2010

Does dual task interference affect concurrent duration production?

Mary Merlau

Does dual task interference affect concurrent duration production?

Thesis Research
By
Mary Merlau
Bachelors of Human Development with a Concentration in Psychology,
Department of Psychology, Empire State College, 2007

Submitted in partial fulfillment of the requirements for the degree of
Master of Science in Applied Experimental and Engineering Psychology
in the Graduate College of the
Rochester Institute of Technology, Rochester, NY

Faculty Adviser and Chair of the Thesis Committee:

____

Andrew M. Herbert, Ph.D.
Chair, Associate Professor,
Psychology Department, Rochester
Institute of Technology

Reader:

____

Nicholas Difonzo, Ph.D.
Professor of Psychology, Rochester
Institute of Technology

Reader:

____

Kirsten Condry Ph.D.
Assistant Professor of Psychology,
Rochester Institute of Technology

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to graciously thank my adviser, Dr. Andrew Herbert for his encouragement, advice, support and guidance throughout this thesis. I would like to thank my committee members Dr. Nicholas Difonzo and Dr. Kirsten Condry for their insightful comments and guidance. I would also like to thank Dr. Esa Rantanen for providing me with some of Sandra Hart's articles. A very special thank you to Jason Plaisted who worked with me for months to code, test, fix, and retest the program for this experiment. I would not have been able to do this experiment without you. My dearest thanks to my mother, Gail Dempsey, for devoting her time and energy to my primary education, loving me unconditionally, supporting me and for having her Master's degree, you made me want to get mine even more. I would like to thank all of my participants for being a part of my study, I could not have done this without you. I would also like to thank Renee Walker for keeping me on track with all the deadlines for my thesis. I owe my deepest gratitude to my husband, William Merlau, for his moral support, love, understanding and patience through my entire college career. I love you.

Abstract

Time perception has been cited as a good measure of workload because it seldom interferes with performance of tasks that do not require time estimation, is easy to implement, and is sensitive to task difficulty (Hart, 1975a; Zakay & Shub, 1998). It was hypothesized that duration productions could be used to measure workload caused by two word tasks that would interfere with one another as outlined by Wickens' (1980) multiple resource theory. To test this hypothesis participants estimated a 15 second duration while performing a visual animal name detection task, along with an auditory animal name detection task, or a tone change task. The mean duration produced for the condition with two word tasks was not significantly different from the duration produced for the condition with a word task and a tone task. There was also interference between tasks that should not have produced interference. The results do not support the use of the concurrent duration production method to measure the workload caused by dual nontemporal task interference.

Does dual task interference affect concurrent duration production?

Multitasking, doing two or more tasks at once, is very common in today's world and has received much attention because of its dangers (e.g., texting while driving). Multitasking requires more attention to perform two or more tasks without failure, which causes more workload for an operator. Workload is a measure of how cognitively busy someone is (objective) or how cognitively busy someone thinks they are (subjective). High workload occurs when there is not enough attention to perform a task. When workload is high, task performance suffers and more errors occur.

Time perception is an interesting measure of workload because attention needed to perform other tasks can change one's perception of time. An example is the watched pot that takes a long time to boil. When one is waiting for a pot of water to boil it seems to take a long time, but if a friend calls and starts a conversation the pot of water will seem to boil sooner. The phone conversation takes attention away from monitoring time, which results in one perceiving less time passing. The more difficult a task is, the less attention time will get, which will result in a shorter perceived duration. The attentional gate model (AGM) of time (Zakay & Block, 1997) has been used to explain why time is perceived as going by faster or slower when attention is given to another task. The AGM explains that time will be perceived as going by faster when more attention is available to time and slower when less attention is available to time. This allows for a measure of how much attention is left over from the task that is being performed.

Temporal duration production (the estimation of a specified amount of time) is an objective measure that has previously been used to measure workload of tasks that do not require attention to time (nontemporal; Zakay & Shub, 1998). Estimates of time are an objective measure of workload because the amount of time that has passed can be observed and is a reflection of the attention resources available (or not) to time perception (Hart, 1976; 1978; Zakay & Shub, 1998). The more difficult a task is, the more workload it causes an operator and the longer duration estimates will be (Zakay & Shub, 1998). Duration estimates become longer with more workload because operators perceive less time passing. Therefore, more actual time would have to pass for it to feel like the correct duration. Even though duration productions have been used to measure single and dual task workload, it has not been researched as a measure of workload caused by dual task interference. The goal of this thesis was to determine if performing two word tasks that use the same verbal information processing resources would interfere more with the resources available for time estimation than a word and a tone task that do not use the same information processing resources as outlined by multiple resource theory (MRT; Wickens, 1980).

