Spring 2019

Comparing the Effects of Mental Workload Between Visual and Auditory Secondary Tasks During Laparoscopy

James P. Corcoran
Old Dominion University, jcorc006@odu.edu

Recommended Citation

Corcoran, James P.. "Comparing the Effects of Mental Workload Between Visual and Auditory Secondary Tasks During Laparoscopy" (2019). Master of Science (MS), Thesis, Psychology, Old Dominion University, DOI: 10.25777/kv21-7v83
https://digitalcommons.odu.edu/psychology_etds/230

COMPARING THE EFFECTS OF MENTAL WORKLOAD BETWEEN VISUAL AND AUDITORY SECONDARY TASKS DURING LAPAROSCOPY

by

James P. Corcoran
B.A. May 2008, College of William and Mary
B.S. December 2014, Old Dominion University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY

May 2019

Approved by:

Mark W. Scerbo (Director)

Matt Judah (Member)

Yusuke Yamani (Member)

ABSTRACT

COMPARING THE EFFECTS OF MENTAL WORKLOAD BETWEEN VISUAL AND AUDITORY SECONDARY TASKS DURING LAPAROSCOPY

James P. Corcoran
Old Dominion University, 2019
Director: Dr. Mark W. Scerbo

The purpose of this study was to test Wickens' Multiple Resource Theory (MRT) by comparing performance and subjective workload on a visual-spatial secondary task with an auditory-spatial analog when paired with visual-spatial laparoscopic primary tasks. Two primary tasks were performed with a laparoscopic box trainer: a high workload task that consisted of transferring rings from one peg to another and a low workload task that consisted of grasping and placing large pencil erasers in a bowl. It was predicted that the visual-spatial secondary task would be more sensitive when paired with the laparoscopic primary task than the auditory analog. Findings from the study mostly supported this prediction. Proportion of correct detections and subjective workload scores indicated that the auditory-spatial task secondary task was less demanding than the visual-spatial task in high workload, dual task conditions. However, no significant differences were found for response time and false alarms. Overall, these results support the modality predictions of MRT under high workload conditions. Additionally, this study provides further evidence supporting the use of the visual-spatial, ball-and-tunnel task as a measure of workload during laparoscopic surgery.

In memory of my father, who taught me the importance of sticking things out regardless of how difficult circumstances may become.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Mark Scerbo for sticking with me and guiding me through the ups and downs of this process. I would also like to thank Dr. Chris Brill for allowing me to get my start in Human Factors in his lab. Last, I would like to thank Dr. Barbara Winstead for providing sound advice during some of the more difficult times.

TABLE OF CONTENTS

Page

Chapter

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

One of the primary areas of interest for human factors researchers is the relationship between task performance and cognitive resource consumption. In this case, cognitive resources may be understood as representations of working memory capacity. As attention and effort are required to perform a task, some portion of working memory is allocated to that end. When working memory is taxed beyond its limit, task performance declines. Over the years, several researchers have attempted to explain specifically how variations in task demand affect performance and describe the mechanisms underlying these processes. One approach that has come to the forefront of performance oriented human factors research is Multiple Resource Theory (MRT; Wickens, 1980; 2002; 2008). MRT posits that cognitive resources are drawn from separate resource pools depending on specific task attributes. The more that these attributes overlap, the greater a demand is placed on a single resource pool. This, in turn, can have a negative effect on performance.

Multiple Resource Theory

Wickens (1980) investigated the effects of time-sharing on multitasking performance. The primary focus of his initial 1980 study was to evaluate two existing explanations for the effects of time sharing: the hemisphere of processing approach and the modality of processing approach. The hemisphere of processing approach was based on the notion that resources were divided into the two hemispheres of the brain by the functions performed by unique anatomical areas (Kinsbourne & Hicks, 1978). In contrast, the modality of processing approach was based on the idea that resource pools are divided into discrete categories by encoding and response modality. Wickens (1991) argued against the inclusion of the hemisphere approach for two reasons. First, although there was clear evidence for resource distribution between hemispheres, the division of resources could not be considered orthogonal, meaning that there was some degree of overlap. Second, the hemisphere approach treated each hand as a separate resource-defined response channel. This would mean that a response involving both hands would consume resources from two independent resource pools. Instead, Wickens (1991) argued that a code-based approach in which the use of both hands drew from a single spatial resource system better described the division of resources.

Wickens' (2002) current multiple resource model is comprised of four dimensions: stages, modalities, channels, and codes. In the model each dimension contains two discrete levels. For the stage dimension, the first level is the perceptual/cognitive stage at which resource consumption is due to processing and organizing information. The second is the response stage at which the execution of a decision is the primary cause of resource expenditure. This suggests that perceiving and organizing information relies on a separate resource pool than responding. However, it should be noted that a decrement in performance may occur when elements of the perceptual and response stages overlap (Liu & Wickens, 1992).

The second dimension is comprised of a dichotomous split between the auditory and visual modalities. MRT predicts that cross-modal time-sharing results in better task performance than intramodal time-sharing (Wickens, 1993). The underlying assumption of the theory is that each modality has its own resource pool. Therefore, performance should be better when time-sharing tasks are split between the modalities because the tasks are drawing from two separate resource pools.

The visual modality is divided further into two levels consisting of a focal channel associated with foveal vision and an ambient channel for processing information in the periphery. As would be expected for foveal vision, the focal channel is primarily used to process detailed information and patterns to which an individual is attending. The ambient channel is primarily used for orienting and movement through an environment (Weinstein & Wickens, 1992). The separation between these two channels is most apparent in a human's ability to move about a space while simultaneously attending objects within that space. The channel dimension is also of importance when evaluating MRT because it explains how two visual tasks may be performed simultaneously with minimal effects on performance.

The fourth dimension, coding, may also be understood in terms of processing methods. Codes are organized dichotomously into spatial and verbal categories. Spatial coding is reserved for tasks related to location and distance while verbal coding is associated with linguistic processes. Codes are important because they account for performance effects when there is minimal overlap among the stages, modalities, and channels dimensions. For example, it is difficult to read and listen to a person talk simultaneously without sacrificing performance on one or both tasks. Although one is a visual task and one is auditory, both are verbal. When evaluating MRT it is important to select task codes to ensure that the correct comparison is made. Performance outcomes from the combination of a visual-spatial task and an auditory-verbal task would likely be significantly different from those of a combination of visual-spatial and auditory spatial task.

Since the introduction of MRT, several researchers have used the theory to predict outcomes in performance in studies requiring shared attention (Burke et al., 2006; Crawford, Watson, Burmeister, & Sanderson, 2002; Sarter, 2007). Most of these studies have dealt with the design and use of multimodal displays. In one example, Crawford and colleagues (2002) evaluated time-sharing performance using multimodal displays in an anesthesiology context. They had anesthesiologists monitor common operating room scenarios and respond to “probes” in which they verbally reported information about different indices regarding the patient’s status. In one condition the physiological information was completely sonified (i.e., conveyed through changes in the acoustic characteristics of the signal), while in three other conditions one or more devices required visual attention. In the conditions with visual displays, participants had to shift their visual attention between the patient and the visual display. However, in the sonified condition, visual attention remained directed toward the patient. The results indicated that the anesthesiologists performed better when they monitored the patient visually and physiological status information was presented using the sonified auditory display.

Burke and colleagues (2006) performed a meta-analysis using 43 studies of multimodal feedback to see how well the results of multimodal display research fit the predictions of the MRT framework. In general, the researchers found that the addition of auditory or tactile feedback to a visual task improves performance. The conclusion that multimodal feedback leads to better performance across a wide range of studies provides support for the predictive ability of MRT. However, the authors also suggested more research was needed to better understand the effects of varying levels of task demand on performance during multimodal tasks. In addition, this meta-analysis did not differentiate between types of auditory displays leaving exploration of that topic to future research. In addition to expanding knowledge surrounding MRT, some researchers have compared the theory with competing ideas.

Numerous studies have provided support for MRT and it is typically considered the preeminent theory concerning the division of cognitive resources in human factors applications

(Burke et al., 2006; Sarter, 2007). However, there have been alternative theories. Paulson and Friedman (1988) challenged MRT with undifferentiated resource theory (URT). Boles and Law (1998) compared MRT and URT and their conclusions mostly supported MRT taking issue only with its dichotomous structure. One way to further test MRT is to evaluate each of its dimensions through constructs for which well-defined metrics already exist. One construct of importance concerning the multiple resource model is workload.

MRT and Workload

MRT is not a theory of workload; however, workload is a key factor when accounting for the resource expenditure side of Wickens' (2008) model. Workload as it applies to MRT may be defined as a construct which describes the relationship between task demand and cognitive resource depletion (Hart & Staveland, 1988). When portions of a task or multiple tasks are divided across the multiple dimensions of the model, MRT predicts that cognitive demand will be distributed among different types of information processing. In turn, this reduces the total demand on any one processing area thereby reducing workload. This reduction in workload may account for improvements in performance (Wickens, 2008). This means that workload is a primary means of explaining how MRT is used to make predictions about performance.

There are several methods for measuring workload; however, the efficacy of a particular measure should be determined using several criteria (Carswell, Clarke, & Seales, 2005; O'Donnell & Eggemeier, 1986). The first of these criteria is sensitivity. For a measurement to be considered adequately sensitive it must be able to differentiate among varying levels of workload. By possessing this property, a sensitive measure allows researchers to isolate and compare the resource demands associated with particular sets of tasks.

A second characteristic is diagnosticity. A measure must match the particular resource pool associated with a specific task. Furthermore, researchers should avoid using techniques that involve measurement across multiple resource pools. Failing to do so may result in an inability to differentiate between specific resources and global mental workload. Another related criterion is selectivity. This refers to the ability of a measure to remain unaffected by extraneous variables. This property may also be understood as resistance to the effects of confounding variables (Carswell, 2005).

Intrusiveness refers to the degree to which a measurement technique interferes with the primary task. Similarly, ease of use must be considered. Ease of use refers to the amount of resources (i.e., time, money, training) needed to obtain the measurement. The final criterion is operator acceptance. It is important that participants are comfortable with the measurement technique involved. As with most experimental criteria there is no single, optimal combination that may be viewed as a standard for use. Instead, the relative importance of these criteria should be considered when evaluating or comparing different measurement techniques or instruments for specific research questions.

