoring ($F(1,75) = 2.90$, $p = .09$, $\eta_p^2 = .04$). Follow up paired t tests confirmed resource error under easy tracking was greater ($M= 594$, $SD= 496$) than error under difficult tracking ($M= 437$, $SD= 356$) for easy-first counterbalancing conditions ($t(33) =$

2.92, $p < .01$). For the difficult-first condition, the pattern was reversed with less error in easy ($M = 404$, $SD = 397$) compared to difficult tracking ($M = 697$, $SD = 560$; $t(44) = -5.03$, $p < .001$). This seems to be an effect of training; for both levels of tracking difficulty, less error in resource management was present when participants completed a prior trial that included the Rman task.

Analysis 3: Task switching overall

A 2 (tracking difficulty) x 2 (priority group) x 2 (counterbalance condition) repeated measures ANOVA was conducted on number of switches overall. Tracking difficulty and counterbalancing condition significantly interacted, $F(1,75) = 6.30$, $p = .01$, $\eta_p^2 = .08$). In the easy-first condition, there were significantly fewer switches under difficult tracking ($M = 40.88$) than under easy ($M = 49.56$; $t(33) = 5.17$, p < .001). This difference was *not* significant in the difficult-first counterbalance condition between easy ($M = 47.84$) and difficult tracking ($M = 45.64$) switches ($t(44) = 1.17$, $p > .05$). This interaction could be evidence of a fatigue effect or a learning effect.

Analysis 4: Time spent in tracking task

A 2 (priority condition) x 2 (difficulty condition) x 2 (counterbalance condition) repeated measures ANOVA was run on the summed time spent in the tracking task on each trial. A significant interaction was revealed between difficulty and counterbalancing. A significant increase in time spent in difficult tracking trials ($M = 532$s) compared to easy ($M = 517$s) was found, but only for the easy-first condition ($t(33) = -4.69$, $p < .001$). The difference between easy and difficult time in tracking was not significant in the difficult-first condition ($t(44) = -.403$, $p = .69$).