MRT has been used to predict how multitask performance will change based on the information processes each task uses. The information processes are stages of processing (perceptual / cognitive [input] or response [output]), modalities of processing (visual [eyes] or auditory [ears]), channels of visual information processing (focal or ambient), and codes of processing (verbal [words] or spatial [pictures or location]). Each information process has limited resources available. Therefore, when the resources are no longer available performance will suffer. The resource in question is an amount of attention or effort available to the information processes (Navon & Gopher, 1979; Wickens, 1981; 2007; 2008). For example, texting and driving will be more difficult to perform without failure than driving and listening to the radio because texting requires the same visual and manual resources as driving. This means that texting and driving must share attention. One cannot pay attention to the road and the phone at the same time, therefore performance on one or both of the tasks would deteriorate. Listening to the radio and driving do not use the same resources. Therefore, one can listen to a favorite song and pay attention to the road at the same time without deteriorated performance on either task.

Measuring the workload caused by two tasks that share resources can inform design or resource allocation changes (changes to the information processes tasks use) to reduce the workload on the operator. In the texting and driving example, workload would be reduced if the manual response resources and the visual resources used for texting could be changed to vocal response and auditory resources. For example, the driver could speak the text message (vocal response) into the phone and say send message. When a new message arrived, the driver could ask the phone to read the new message aloud (vocal response) and listen to the message (auditory).

A brief description of mental workload , and a brief review of using time perception to measure mental workload follow. Wickens' MRT (1980) provides the framework to explain which information processes share attention resources and cause performance decrements (task interference). Zakay and Block's (1997) AGM explains how attention affects time perception. Both will be discussed in detail.

What is Mental Workload?

Wickens (2002) stated, "mental workload describes the relation between the (quantitative) demand for resources imposed by a task and the ability to supply those resources by the operator" (p.161), while Hart and Staveland (1988) define workload as "a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance" (p.46). Both definitions specify that human operators must give something (resources/cost) of themselves to perform a task. The attention or effort demands (cost) placed on an operator are dependent on the difficulty of a task. Poorer performance and more workload occur when a task is so difficult that more resources are required than available. Difficulty refers to demands of individual task, while workload is what the operator experiences and can reflect loads imposed by one or more tasks.

There are many ways to make a task more difficult and impose more workload on an operator. Two examples are making a task faster, or adding more stimuli for an operator to observe. Multitasking adds more stimuli for an operator to observe. In the experiment reported below two tasks that use the same resources (as defined in MRT) were used to impose greater workload on an operator.

Measuring Workload with Time Perception

The measurement of workload is extremely important in government, aviation, military and medical applications. The measurement of workload allows for design changes or changes in task allocation for mentally demanding tasks. Such changes can be tested to determine if they reduced the perceived workload and increased operator performance. The next section reviews measuring workload with time perception, the concurrent duration production method, and limitations of time perception as a workload measure.

Time perception as a measure of workload.

Temporal performance becomes less accurate when a nontemporal task requires more attention (Hart, 1975a; 1975b; 1976; 1978; Hart & Bird, 1980). A difficult task requires more attention than an easy task and therefore will cause poorer accuracy in the perception of time. When performing two tasks, people consider time perception as the less important task (secondary task) unless instructed otherwise (Hart, 1975b; Zakay & Block, 1997; Zakay & Shub, 1998). Although there are mixed results (see Brown, 1997; 2006), a majority of the research confirms that a temporal task does not significantly affect the performance of a nontemporal task, while the nontemporal task affects the performance of the temporal task (Brown, 1997; Casali & Wierwille, 1983; Hart, 1975a; Zakay & Shub, 1998). The conclusion has been that time perception receives left over resources from the more important task. Time perception is a good measure of workload because the amount of left over resources from the primary task affects time perception, yet time perception does not affect primary task performance.

Concurrent duration production.

There are several time estimation methods within the time perception research. The experiment described below used a prospective concurrent duration production method. Prospective means that the participant knew in advance that they would be estimating a duration. Concurrent duration production means that the participant estimates a specified amount of time while simultaneously performing a nontemporal task. For example, a participant could be asked to estimate a 15 second duration by pressing a button to start the interval and pressing the button again when the 15 seconds has passed. At the same time, the participant is asked to perform a change detection task. Time perception is measured along with performance on the nontemporal task. Different combinations of nontemporal tasks can be used, and changes in duration estimates provide a measure of residual resources (attention).