Approaches to measuring workload which can be divided into three classes: psychophysiological, performance-based, and subjective. Psychophysiological measures of workload record changes in some aspect of a subject's physiological state in response to task demands (O'Donnell & Eggemeier, 1986; Wierwille & Connor, 1983; Young & Stanton, 2004). Subjective workload measures are generally presented in the form of surveys or rating scales in which participants are asked report the level of demand they felt while performing the task (Reid & Nygren, 1988; Young & Stanton, 2004). The primary advantage of subjective workload measures is that they are easily administered and can be used to measure workload across a variety of tasks (Stanton, Salmon, Walker, Baber, & Jenkins, 2005).

One of the most commonly used subjective workload measures among human factors researchers is the National Aeronautics and Space Administration Task Load Index (NASA-TLX; Hart, 2006; Hart & Staveland, 1988). The NASA-TLX is comprised of six subscales (mental demand, physical demand, temporal demand, effort, performance, and frustration level) scored on a scale of 0 to 20. Since its adoption, the NASA-TLX has provided researchers with a flexible, inexpensive, measure of workload that can be administered with minimal training (Hart, 2006).

The last category of measurement is based on task performance. Performance-based measures are divided into two categories: primary and secondary. Primary task measurements record an individual's performance during a task of interest. These performance measurements should directly reflect the level of demand placed on the participant by the primary task alone. Task demand is often a reflection of task difficulty or complexity. However, it may also be manipulated in other ways such as imposing time constraints for task completion. On the other hand, secondary task measurements are often used in conjunction with primary tasks. Secondary task performance is an index of workload obtained through the addition of another task performed simultaneously with the primary task (O'Donnell & Eggemeier, 1986).

The theoretical basis for the use of secondary tasks relies on the concept that primary tasks require a certain amount of cognitive resources. The cognitive resources that remain unused are then available to an individual for performing an additional task (Eggemeier, Wilson, Kramer, & Damos, 1991). Demanding primary tasks should consume more resources thereby leaving little in reserve for additional tasks. The inclusion of a secondary task should have a measurable effect on performance. These changes in performance may be observed in one of two ways. For the loading method, an individual must maintain his or her level of secondary task performance while remaining indifferent to possible effects on primary task performance. Alternately, the subsidiary method requires an individual to maintain his or her level of primary task performance without regard for possible effects on secondary task performance (Wierwille & Eggemeier, 1993).

Secondary tasks may also be independent or embedded. Independent, secondary tasks are those not typically associated with the performance of the primary task. If carefully selected, an independent, secondary task can be highly sensitive to differences in workload. However, it is important to consider the intrusiveness of this type of task. A highly intrusive, secondary task may create additional demand that could introduce an artificial performance decrement. In contrast, an embedded secondary task is one that is performed as a normal part of the primary task procedure. For example, a radio communication task added to a flight simulation would be considered embedded because maintaining radio contact is a normal part of operating an aircraft. An advantage of using an embedded, secondary task is that it is minimally intrusive. However, because this kind of task is typically performed with the primary task, it may not be sensitive enough to provide a useful measure of workload.

MRT and Laparoscopic Surgery

The multiple resource model has become a common feature for many applied, multitask studies; however, given the complexity of the topic area such testing requires the use of a specific, well-defined context such as laparoscopic surgery. A number of studies have been conducted to assess mental workload in the context of laparoscopic surgery (Britt et al., 2015; Prytz et al., 2012; Scerbo et al., 2013, Stefanidis et al., 2007) relying very heavily on the modality predictions of MRT. These studies of laparoscopic surgery provide one way to assess the multiple resource model directly.

Laparoscopic surgery is a form of surgery in which a surgeon makes several small incisions. A camera is inserted into one of the incision sites and the images from the camera are projected onto a screen. The surgeon then inserts instruments into the other incision sites to perform the procedure. The small incisions can greatly reduce recovery time because less of the patient's body is directly affected by the procedure and may also reduce complications following the surgery (Braga, et al., 2005, King, et al., 2005). For these reasons, laparoscopy has become a popular approach compared to open surgeries. However, laparoscopy also presents a number of challenges because it is a less direct method for performing surgery (Berguer, Smith, & Chung, 2001). Visual challenges arise because the surgeon is working from a projected image instead of direct line-of-sight (Cuschieri, 1995, 2006; Tendick, Bhoyrul, & Way, 1997). Orientation issues occur because the camera is inserted through an incision and its path to the operating site is not the same as the surgeon's natural point of view (Conrad et al., 2006; Gallagher et al., 2009; Klein, Warm, Riley, Matthews, & Parsons, 2004). The loss of tactile information makes the procedure more difficult because the surgeon is less able to sense whether he or she is applying the appropriate amount of pressure and must rely more on visual cues, which as noted above, are distorted compared to natural viewing conditions (Mohr et al., 2001). These issues make laparoscopy highly demanding and therefore a high workload task (Prytz et al., 2012, Scerbo et al., 2013).

Because laparoscopic surgery is a visual-spatial task, researchers determined that the secondary task should overlap on the spatial dimension of the multiple resource model. Stefanidis, Scerbo, Korndorffer, and Scott (2007) developed a task in which participants had to monitor squares presented briefly on a laptop while simultaneously completing a suturing task presented on another visual display. Participants were required to respond with a foot pedal every time three squares were presented on the right side of the laptop screen. Results indicated that the squares task was sensitive to workload on the dimensions of interest. However, the use of two displays may have added additional difficulty by requiring participants to shift their gaze away from the primary display.

To improve the measurement of visual-spatial workload, Scerbo and colleagues (2012; Prytz et al., 2012) devised a new secondary task that required individuals to monitor the positional changes of a set of balls projected over a laparoscopic primary task. The “ball-and-tunnel” task uses four different colored balls presented in a simulated three-dimensional tunnel. The balls can rotate clockwise or counter-clockwise and move closer or farther down the tunnel. Participants are required to respond when one of the balls changes orientation by pressing a foot pedal. Another advantage of the ball-and-tunnel task was that it eliminated the need for a participant’s gaze to shift between two displays because the secondary task was projected onto the primary task display. This increased the overlap in resource demand because both tasks address the same focal-visual resource pools simultaneously.

Prytz and colleagues (2012) conducted a follow-up study confirming the efficacy of this new technique. For this study, participants were asked to complete the ball-and-tunnel task alone and then concurrently with three laparoscopic surgical tasks. The researchers found that performance on the ball-and-tunnel task declined significantly when paired with the laparoscopic primary tasks. In addition, there were significant differences in performance on the ball-and-tunnel task depending on the difficulty of the individual laparoscopic tasks. These findings support earlier work indicating that the ball-and-tunnel task causes significant overlap in the focal-visual dimensions of the multiple resource model. In addition, differences in secondary task performance provided evidence that the ball-and-tunnel task was sensitive to difference in primary task workload.

Britt and colleagues (2015) had participants perform laparoscopic suturing on simulated bowel and on a cadaver bowel using the ball-and-tunnel task to measure workload. They found that completion times were longer and performance on the ball-and-tunnel task was significantly lower when the suturing task was performed on the cadaver as opposed to the simulation. These findings suggest that the ball-and-tunnel task is a reliable measure of workload for laparoscopic surgery and may provide a useful platform from which analogs can be created to assess mental workload across the other modalities. In doing so, comparisons against an established visual-spatial task may help clarify some of the ambiguities of MRT involving cross-modal performance.

MRT and Auditory Processing

According to MRT, dividing tasks across modality may decrease mental workload by reducing demand on a particular channel and thereby reducing the degree to which a resource pool is depleted. One of the more common approaches is the use of auditory displays which can reduce the demand on the visual channel. This shift in modality may help to declutter primarily visual workspaces and reduce mental workload by spreading task demands across different modalities (Wickens, 2008).

One gap in current human factors literature concerns investigation of the relationship between auditory-spatial processing and workload. With respect to MRT, this gap raises questions regarding how to predict performance on auditory-spatial tasks compared to tasks on other dimensions and makes its role in the multiple resource model somewhat ambiguous despite its conceptual simplicity. To examine the effects of visual and auditory-spatial displays on workload, two important questions arise. The first is whether visual-spatial and auditory-spatial displays can convey the same information. The second is what limitations may exist concerning the psychophysical equivalence of information presented on both displays.

Brown, Newsome, and Glinert (1989) addressed these initial questions by conducting a study to test the effectiveness of auditory and visual cues during a task in which participants had to detect the presence or absence of a target. For the both tasks, participants were required to locate a specific target among 30 distractors. In the visual condition, participants were presented with a cue regarding the location of the target on another screen. For the auditory condition, participants were trained to associate specific sounds with columns on a screen. These sounds provided the auditory analog of the visual cue. Their findings indicated no difference in performance between modalities, but that auditory signals took more time to process than visual signals. These results suggest that auditory information could be used to replace visual information during some visual tasks as long as the speed of responding is not paramount. These findings also suggest that an auditory analog of a visual display could potentially be used to facilitate performance.

The findings of Brown and Glinert (1989) provide some evidence that the visual and auditory modalities are comparable, but it is important to assess the extent to which splitting modalities affects workload. MRT posits that the visual and auditory channels rely on separate pools in multitask conditions (Wickens, 1980; 2002; 2008). Several studies using multimodal displays in applied settings have investigated the extent to which research supports this theory. A study by Bronkhorst, Vetman, and Van Breda (1996) investigated the effects of adding a three-dimensional, auditory display to assist in target acquisition and pursuit during a flight task. Participants in all conditions worked with a three-dimensional, visual display which provided information about incoming targets at multiple distances within a limited field of view. The participants in the three-dimensional auditory condition wore headphones which presented a warning sound originating from the direction of the incoming target. The results showed that inclusion of a three-dimensional auditory display significantly reduced search time when compared to the three-dimensional, visual display alone. Following MRT, if auditory displays can replace visual displays for certain tasks, then it follows that the auditory and visual channels could be used in conjunction with minimal increase to workload.