The prospective concurrent duration production task was chosen over other temporal tasks because it does not require the participant to recall the amount of time that passes. Temporal reproduction or other recall tasks rely on short term memory rather than attention alone. In addition, the concurrent duration production task does not interfere with nontemporal task performance or perceived workload and has been positively correlated with subjective workload (Zakay & Shub, 1998). In this thesis, the NASA-TLX measured subjective workload.

Limitations of time perception as a workload measure.

Because time perception does not rely on spatial, verbal, visual, auditory, manual or vocal response resources, Brown (1997) suggested that timing tasks use perceptual / cognitive information processing resources within the MRT framework. Brown also suggested that when people count to themselves they use verbal processing resources. Thus, one would expect to see duration estimates become more variable when people count to themselves because of the shared verbal resources. Hart (1978) found that counting aloud only decreased the variability in duration productions and counting sub-vocally was no different than not counting at all. Thus, it is doubtful that time perception uses verbal processing resources.

For higher workloads, operators switch from a prospective method of timing to a retrospective method of timing (where they are not paying attention to time and must use memory to determine how much time has passed; Hart, 1978). Thus, when producing a 15 second duration the operators might get five seconds into the task, put the temporal task on hold (not paying attention to time), and then come back to the temporal task later trying to remember how long they had put it on hold. Participants would then overestimate how much time had passed stopping the trial before the 15 seconds had passed. When workload is increased further, operators tend to forget about the temporal task altogether. Forgetting a task is also known as task-shedding, and occurs when operators are overloaded (Raby, Mireille & Wickens, 1994). If overload did occur, one would expect shorter duration estimates than requested or extremely long durations caused by forgetting the temporal task (Hart, 1978).

Temporal tasks have been shown to interfere with math tasks such as mental subtraction, finding incorrect answers among correct answers to subtraction problems, random number generation tasks and other tasks that integrate information, oversee multitasking, and coordinate actions (Brown, 1997; 2006). The use of perceptual / cognitive resources for time perception would explain why timing tasks interfere with math performance, but would not explain why temporal tasks would not interfere with other tasks such as visual search and pursuit rotor tracking that also use perceptual / cognitive information processing. Brown (1997) suggested that a temporal task does not use enough resources to have an effect on nontemporal task performance, but the nontemporal task uses enough resources to have an effect on temporal performance. Math tasks were not used in the experiment reported below given that Brown (1997; 2006) has shown that time perception interferes with math tasks.

Time Perception

Attentional gate model.

The AGM is the only prospective time perception model that can explain the effects attention has on time perception. The AGM is very similar to the temporal information processing model (TIP; Church, 1984; Gibbon & Church, 1984) but added an attention component to time perception. As shown in Figure 1, the AGM has a pacemaker (time generator influenced by arousal) that emits pulses, an attentional gate that would allow pulses through based on available attention resources, a switch which closes and opens based on the beginning of a target interval, an accumulator that gathers and counts the pulses, a pathway to either working memory or reference memory, and a cognitive comparison process to determine if the duration of time passed matches a previous duration stored in either working memory or reference memory (Block & Zakay, 2006).