Oshima and Wickens (1992) tested whether redundant, spatial auditory cues could improve flight performance and reduce workload. Their findings indicated no significant improvement, but they concluded that the absence of effects was due to the limitations of the audio equipment at the time. However, a later study by Begault (1993) on the effectiveness of simulated, auditory-spatial cues on target acquisition had a different result. This investigator found that the use of three-dimensional auditory displays improved participants’ acquisition times for targets. Studies by Mckinley and Ericson (1997) and Pavlovic, Keillor, and Hollands (2009) in which spatial auditory cues were added to conventional heads-up displays in cockpits showed improved performance and decreased mental workload. These findings suggest that the inclusion of spatial audio can reduce workload in applied settings such as flight tasks. However, it should be noted that in the above studies, the auditory cues were used in conjunction with visual displays. Therefore, the effect was more that of reinforcement than a comparison of how the use of different modalities affects performance. Collectively, these findings suggest that shifting a portion of a task from the visual to the auditory modality can improve performance and reduce workload with visually cluttered or complex environments. However, while there is clear support for MRT for dual task conditions in which one task is auditory and one is visual, little research has examined the effects of sound localization in and of itself on mental workload. MRT would predict that performance on a visual-spatial task should be significantly worse and mental workload would significantly higher when combined with another visual-spatial task than with an auditory-spatial task.

Proposed Study

The purpose of the present study was to test MRT by comparing performance on a visual-spatial secondary task with that of an auditory-spatial secondary task under high and low workload conditions. The theoretical goals of this study were to explore and clarify the applicability of MRT and to compare visual and auditory sensitivity concerning secondary tasks in relation to laparoscopic surgery. The direct comparison of a well-established, visual-spatial task to an auditory-spatial task aided in determining whether the multiple resource model can be used to predict auditory-spatial performance. Given the predictions of MRT, auditory secondary tasks should be less sensitive than visual tasks when paired with a high workload, visual laparoscopic primary task. This outcome would evince the demanding visual-spatial nature of laparoscopic surgery. In addition, the relationship between sound localization and workload has not been fully explored.

For this study, participants completed high and low workload primary tasks. They also completed visual and auditory secondary tasks independent of the primary tasks and then in conjunction with each primary task. The secondary tasks performed independently served as control conditions for assessing primary task workload. Performance was measured using the proportion of correct responses, d’, and response time.

Hypotheses

Previous studies have found that splitting complex tasks between the auditory and visual channels can reduce workload and improve performance supporting Wickens' (2008) multiple resource model (Begault, 1993; Jeon, et al.,2015, Liu, 2009; Mckinley & Ericson,1997; Pavlovic, Keillor, & Hollands, 2009). In comparing performance outcomes with visual-spatial and auditory-spatial secondary tasks when each was paired with a visual-spatial primary task, the model would predict the auditory-spatial task to be the least sensitive of the two secondary tasks. Therefore, participants should have a greater proportion of correct target detections, a higher d', and shorter response times in the auditory secondary task condition because there would be less overlap for attentional resources between the primary and secondary tasks than in the visual secondary task condition. This reduction in resource overlap would mean that secondary task performance measured in terms of detection accuracy, d', and response time should be better with the auditory analog when compared to the ball-and-tunnel task. In addition, it was expected that measures of subjective workload will corroborate the differences observed in performance for the two secondary tasks.

H1: It was expected that there would be a main effect for workload for the proportion of correct detections, d', and response times. More specifically, performance in the high workload conditions was expected to be significantly different from the low workload and baseline conditions for both the ball-and-tunnel task and the auditory analog tasks. This would be reflected in a higher proportion of correct detections for baseline and low workload conditions and a lower number of false alarms and shorter response times for those same conditions when compared to the high workload condition.

H2: Because overlap in resource consumption was lower for an auditory analog secondary task than a ball-and- tunnel task, it was expected that the proportion of correct detections and d’ scores would be be significantly higher and response times would be be lower for the auditory secondary task when compared to the visual ball-and-tunnel task in the high workload condition.

H3: It was expected that there would be a significant simple effect for Subjective workload scores between the ball-and-tunnel task and auditory analog in the high workload conditions. Additionally, it was expected that there would be a significant main effect for subjective workload across the two modalities.

RQ: 1 Given the novelty of research into the specific workload demands related to spatial-auditory displays, it was difficult to make specific predictions for the NASA-TLX subscales. However, it was thought that there would be differences in mental demand, performance, effort, and frustration between the two dual-task conditions. Therefore, these four subscales were examined to determine whether there were any consistent differences among the subscales between the two secondary task condition

CHAPTER II

METHOD

Participants

A power analysis using G*Power 3.1.9.2 indicated that this study required 31 participants. To provide an even number for counterbalancing, 32 participants were recruited consisting of 25 women and 7 men with a mean age of 18.76 (SD = 3.36). Due to the novelty of this experiment, there were no studies from which appropriate effect sizes could be used for reference. Therefore, using Cohen's d_z , a moderate effect size of .5 was selected for the power analysis calculations with power set at .85 (Cohen, 1992)) and alpha at .05. Cohen's d_z was selected as the effect size measure because the goal was to measure the size of the difference as opposed to the proportion of total variance. The sample consisted of undergraduate students attending Old Dominion University. Participants will be compensated for their time with SONA credits which may satisfy class requirements or count toward extra credit. The study was be performed in compliance with the Old Dominion IRB and participation will be completely voluntary.

Equipment

***Laparoscopic box trainer*.** The box trainer is a 42 cm x 36 cm x 25 cm plastic box with a drawer used to prevent participants from having a direct view of the primary task. Within the box there is a pegboard with 12 pegs. Small rubber rings are placed on six of those pegs. On the top of the interior of the box is a Microsoft LifeCam VX-5000 USB which was used to record actions inside the box. The video feed from the camera was transferred to an Alienware OPTX

AW2210 monitor placed on top of the trainer box. The video image and the ball-and-tunnel task (see below Figure 1 below) is presented on the Alienware laptop.

Figure 1. The experimental setup with display.

Audio Configuration. Auditory signals were presented over six American Audio ELS 8GO LTW speakers. The ELS 8GO LTW is an 8 inch, 2-way battery-powered speaker. The six speakers surrounded the listener and be placed at the 1 o’clock, 3 o’clock, 5 o’clock, 7 o’clock, 9 o’clock, and 11 o’clock positions. The distance from the voice cone to the participant will be 32 inches and the distance from the floor to the central voice cone will be 64 inches. Current output for the speakers was set at 70 decibels (dB) and the sounds played were two synthesizer-generated, complex waveforms set 45-semitones apart. The signals sent to the speakers were controlled by a Presonus Audiobox 1818 VSL USB digital audio interface with 8 analog outputs.

Materials

Software. This study used two types of software. Superlab 5.0 is a stimulus presentation program capable of recording responses across a wide range of experimental applications. For this study, 6-channel sound files were stored in a folder and randomly selected and presented to participants. Reaper 5.0 is a digital audio workstation (DAW) designed for recording and generating sound files. With its plugin, RealSurround, Reaper will be used to create 6-channel audio to be presented using Superlab.

Subjective measures. The NASA Task Load Inventory Index (Hart & Staveland, 1988) is a self-report measure designed to evaluate perceived workload. The NASA- TLX is divided into six dimensions of workload: mental demands, physical demands, temporal demands, own performance, effort, and frustration. Participants respond on verbally weighted (low to high) visual analog scales. Test-retest reliability for the NASA-TLX is r = .83 (Hart & Staveland, 1988).

Primary Tasks

Peg transfer task. The high workload, primary task is the peg transfer task from the Fundamentals of Laparoscopic Surgery (FLS). The peg transfer task requires participants to transfer six rubber ring objects from one side of a peg board to another using two Johnson & Johnson Ethicon dissector/graspers (see Figure 2). Participants must first grasp one of the rubber rings with their nondominant hand and transfer it to their dominant hand before placing it on a peg on the opposite side of the peg board. They must continue these actions until all of the rubber rings have been placed on the opposite side of the board. When this is complete, participants must perform the same actions starting with their dominant hand and transfer each peg to their non-dominant hand before placing the ring on a peg. Participants are to perform as many peg transfers as they can in ten minutes. For this task, there are no requirements regarding the color of the rings or the specific peg to which they are transferred. Timing for the peg transfer task begins when the first ring is grasped and ends when ten minutes have elapsed. Additionally, transfers must occur in mid-air and dropping of rings will be recorded and counted as an error. If a ring is dropped participants will be asked to resume the task from the point just before dropping the ring.

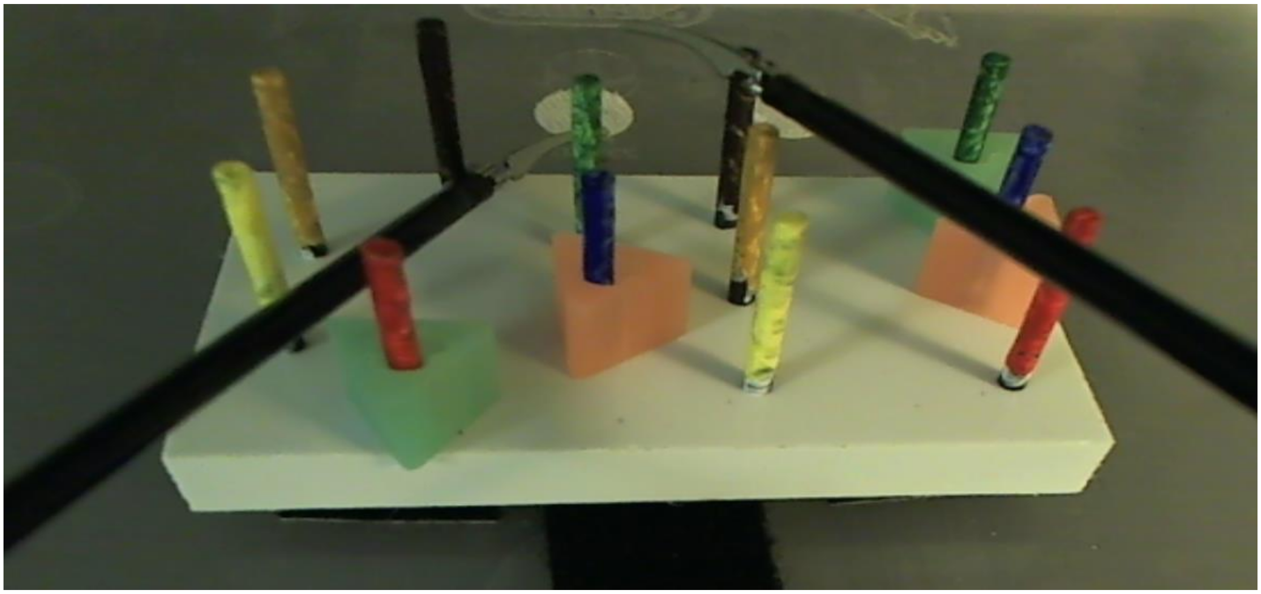

Figure 2. An image of the peg transfer task.

Eraser and bowl task. For the low workload, primary task participants must grasp 12 large, white pencil erasers one at a time and place them in a bowl using one Johnson & Johnson Ethicon dissector/grasper, then remove them (see Figure 3). This process will be repeated as many times as possible within five minutes. Timing for this task begins when the first eraser is successfully grasped. Participants will use only their dominant hand to complete this task. Should a participant unintentionally drop an eraser or miss the bowl it will be counted as an error.

Figure 3. An image of the eraser and bowl task

Secondary Tasks

Ball-and-tunnel task. The original ball-and-tunnel task (Prytz et al., 2012) consists of four balls, each of a different color. These balls are presented over a background designed to simulate a 3-dimensional tunnel. The illusion of depth is accomplished using dots that became smaller and closer together toward the center of the screen. The neutral or standard position of the balls was at the twelve, three, six, and nine o’clock positions. The image of balls is presented at random intervals every two to four seconds. Participants are required to respond using a foot pedal when one of the balls appears to have changed position relative to the others. The perceived change in position is conveyed by changing the diameter of a ball so that it appears closer (larger and nearer the edge of the tunnel) or smaller (farther away and closer to the center of the tunnel).

To be consistent with the auditory task, the ball-and-tunnel task was modified to display only two balls: one on the right side at 90 degrees (3 o’clock) and the other on the left side at 270 degrees (9 o’clock). Targets will change position by 15 degrees. For the right ball, the change will be to either the 75-degree or 105-degree positions and for the left ball, the change will be to either the 285-degree or 255-degree positions (see Figure 4 below).

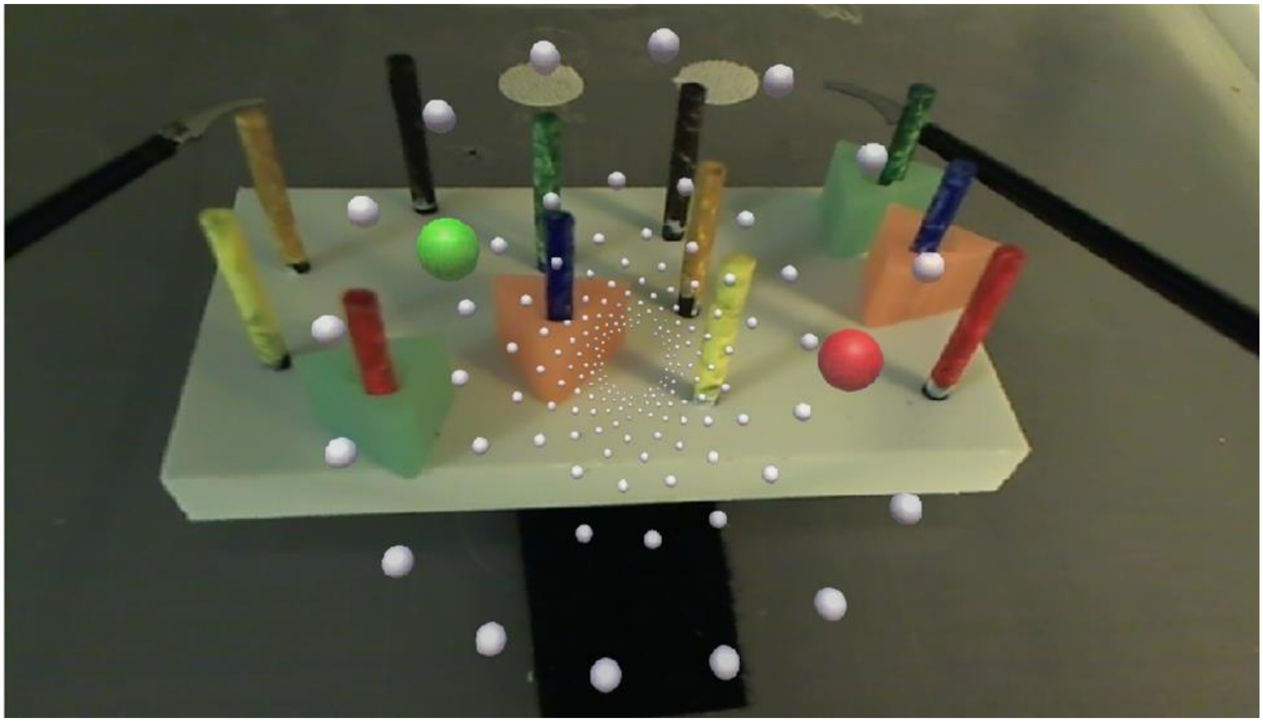

Figure 4. Dual task condition with ball-and-tunnel task projected over peg transfer task

Auditory task. Like the original ball-and-tunnel task, participants must respond when the auditory signals changes location. In the auditory version, the signal is played from a different speaker. The present configuration consists of two speakers positioned at 3 o'clock and 9 o'clock analogous to the balls in their neutral positions at the beginning of the ball-and-tunnel task. The other four speakers are placed at the 1 o'clock, 5 o'clock, 7 o'clock, and 11 o'clock positions. To improve discriminability there is a 45-semitone difference between the sounds presented on the right side versus the left.

Pilot Test

A pilot test was conducted with seven participants and revealed that the modified ball-and-tunnel task resulted in a high proportion of correct detections (*M* = .93, *SD* = 4.77) with a range of .86 to .99. Similarly, the auditory-spatial task produced a comparable proportion of correct detections (*M* = .94, *SD* = 3.84) with a range of .88 to 99. An equivalence test indicated that performance on the two tasks was not significantly different suggesting that the psychophysical performance was approximately equivalent for both tasks, paired-samples, *t*(6) = -.43, *p* =.68, 95% CI [-6.28, 4.28]. One additional finding was that there was a .41 second mean difference in response time *(SD* = .13) between the auditory and visual conditions. Initially this difference was thought attributable to modality. However, it was likely due to a hardware problem that will be addressed later in the text.

Procedure

Participants were given an informed consent form to read and sign. During this time, they were also informed that no personally identifiable information would be collected and that they could cease participation at any time. After signing the informed consent document, participants were given a background questionnaire (see Appendix A). They were then randomly assigned to begin with either the visual or auditory task condition and to either the high or low workload primary task condition. The purpose of group assignment by task was to control for order effects by counterbalancing. Instructions for each task were read aloud to the participants (see Appendix B). Next, participants performed the secondary task by itself to establish baseline measurements. The participants were then introduced to the peg transfer task or the eraser and bowl task and given 5 minutes of practice to familiarize themselves with the task. At this time, they had the opportunity to ask questions and receive feedback regarding aspects of the task. After the practice session, participants performed one of the two primary tasks simultaneously with the secondary task. Participants were given 5 minutes to perform the eraser and ball task and 10 minutes to perform the peg transfer task during the dual task portion of the experiment.

Participants completed each of the four dual task and secondary task pairings in this manner. The NASA-TLX was administered after each task is completed.

Dependent Measures

In both the visual and auditory ball-and-tunnel tasks, performance was measured using response time, proportion of correct responses, and d'. Subjective workload was recorded using the NASA-TLX. The composite workload score was of primary interest; however, individual scale scores were analyzed as well. Primary task performance was also measured as another index of the workload manipulation by the number of successful peg transfers in ten minutes for the peg transfer task and the number of eraser transfers in five minutes for the eraser and bowl task. The number of errors in the form of dropped rings was also counted.

Design

This study used two 2 x 3, within-subjects designs. The first, a 2(demand) x 3 (condition) design, was used to assess the effects of the workload manipulation. Determining that the primary tasks placed different levels of demand on participants was key in interpreting the results of the secondary task. The second design was a 2 (modality) x 3 (workload) within-subjects design. Modality was split into two secondary task conditions: auditory and visual. Workload was split into high and low conditions with performance on the secondary tasks by themselves serving as control conditions. The high workload primary task was a standard laparoscopic peg transfer task and the low workload primary task is the eraser task. The visual secondary task consisted of a modified version of ball-and-tunnel task and the auditory secondary task used sounds presented in a spatial array.

CHAPTER III

RESULTS

Visual inspection of the distributions of difference scores using histograms and q-q plots indicated that the data were approximately normally distributed. However, the sphericity assumption was violated for multiple tests due to a significant increase in variability between the lowest and highest workload conditions. In these cases, the results of Mauchley's test was reported and the Greenhouse-Geisser correction was used. A priori, pairwise comparisons were conducted to analyze secondary task performance using paired samples t-tests. In instances in which comparisons were not planned, post hoc pairwise comparisons were conducted using a Bonferroni correction. Examination of the proportion of false alarms revealed that several participants had no false alarms for one or more conditions. The absence of false alarms prohibits the use of signal detection measures. Although d' can be calculated in some instances, the results cannot be readily interpreted. Therefore, sensitivity analyses were excluded from further analysis. Because correct detections were measured as a proportion of the total number of targets presented and false alarms as a proportion of incorrect responses to non-targets, the mean for the total number of targets and non-targets presented for each condition is reported below. For the secondary tasks, the mean number of target presentations for the ball-and-tunnel task during the low workload condition was 41.31 (SD = 1.53) and 80.43 (SD = 2.72) for the high workload condition. The mean number of non-targets presented was 58.99 (SD = 1.77) in the low workload condition and 117.92 (SD = 2.37) for the high workload condition. For the auditory tasks the number of target presentations was 33.24 (SD = 2.6) in the low workload condition and 98.33 (SD = 2.71) in the high workload condition. The mean number of non-targets presented in the auditory task was 46.34 (SD = 1.91) in the low workload condition and 102.33 (SD = 2.04) in the high workload condition.`21 In addition, the mean amount of time between presentations of stimuli or the interstimulus interval (ISI) was also recorded. The mean interstimulus interval duration was 2.97 (SD = .56) for the ball-and-tunnel task and 3.1(SD =.47) for the auditory analog.

Primary Task Performance

Primary task performance was recorded as an index of the workload manipulation. A 2 (demand) x 3 (condition) ANOVA was used to examine the effects of the manipulation for the number successful transfers. Mauchley's test revealed that the sphericity assumption was violated for condition and the interaction between condition and demand. Therefore, a Greenhouse-Geisser correction was used. Results indicated that there was a significant difference in the number of successful transfers between the low and high levels of demand, $F(1, 31) = 529.11$, $p < .001$, $\eta^2 = .945$. However, no significant difference was found among the conditions (single task, dual task with ball-and-tunnel, and dual task with auditory analog), $F(1.601, 31) = 2.41$, $p = .110$, $\eta^2 = .072$, nor was there a significant demand by condition interaction, $F(1.193, 36.989) = 2.77$, $p = .098$, $\eta^2 = .082$ (see Table 1).

Table. 1

Results of Analysis of Variance for Primary Task Performance: Successful Transfers

	SS	df	MS	*F*	*p*	η^2
Demand	7537.547	1	7537.547	529.107	.000*	.945
Error	441.620	31	14.246			

Condition	20.906	1.601	13.061	2.414	.110[a]	.072
Error	268.427	31	8.659			
Demand x Condition	39.031	1.193	32.712	2.773	.098[a]	.082
Error	436.302	36.989	11.796			

Note. *$p < .05$, [a] Greenhouse-Geisser corrected

Another 2 (demand) x 3 (condition) ANOVA was used to examine the effects of the workload manipulation on the number of drops. The analysis of the number of drops also indicated that there was a significant difference between the low and high demand conditions, $F(1,31) = 143.69$, $p < .001$, $\eta^2 = .823$. Again, there was no difference observed among the three conditions, $F(2, 62) = .125$, $p = .883$, $\eta^2 = .004$, nor was there a significant demand by condition interaction, $F(2, 62) = 1.42$, $p = .249$, $\eta^2 = .044$ (see Table 2).

Table 2

Results of Analysis of Variance for Primary Task Performance: Drops

	SS	df	MS	*F*	*p*	η^2
Demand	985.547	1	985.547	143.693	.000*	.823
Error	212.620	31	6.859			

Condition	.448	2	.224	.125	.883	.004
Error	111.219	62	1.794			
Demand x Condition	5.281	2	2.641	1.423	.249	.044
Error	115.052	62	1.856			

Note. $^*p < .05$

Secondary Task Performance

Proportion of correct detections. To test hypotheses 1 and 2, a 2 (modality) x 3 (workload) repeated-measures ANOVA was used to compare the proportion of correct detections among modality and workload conditions. The sphericity assumption was violated for workload and the interaction between modality and workload. Here again, the Greenhouse-Geisser correction was applied. The results indicated that there was not a significant effect for modality, $F(1, 31) = 1.55$, $p = .223$, $\eta^2 = .048$; however, there was a significant effect for workload; $F(1.467, 45.484) = 109.80$, $p < .001$, $\eta^2 = .780$; and a significant interaction between modality and workload, $F(1.652, 51.197) = 13.13$, $p < .001$, $\eta^2 = .298$ (see Table 3).

Table 3

Results of Analysis of Variance for Proportion of Correct Responses

	SS	df	MS	*F*	*p*	η^2
Modality	4.6	1	4.6	1.547	.223	.048
Error	92.184	31	2.974			
Workload	64410.920	1.467	43900.181	109.803	.000 $*^a$	.780
Error	18184.754	45.484	399.809			
Modality x Workload	56.358	1.652	34.125	13.129	.000 $*^a$	.298
Error	133.076	51.197	2.599			

Note. $*p < .05$, [a] Greenhouse-Geisser corrected

A priori, pairwise comparisons with a Bonferroni correction were used to assess differences in correct responses between the high and low workload conditions and the high workload condition and baseline while controlling for alpha inflation due to the number of comparisons. An additional post hoc comparison was used to address differences between the low workload condition and baseline. Results showed that there were significant differences between the baseline and low workload condition, the baseline and the high workload condition, and the low workload condition and the high workload condition (see Table 4). Next, a paired sample t test was used to analyze differences in the proportion of correct responses between the auditory and visual modalities in the high workload condition. The results indicated that there was a significant difference between the visual (*M* =47.37, *SD* = 21.22) and auditory (*M* = 49.21, *SD* = 20.28) modalities in the high workload conditions, *t*(31) = -3.73, *p* = .001,d_z = .67. There were no other significant effects (see Table 5).

Table 4

Pairwise Comparisons for Workload with Bonferroni Correction: Correct Detections

				95% CI	
	Mean Difference	Std. Error	*p*	Lower Bound	Upper Bound
Baseline – Low	19.353	1.929	.000*	14.471	24.235
Baseline – High	44.730	3.563	.000*	35.713	53.747
Low – High	25.377	3.329	.000*	16.951	33.803

Note. **p* <.05

Table 5

A Priori Pairwise Comparisons for Modality

							95% CI	
Visual - Auditory	Mean Difference	*SD*	Std. Error Mean	*t*	df	*p*	Lower	Upper
Baseline	.40	1.59	.28	1.42	31	.167	-.17	.975
Low	.51	2.10	.37	1.37	31	.179	-.24	1.27
High	-1.83	2.73	.43	-3.73	31	.001*	-2.81	-.84

*Note. *Significant with Bonferroni correction*

Proportion of false alarms. Upon inspection, it was found that participants made very few false alarms. The proportions of false alarms did not vary enough to justify further analysis; therefore, descriptive statistics are presented without statistical tests (see Table 6).

Table 6

Descriptive Statistics for the Proportion of False Alarms

	N	Min.	Max.	*M*	*SD*
Visual Baseline	32	0.00	4.8	2.59	1.37
Auditory Baseline	32	0.00	5.1	2.72	1.40
Visual Low Workload	32	0.00	4.4	1.81	1.61
Auditory Low Workload	32	0.00	4.6	2.30	1.57
Visual High Workload	32	0.00	5.0	1.48	1.94
Auditory High workload	32	0.00	6.1	2.57	1.63

Response Time. To test hypotheses 1 and 2, a 2 (modality) x 3 (workload) repeated-measures ANOVA was used to assess differences in response time. During initial analysis it was found that a time delay had been introduced by the hardware used for the auditory task. The mean delay duration was .046 seconds (*SD* = .036). This time was subtracted from the response times for auditory tasks for each participant. After applying this correction, a significant effect for workload was detected, *F*(1.776, 55.050) = 17.14, $p < .001$, $\eta^2 = .356$, with significant mean differences in response time between baseline and the low workload condition and between the baseline and high workload condition (see Tables 7 and 8).

Table 7

Results of Analysis of Variance for Response Time

	SS	df	MS	*F*	*p*	η^2
Modality	.008	1	.008	.327	.571	.010
Error	.744	31	.024			
Workload	2.166	1.776	1.083	17.137	.000*	.356
Error	3.917	55.050	.063			
Modality x Workload	.032	1.162	.028	.680	.436[a]	.021
Error	1.458	36.009	.040			

Note. *$p < .05$, [a] Greenhouse-Geisser corrected

Table 8

Pairwise Comparisons for Workload with Bonferroni Correction: Response Time

	95% CI

	Mean Difference	Std. Error	p	Lower Bound	Upper Bound
Baseline – Low	-.179	.036	.000*	-.269	-.088
Baseline – High	-.253	.047	.000*	-.372	-.134
Low Workload – High	-.074	.049	.424	-.199	.050

Note. *p < .05

Subjective Workload

Global Workload. To test hypothesis 3, a 2 (modality) x 3 (workload) repeated-measures ANOVA was used to assess subjective workload ratings on the NASA-TLX. Mauchley's test revealed that the sphericity assumption had been violated for workload and a Greenhouse-Geisser correction was applied. Results of the ANOVA for the mean total scores on the TLX as a measure of global workload indicated that there was a significant main effect for modality, *F*(1, 31) = 4.81, p = .036 η^2 = .010. The results for workload also indicated that there was a significant difference among the workload conditions, *F*(1.48, 45.75) = 61.35, p < .001, η^2 = .664. The interaction between modality was not significant. (see table 9).

Table 9

Results of the Analysis of Variance for NASA-TLX Total Scores

	SS	df	MS	F	p	η^2
Modality	517.253	1	517.253	4.812	.036	.010

Error	3332.133	31	107.488			
Workload	22675.984	1.476	15364.742	61.347	.000*[a]	.664
Error	11458.667	45.751	250.456			
Modality x Workload	94.097	2	47.049	1.857	.165	.057
Error	1570.878	62	25.337			

Note. *p < .05, [a] Greenhouse-Geisser corrected

Here again, the workload data were analyzed with a priori, pairwise comparisons for the high workload condition and the low workload condition and the high workload condition and baseline using a Bonferroni correction to address potential alpha inflation. A post hoc comparison was used to assess differences between the low workload condition and baseline. Results showed that there were significant differences in workload scores between each of the three workload conditions (see table 10). Although the interaction between modality and workload was not significant for the omnibus test, a difference between the visual and auditory modalities in the high workload condition was hypothesized and therefore analyzed using preplanned paired sample t tests. The results revealed that the mean TLX score for the visual high workload condition (M = 75.96, 17.29, SD = 3.06) was significantly higher than the auditory high workload condition (M = 70.78, SD = 3.31), t(31) = 3.15, p = .004, d_z = .56. There were no other significant differences found (see table 11).

Table 10

Pairwise Comparisons for Workload with Bonferroni Correction: NASA-TLX Total Scores

				95% CI	
	Mean Difference	Std. Error	*p*	Lower Bound	Upper Bound
Baseline – Low	-15.766	2.126	.000*	-21.147	-10.385
Baseline – High	-26.458*	3.030	.000*	-34.126	-18.791
Low Workload – High	-10.692*	1.905	.000*	-15.513	-5.872

Note. $*p < .05$

Table 11

A Priori Pairwise Comparisons for Modality

							95% CI	
Visual - Auditory	Mean Difference	*SD*	Std. Error Mean	*t*	df	*p*	Lower	Upper
Baseline	1.84	12.61	2.23	.829	31	.414	-2.70	6.39
Low	2.81	8.38	1.48	1.899	31	.067	-5.87	5.84
High	5.18	9.31	1.64	3.149	31	.004*	1.82	8.53

Note. $*p < .05$

NASA-TLX Scale Scores. To test research question 1, four 2 (modality) x 3 (workload) repeated measures ANOVAs were used to assess differences among conditions for four of the NASA-TLX subscales: mental demand, performance, effort, and frustration (see table 12). Mauchley’s test indicated that the sphericity assumption had been violated on each of the four subscales, and for the interaction between modality and workload on the mental demand, effort, and frustration subscales. In each of these cases a Greenhouse-Geisser correction was applied.

For the mental demand subscale, analyses revealed that there was a significant main effect for modality, $F(1,31) = 6.03$, $p = .020$, $\eta^2 = .163$, and workload, $F(1.621, 62) = 46.17$, $p < .001$, $\eta^2 = .598$. However, the interaction effect was not significant. Post hoc pairwise comparisons for workload showed significant differences in subjective workload ratings between each of the three conditions (see Table 13). Pairwise comparisons using Bonferroni corrected paired-sample t tests showed that mental workload scores were significantly higher for the visual modality ($M = 86.09$, $SD = 17.03$) than the auditory modality under the high workload condition ($M = 78.28$, $SD = 23.61$), $t(31) = 2.78$, $p = .009$, dz = .287. Pairwise comparisons for modality for each of the scales can be found in Table 14.

Turning to performance, scores on this subscale were inverted meaning that lower performance ratings reflect higher workload. The analyses for the performance subscale showed that there were no significant effects (see Table 13).

Results for the analysis of the effort subscale indicated that there was a significant main effect for workload, $F(1.340, 41.535) = 17.11$, $p <.001$, $\eta^2 = .356$. Pairwise comparisons using a Bonferroni correction indicated that there were significant mean differences in effort ratings between each level of workload (see table 12).

The last subscale to be analyzed was frustration. The analysis revealed that there was a significant main effect for modality, $F(1, 31) = 4.75$, $p = .037$, $\eta^2 = .133$, and a significant main effect for workload, $F(1.866, 57.839) = 44.16$, $p < .001$, $\eta^2 = .588$. The interaction effect was not significant (see Table 12). Post hoc pairwise comparisons with Bonferroni correction indicated that there were significant mean differences between each of the workload conditions (see Table 13). In addition, pairwise comparisons were used to examine differences between the two modalities in each workload condition. The results indicated that there was a significant difference in mental demand, t(31) = 2.78, p = .009, d_z = .40 (see Table 14). There were also differences between the visual and auditory modalities for performance in the high workload condition and frustration in the low and high workload conditions. However, there was insufficient power to find significance with the corrected alpha (see Table 14).

Table 12

Results of Analysis of Variance for NASA-TLX Scale Scores

		SS	df	MS	*F*	*p*	η^2
Mental	Modality	1354.688	1	1354.688	6.025	.020	.163
	Error	6970.313	31	224.849			
	Workload	32388.281	1.621	19982.835	46.173	.000*[a]	.598
	Error	21745.052	62	350.727			
	Modality x Workload	150.781	1.525	98.879	.619	.500[a]	.020
	Error	7549.219	47.272	159.697			
Performance	Modality	963.021	1	963.021	4.102	.052	.117
	Error	7278.646	31	234.795			
	Workload	4782.292	1.584	3019.155	3.715	.041*	.107
	Error	39901.042	49.103	812.591			

	Modality x Workload	226.042	2	113.021	.631	.533	.020
	Error	11107.292	62	179.150			
Effort	Modality	287.630	1	287.630	1.718	.200	.052
	Error	5191.536	31	167.469			
	Workload	14232.292	1.340	10622.512	17.111	.000*[a]	.356
	Error	25784.375	41.535	620.794			
	Modality x Workload	54.167	2	27.083	.226	.744[a]	.007
	Error	7429.167	48.613	152.822			
Frustration	Modality	1518.750	1	1518.750	4.745	.037	.133
	Error	9922.917	31	320.094			
	Workload	40434.635	1.866	21671.590	44.164	.000*[a]	.588
	Error	28382.031	57.839	490.703			
	Modality x Workload	436.719	2	218.359	1.133	.328[a]	.035
	Error	11946.615	61.139	195.400			

Note. *$p < .001$, [a] Greenhouse-Geisser corrected

Table 13

Pairwise Comparisons for Workload with Bonferroni Correction: Scale Scores

95% CI

		Mean Difference	Std. Error	*p*	Lower Bound	Upper Bound
Mental	Baseline – Low	-21.094*	3.442	.000*	-29.804	-12.384
	Baseline – High	-31.172*	3.879	.000*	-40.989	-21.355
	Low – High	-10.078*	2.448	.001	-16.273	-3.883
Performance	Baseline – Low	-2.344	3.298	1.000	-10.690	6.002
	Baseline – High	-11.563	5.356	.116	-25.119	1.994
	Low – High	-9.219	4.558	.155	-20.754	2.316
Effort	Baseline – Low	-12.656*	3.780	.006	-22.223	-3.089
	Baseline – High	-20.938*	4.491	.000*	-32.304	-9.571
	Low – High	-8.281*	2.128	.001	-13.668	-2.895
Frustration	Baseline – Low	-17.813*	3.864	.000*	-27.591	-8.034
	Baseline – High	-35.547*	4.156	.000*	-46.065	-25.028
	Low \– High	-17.734*	3.273	.000*	-26.019	-9.450

Note. *$p < .05$

Table 14

Pairwise Comparisons for Modality: Scale Scores

								95% CI	
	Visual - Auditory	Mean Difference	*SD*	Std. Error Mean	*t*	df	*p*	Lower	Upper
Mental	Baseline	4.219	21.631	3.824	1.103	31	.278	-3.58	12.018
	Low	3.906	14.687	2.596	1.505	31	.143	-1.39	9.201
	High	7.813	15.910	2.813	2.778	31	.009	2.08	13.549
Performance	Baseline	2.500	22.540	3.985	.627	31	.535	-5.63	10.627
	Low	3.438	19.404	3.430	1.002	31	.324	-3.56	10.433
	High	7.500	17.367	3.070	2.443	31	.020	1.24	13.761
Effort	Baseline	1.406	20.683	3.656	.385	31	.703	-6.05	8.863
	Low	3.906	14.687	2.596	1.505	31	.143	-1.39	9.201
	High	2.031	13.067	2.310	.879	31	.386	-2.68	6.742
Frustration	Baseline	1.406	23.869	4.219	.333	31	.741	-7.19	10.012

Low	8.281	22.490	3.976	2.083	31	.046	.173	16.390
High	7.188	18.313	3.237	2.220	31	.034	.585	13.790

Note. Alpha adjusted to .016

CHAPTER IV

DISCUSSION

The purpose of the present study was to test Multiple Resource Theory (MRT; Wickens, 1980, 1998, 2008) by comparing performance on a visual-spatial secondary task with that of an auditory-spatial secondary task during high and low workload laparoscopic tasks. Each participant performed combinations of two primary and two secondary tasks along with each primary and secondary task alone for a total of eight task blocks. The primary tasks were divided into high and low workload levels based on task complexity. The secondary tasks consisted of a visual-spatial task and its auditory analog.

Primary Task Performance

Performance on the primary laparoscopic tasks was examined as an index of the workload manipulation. This is of particular importance because Wickens (1998) posited that performance differences due to overlapping demand on resource pools should become more pronounced as workload increases. As a result, performance on one or both tasks should decline when they both rely on the same resource pool. The results of the present study indicated that the low workload task was significantly less demanding than the high workload task with a partial η^2 of .945 for successful transfers and .823 for drops. This finding suggests that the workload manipulation was successful. Furthermore, there were no significant differences among conditions in which primary tasks were paired with secondary tasks. The lack of significant effects for primary tasks outside of workload provided evidence that participants followed directions and maintained their performance across all conditions. These performance results also suggest that the secondary tasks were not intrusive and therefore did not have a negative effect on primary task performance. Additionally, these results support the findings of previous researchers indicating that differences in primary task performance are due to the difficulty associated with those tasks and not secondary task intrusion (Britt, et al., 2015; Warvel, 2015). Both of these findings are essential for drawing conclusions from secondary task performance.

Secondary Task Performance

The primary goal of the study was to compare performance differences between auditory and visual modalities under differing workload conditions using secondary tasks. Consequently, it was necessary at the outset to establish that performance was similar for both modalities at baseline. Maintaining similar performance in single task conditions would ensure that performance differences observed in dual task conditions were due to an increase in workload from time sharing as predicted by MRT. To this end, the first hypothesis predicted that performance would not differ between modalities in the baseline and low workload conditions and that performance would decline in high workload conditions when compared to baseline and the low workload conditions. Results of the present study showed that performance in the auditory and visual single task conditions did not differ significantly supporting the first hypothesis and allowing for the conclusion that observed differences in performance were due to an increase in workload associated with the addition of another task as would be predicted by MRT. The results also showed that performance was poorer in the high workload conditions when compared to baseline and low workload conditions. Support for the first hypothesis provides corroborating evidence for Wickens’ (1980) finding that the negative effects of time sharing on performance are not always observable in low workload conditions. The degree of consistent performance among low workload conditions demonstrated that dual task settings alone did not place enough demand on resource pools to have a detrimental effect on an individual’s ability to perform. Instead, as Wickens’ (1980) predicted, demand would need to be increased in one or both tasks to observe a decline in performance.

MRT predicts that performance should decline when multiple tasks place demands on the same resource pool under high workload conditions. Therefore, the second hypothesis stated that detections would be significantly lower for the visual modality than the auditory modality during high workload conditions and response times would be significantly longer. The results revealed partial support for the second hypothesis. Participants had a significantly higher proportion of correct detections in the auditory condition than in the visual condition with Cohen’s $d_z = .67$, indicating a large effect. These findings suggest that participants were better able to detect auditory targets than visual targets while simultaneously performing a visual-spatial primary task. This finding directly supports the predictions of MRT regarding modality. Under high workload conditions, a visual-spatial secondary task was more sensitive to a high workload, visual-spatial primary task than an auditory-spatial task secondary task. However, response times did not differ significantly between the two modalities. One issue that affected the response times was the introduction of a delay caused by the hardware used for the auditory condition. This problem will be discussed further in the limitations section.

To corroborate findings from performance measures, subjective workload measures were used to indicate which tasks were perceived to be more demanding. For the present study, the NASA-TLX (Hart & Staveland, 1988) was used to determine whether there were differences in perceived demand among the workload conditions and between the two modalities. The third hypothesis predicted that subjective workload scores would be significantly lower at baseline than in the low and high workload conditions across modality. In addition, it was expected that subjective workload scores would be significantly higher in the high workload, dual task conditions when compared to the low workload conditions and baseline. Consistent with this prediction, the results revealed significant differences in subjective workload with ratings increasing from baseline to low workload and from low workload to high workload. This finding provides further evidence that the workload manipulation was effective. Further, it was predicted that subjective workload scores would be lower for the auditory modality than the visual modality in the low and high workload conditions. This aspect of the hypothesis was partially supported. Subjective workload ratings were significantly lower for the auditory modality in the high workload condition with Cohen's $d_z = .56$. Subjective workload scores for the auditory modality were also lower than for the visual modality in the low workload condition. However, this difference was not significant suggesting that demand was not sufficient to affect the ratings. Overall, analysis of global workload supported the predictions of MRT that secondary tasks presented using a different modality than the primary task is perceived as being less demanding.

In addition to a global workload score, the NASA-TLX is comprised of six subscales. A secondary goal of this study was to explore perceived workload differences between modalities on four of the six subscales: mental demand, performance, effort, and frustration. Little research has been done on scale score differences for this specific type of secondary task comparison. The results showed that there were significant differences for each of the levels of workload on all four subscales providing further evidence for the effects of the workload manipulation. In terms of individual subscale scores, participants found the visual secondary task to be significantly more mentally demanding than the auditory secondary task in the high workload condition with Cohen's $d_z = .40$. There were also noteworthy differences for performance and frustration ($p<.05$ uncorrected), but there was insufficient power to find significance with the correction applied.

Theoretical Implications

One limitation of MRT as presented by Wickens (2008) is that it does not fully explore predictions as they relate to specific dimensions. To address this issue, the primary goal of this study was to directly test performance differences attributed to modality as predicted by MRT. This was achieved by using two matched secondary tasks as opposed to presenting a portion of the primary task in another modality or using cues in another modality to direct attention to a specific part of the task. By creating an auditory analog of an established visual-spatial secondary task, this study effectively isolated the modality dimension. The findings of the present study indicated that the auditory-spatial secondary task was less sensitive than the visual-spatial secondary to the demands of a visual-spatial primary task supporting the idea that processing is different between the two modalities.

However, the results also run counter to the findings of Wickens and Liu (1988) and Latorella (1998) who found that discrete auditory tasks were more disruptive than visual tasks when paired with continuous visual tasks due to the auditory input "pre-empting" visual processing. Given that primary task performance for the current study was unaffected by either secondary task, these earlier findings are likely the result of an important difference between the present study and the two previously mentioned. The Wickens and Liu (1988) and the Latorella (1998) studies involved pairing a continuous, visual, primary task with a discrete, auditory, secondary task. In this case, the discrete secondary task interrupted the continuous, primary task. In contrast, the present study paired a continuous, laparoscopic task with a continuous monitoring task. Therefore, the previous studies speak more to the effects of interrupting a visual primary task with an auditory task as opposed to the effects of both tasks being performed simultaneously. Another important difference is that the two previous studies used complex, verbal auditory tasks which were likely more demanding than the ball-and-tunnel task or its auditory analog. The present study supports the notion that different modalities may draw from separate resource pools and that this may be observable at a basic research level. Yet, the predictions of MRT remain context dependent.

Applied Implications

Laparoscopic surgery is more difficult than open surgery due to visual, tactile, and spatial orientation differences. Additionally, laparoscopy is a complex, visual-spatial task (Berguer, Smith, & Chung, 2001; Braga, et al., 2005; King, et al., 2005). One of the major contributions of MRT is that it provides guidance for improving performance by identifying aspects of tasks drawing on a common resource pool and instead distributing them across multiple resource pools (Wickens, 2008). Other studies have also demonstrated the efficacy of using auditory displays to reduce demand placed on visual-spatial resources (Begault, 1993; Bronkhorst, Vetman, & Van Breda 1996; Mckinley & Ericson, 1997).

In medical contexts, auditory displays have been evaluated and found to be effective when paired with visual monitoring devices that display patient information in numbers, as histograms, or polygons (Loeb & Fitch, 2002; Sanderson, Liu, & Jenkins, 2009). However, similar approaches have not been taken to reduce the demand placed on surgeons by laparoscopic surgery. The findings of the present study suggest that an auditory task is significantly easier to perform than a visual task when coupled with a a primary, laparoscopic task. Given that auditory tasks are less sensitive to the visual demands of laparoscopy than visual tasks, some information may be better conveyed in the auditory domain. Presenting some information needed during laparoscopic surgical tasks using an auditory display could result in reduced visual-spatial demand which could improve surgical performance and reduce errors. By reducing the amount of information that a surgeon needs to process visually, overall visual-spatial demand could also be reduced. In turn, this could reduce the perceived workload associated with performing laparoscopic tasks.

Another purpose of the present study was to further evaluate different measures of workload during laparoscopic surgery. In this context, the present study may be considered a continuation of research by Stefanidis, Scerbo, Korndorffer, & Scott (2007), Prytz and colleagues (2012), and Warvel (2015). The findings from these earlier studies support the predictions of MRT by showing that the visual-spatial, ball-and-tunnel task was sensitive when paired with visual-spatial, laparoscopic primary tasks making it a good choice for measuring workload during laparoscopy. In contrast, the goal of the present study was to determine whether presenting an analog of the ball-and-tunnel tasks in the auditory modality would yield different results. The results showed that the auditory analog task was less sensitive than the ball-and-tunnel task when paired with laparoscopic, primary tasks. This result is consistent with predictions based on MRT and suggests that an auditory secondary task is a poorer measure of the workload demanded by visual-spatial primary tasks.

Limitations

There were several limitations with the study that bear consideration. One notable limitation was that the hardware used to present the auditory stimuli introduced a time delay which affected the recording of response times. This was addressed by subtracting the average delay duration from each score prior to analyzing the data. Analysis of the delay times introduced by the hardware indicated that they did not vary significantly, nor did response times vary significantly within task blocks. This finding suggests that the delay was approximately the same across all participants for each task block. In addition, there were no trends that would indicate that the duration of the delay changed over the course of the experiment. However, response times recorded without a delay would be more accurate than those corrected after the fact. Further research should make use of more up-to-date audio equipment and accuracy of response times should be checked to ensure that the devices are operating correctly.

Another possible limitation is that the present study dealt only with novice participants. The advantage of limiting participation in the study to novices was that workload differences were very clear. However, many participants struggled with the novelty of the primary tasks which likely affected how subjective workload was reported. There may be some benefit to including more experienced participants to compare to novices or having participants complete the tasks across multiple sessions in future research (Patten,Kircher, Ostlund, Nilsson, & Svenson, 2006; Tole, Stephens, Harris, & Ephraph, 1982) .

One final limitation was that there were substantially more female than male participants. Given that there are differences in spatial processing between men and women, this issue should not be overlooked (Simon-Dack, Friesen, & Teder-Sälejärvi, 2011; Vecchi & Girelli, 1998). That said, there was no evidence that gender accounted for differences in performance in the present study.

Future Research

Some of the most important aspects of MRT to consider are the underlying processing differences that could potentially explain why the model works. The current study provided evidence supporting the notion that visual-spatial and auditory-spatial processing are different enough that pairing them in dual task conditions results in observable differences in performance and subjective workload. The explanation for these observations provided by MRT (Wickens, 1980,1984, 2002, 2008) is that each modality pulls from a separate resource pool. However, the physiological mechanisms underlying this phenomenon remain unexplained. One of the goals of future research on this topic should be to further assess the physiological differences in processing among modalities to further explain differences in resource expenditure.

An important step to expand on the findings of this study would be to perform a similar experiment using a tactile analog. The tactile modality is a relatively recent addition to the Multiple Resource Model (Wickens, 2008). Therefore, conducting a study using the same paradigm, changing only the modality of the analog would provide directly comparable results among all three modalities. These comparisons could provide further support for predictions relying on the modality dimension of the Multiple Resource Model. Alternately, they could help reveal aspects of the model that are not consistent across modalities.

One final recommendation for future work is to include expert-level participants. For the initial study, it was important for participants to be unfamiliar with the primary tasks to ensure that the high workload task was sufficiently demanding. While this approach was effective, it also limited the generalizability of the study, particularly considering that practicing laparoscopic surgeons are experts. The primary challenge of generalizing the results of the present study to practicing laparoscopic surgeons is that the high workload primary tasks used in the present study would not likely place much demand on expert surgeons. With adjustments to the workload manipulation, it should be possible to determine the extent to which the findings of the current study extend to those with experience performing laparoscopic procedures. In the present study, novice participants struggled to complete the peg transfer task in 10 minutes, whereas the Fundamentals of Laparoscopic Surgery (FLS) program requires surgeons to complete the task in under 48 seconds to be considered proficient (FLS Proficiency-Based Training Curriculum, 2014). A performance difference this large suggests that a task for surgeons would need to be much more complex to achieve the same level of demand as the novices experienced.

CHAPTER 5

CONCLUSION

The present study compared performance on a visual-spatial secondary task with that of an auditory analog to test predictions of MRT. Specifically, this study sought to compare performance on two tasks that differed only in modality. Much of the previous research exploring differences between visual-spatial and auditory-spatial processing applying MRT has consisted of adding auditory displays to existing visual tasks. One limitation of this approach is that the demand of each task or task component is not known. The goal of the present study was to use a well-established measure of workload for laparoscopic surgery and its auditory equivalent to isolate modality differences. This study demonstrated the efficacy of using analogues of established tasks to more directly investigate performance differences among the dimensions of the Multiple Resource Model.

Overall, the findings of present study support the prediction that the auditory analog would be less sensitive than the visual spatial secondary task. This finding provides additional evidence for the effects of modality on performance in multitask settings. More specifically it supports the notion that task performance under high workload conditions is better when multiple tasks do not share the same modality. Further, the findings support the notion that there may be different resource pools for each modality and by extension for each stage, code, and channel of the Multiple Resource Model.

REFERENCES

Begault, D. R. (1993). Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation. *Human factors*, *35*, 707-717.

Berguer, R., Smith, W. D., & Chung, Y. H. (2001). Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. *Surgical Endoscopy, 15,* 1204-1207. doi: 10.1007/s004640080030

Braga, M., Vignali, A., Zuliani, W., Frasson, M., Di Serio, C., & Di Carlo, V. (2005). Laparoscopic versus open colorectal surgery. *Annals of Surgery, 242,* 890-896. doi: 10.1097/01.sla.0000189573.23744.59

Bronkhorst, A. W., Veltman, J. A., & Van Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human factors*, *38*, 23-33.

Brown, M. L., Newsome, S. L., & Glinert, E. P. (1989, March). An experiment into the use of auditory cues to reduce visual workload. In *ACM SIGCHI Bulletin*, 20, 339-346.

Boles, D. B., & Law, M. B. (1998). A simultaneous task comparison of differentiated and undifferentiated hemispheric resource theories. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 204-215. doi: 10.1037/0096-1523.24.1.204

Britt, R. C., Scerbo, M. W., Montano, M., Kennedy, R. A., Prytz, E., & Stefanidis, D. (2015). Intracorporeal suturing: Transfer from Fundamentals of Laparoscopic Surgery to cadavers results in substantial increase in mental workload. *Surgery, 158*, 1428-1433. doi:10.1016/j.surg.2015.03.032

Carswell, C. M., Clarke, D., & Seales, W.B. (2005). Assessing mental workload during laparoscopic surgery. *Surgical Innovation*, *12*, 80-90. doi: 10.1177/155335060501200112

Crawford, J., Watson, M., Burmeister, O., & Sanderson, P. (2002). Multimodal displays for anaesthesia sonification: timesharing, workload, and expertise. In *Proceedings of the joint ESA/CHISIG Conference on Human Factors (HF 2002). Melbourne, Australia.*

Cuschieri, A. (1995). Visual displays and visual perception in minimal access surgery. *Surgical Innovation, 2*, 209-214. doi: 10.1177/155335069500200030

Gallagher, A., McClure, N., McGuigan, J., Ritchie, K., & Sheehy, N. (1998). An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy*, *30*, 617-620. doi: 10.1055/s-2007-1001366

Gallagher, A. G., Al-Akash, M., Seymour, N. E., & Satava, R. M. (2009). An ergonomic analysis of the effects of camera rotation on laparoscopic performance. *Surgical Endoscopy, 23,* 2684-2691. doi: 10.1007/s00464-008-0261-x

Hart, S. G. (2006). NASA-task load index (NASA-TLX): 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50, 904-8

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52,* 139-183. doi: 10.1016/S0166-4115(08)62386-9

Jeon, M., Gable, T. M., Davison, B. K., Nees, M. A., Wilson, J., & Walker, B. N. (2015). Menu navigation with in-vehicle technologies: Auditory menu cues improve dual task performance, preference, and workload. *International Journal of Human-Computer*

Interaction, 31, 1-16.

King, P.M., Blazeby, J.M., Ewings, P., Franks, P.J., Longman, R.J., Kendrick, A.H, Kipling, R.M., & Kennedy, R.H. (2006). Randomized clinical trial comparing laparoscopic and open surgery for colorectal cancer within an enhanced recovery programme. *British Journal of Surgery, 93*, 300-308. doi: 10.1002/bjs.5216

Klein, M. I., Warm, J. S., Riley, M. A., Matthews, G., & Parsons, K. S. (2004). The multi-dimensional nature of stress in an endoscopic surgery simulator. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 48*, 1700-1703. doi: 10.1177/154193120404801515

Latorella, K. A. (1998). Effects of Modality on Interrupted Flight Deck Performance: Implications for Data Link. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*, 87–91. https://doi.org/10.1177/154193129804200120

Liu, Y. C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, *44*, 425-442.

Loeb, R. G., & Fitch, W. T. (2002). A laboratory evaluation of an auditory display designed to enhance intraoperative monitoring. *Anesthesia & Analgesia*, *94*, 362-368.

Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Linton, P.M., Wierwille, W.W., Zaklad, A.L., Bittner, A.C., & Wherry, R.J. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies* (DTIC Publication No. ADA212879). Retrieved from DTIC Online website:

http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA212879

McKinley, R. L., & Ericson, M. A. (1997). Flight demonstration of a 3-D auditory display.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual review of psychology, 42*, 135-159.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*, 214-255. doi: 10.1037/0033-295X.86.3.214

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance, vol. II: Cognitive processes and performance* (pp. 1-49). Oxford, England: John Wiley & Sons.

Oshima, C., & Wickens, C. D. (1992). Intuitive display features for situation awareness: The effects of frame-of-reference, modality, and redundancy. *Aviation Research Laboratory Institute of Aviation*.

Patten, Kircher, Östlund, Nilsson, & Svenson. (2006). Driver experience and cognitive workload in different traffic environments. *Accident Analysis and Prevention, 38*, 887-894.

Pavlovic, N. J., Keillor, J., Hollands, J. G., & Chignell, M. H. (2009). Reference Frame Congruency in Search-and-Rescue Tasks. Human factors, 51, 240-250.

Prytz, E, Montano, M., Kennedy, R., Scerbo, M., Britt, R., Davis, S., & Stefanidis, D. (2012). Using a spatial task to measure laparoscopic mental workload: Initial results. *Simulation in Healthcare, 7,* 511.

Reid, G.B. & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in Psychology, 52*, 185-218. doi: 10.1016/S0166-4115(08)62387-0

Sanderson, P. M., Liu, D., & Jenkins, S. A. (2009). Auditory displays in anesthesiology. *Current Opinion in Anesthesiology*, *22*, 788-795.

Scerbo, M. W., Kennedy, R. A., Montano, M., Britt, R. C., Davis, S. S., & Stefanidis, D. (2013). A spatial secondary task for measuring laparoscopic mental workload: Differences in surgical experience. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*, 728-732. doi: 10.1177/1541931213571159

Simon-Dack, S. L., Friesen, C. K., & Teder-Sälejärvi, W. A. (2009). Sex differences in auditory processing in peripersonal space: an event-related potential study. *Neuroreport*, *20*, 105-10.

Stefanidis, D., Scerbo, M. W., Korndorffer, J. R., Jr., & Scott, D. J. (2007). Redefining simulator proficiency using automaticity theory. *American Journal of Surgery, 193*, 502-506. doi: 10.1016/j.amjsurg.2006.11.010

Tendick, F., Bhoyrul, S., & Way, L. W. (1997). Comparison of laparoscopic imaging systems and conditions using a knot-tying task. *Computer Aided Surgery, 2*, 24-33. doi:10.3109/10929089709149080

Tole, J. R., Stephens, A. T., Harris, R. L., & Ephrath, A. R. (1982). Visual scanning behavior and mental workload in aircraft pilots. *Aviation, Space, and Environmental Medicine, 53*, 54-61.

Vecchi, T., & Girelli, L. (1998). Gender differences in visuo-spatial processing: The importance of distinguishing between passive storage and active manipulation. *Acta Psychologica*, *99*, 1-16.

Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception (TAP), 2*, 407-412.

Warvel, L. (2015). Assessing the effectiveness of an interval estimation and a visual-spatial secondary task as measures of mental workload during laparoscopy. In M. W. Scerbo, C. Brill, & B. Winstead (Eds.): ProQuest Dissertations Publishing.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, *94*, 111-123.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 239-257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3, 159-177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*, 449-455. doi: 10.1518/001872008x288394

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, *30*, 599-616.

Wierwille, W. W., & Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 25*, 1-16. doi: 10.1177/001872088302500101

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 35*, 263-281. doi: 10.1177/001872089303500205

APPENDIX A

PARTICIPANT BACKGROUND INFORMATION FORM

Participant #:_____ Group:_____ Date:_____ Time:_____

The purpose of this questionnaire is to obtain background information on the participant that will be used for research purposes only.

1. Age______
2. Gender______

0 = Female
1 = Male

3. Do you have any hearing impairments?____

0 = Yes
1 = No

4. Do you have normal or corrected-to-normal vision?_____

0 = Yes
1 = No

5. What is your dominant hand?_____

0 = Right
1 = Left
2 = Ambidextrous

APPENDIX B

PRIMARY TASK INSTRUCTIONS

Peg Transfer Task

1. Take a grasping tool in each hand.
2. Using the grasping tool in your nondominant hand, pick up a ring from one of the pegs.
3. Transfer the ring from the grasping tool in your nondominant hand to the one in your dominant hand.
4. Using the grasping tool in your dominant hand, place the ring on the side of the board that matches that hand.
5. Repeat these steps until all rings have been transferred.
6. When all of the rings have been transferred, complete the process again beginning with the dominant hand and ending when all of the rings are placed on the nondominant hand side of the pegboard.

Eraser and Bowl task

1. Take a grasping tool in your dominant hand.
2. Using the grasping tool pick up one of the erasers and place it into the bowl.
3. Repeat this process until all of the erasers are placed in the bowl.

APPENDIX C

SECONDARY TASK INSTRUCTIONS

Ball-and-Tunnel Task

1.You will be presented with the image of two balls.

2. The neutral positions for these balls will be at 3 o’clock and 9 o’clock.

3.Press the left foot pedal when one of the balls appears to have moved from the neutral position.

4.Press the left foot pedal when you are ready to begin.

5.Press the right foot pedal to exit.

Auditory Analog

1.You will be presented with two tones.

2. The neutral positions for these tones will be in the speakers directly to your left and right.

3.Press the left foot pedal when one of the tones appears to have moved from the neutral position.

4.Press the left foot pedal when you are ready to begin.

5.Press the right foot pedal to exit.

APPENDIX D

NASA-TLX WORKLOAD QUESTIONNAIRE

Participant #:______ Group:______

NASA-TASK LOAD INDEX (TLX) WORKLOAD QUESTIONNAIRE

(Hart & Staveland, 1988)

MENTAL DEMAND

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

PHYSICAL DEMAND

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

TEMPORAL DEMAND

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

PERFORMANCE

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

EFFORT

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

FRUSTRATION

Low High

|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

VITA

James P. Corcoran
Department of Psychology
106 Mills Godwin Building
Old Dominion University
Norfolk, VA 23529
E-mail: jcorc006@odu.edu

EDUCATION

MS, Experimental Psychology Old Dominion University, VA
May 2019 (expected)

BS, Psychology Old Dominion University, VA
December 2014

BA, Music College of William and Mary, VA
May 2008

WORK EXPERIENCE

Teaching Assistant, Old Dominion University August 2017 – May 2019

Teaching Assistant, Old Dominion University August 2016 – May 2017

Teaching Assistant, Old Dominion University August 2015 – May 2016

Teaching Assistant, Old Dominion University August 2014 – May 2015

RESEARCH

Corcoran, J.P. (2019) Comparing the effects of mental workload between visual and auditory secondary tasks during laparoscopy. Unpublished master's thesis. Old Dominion University. Norfolk, VA.