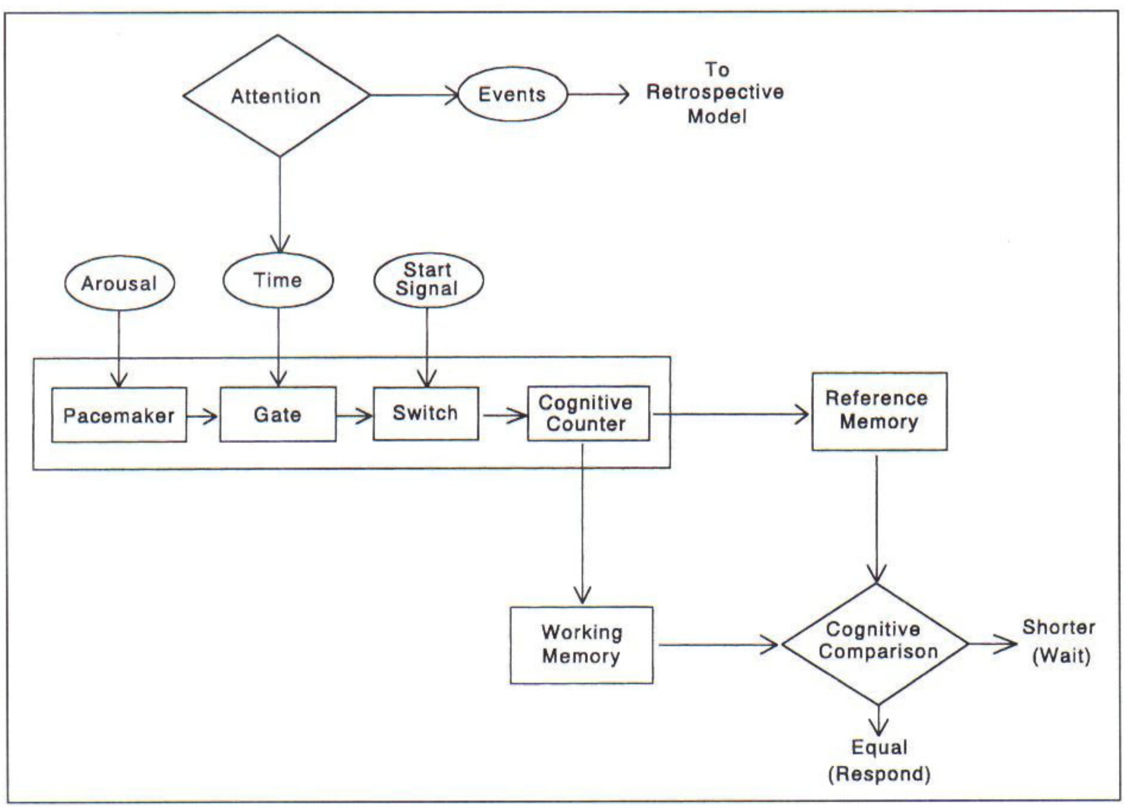


Figure 1. The attentional gate model of time, From "Temporal Cognition," by D. Zakay and R.A. Block, 1997, Current Directions in Psychological Science, 6, p.14. Copyright [1997] by the American Psychological Society. Reprinted with permission.

The focus here is on the attentional gate because it explains how and why time perception is affected by sharing attention with other tasks. The AGM is analagous to crowds trying to enter Disney World. Only one person (pulse) can get through a turnstile (attentional gate) at a time. The number of people who get into the park at a time is dependent on the number of turnstiles open (amount of attention available to the attention gate). If more turnstiles were open (more attention is available to time), then more people could get in at once (more pulses would be allowed through the gate). According to the AGM, when a task is difficult fewer attention resources will be available to pay attention to time. This results in a duration perceived as shorter than it actually is. For example, if asked to estimate a 15 second duration while also performing a difficult task it would take more than 15 seconds to perceive 15 seconds passed because less attention is available to perceive time passing, which means 18 seconds may pass but only 15 seconds would be perceived as passing. This is expected to occur when two tasks use the same attention resources as outlined by MRT, because they would be more difficult to perform than two tasks that did not use the same resources.

Time perception and arousal.

As mentioned above, arousal level influences the number of pulses the pacemaker generates. If one's arousal level increases, it causes the pacemaker to produce more pulses. More difficult trials are expected to be more arousing, causing the pacemaker to produce more pulses, which would result in participants perceiving more time passing and thereby producing a shorter duration. Even though the AGM shows arousal affecting the number of pulses the pacemaker emits, the attentional gate would only allow so many pulses through depending on the attentional resources available to it. Going back to the

Disney example, it would not matter how many people were waiting to get into the park (the number of pulses generated by the pacemaker), it would only matter how many turnstiles were open (how much attention was given to time). Therefore, it was hypothesized that if arousal levels increased they would not have an effect on durations produced. The modified version of Thayer's (1967) self-report rating scale (Wearden, Pilkington, & Carter, 1999) measured arousal in this study.

Task Interference

Wickens' multiple resource theory.

Wickens (1980) performed a meta-analysis of literature on structural and capacity theories of attention and dual-task performance and developed MRT. According to MRT, the more two tasks share resources the greater the performance degradation would be on one or both of the tasks (Wickens, 2007).

Task interference occurs when there are not enough resources to perform two or more tasks at the same level simultaneously as independently (Wickens, 1981). The relevance of MRT to mental workload lies in its ability to predict performance decrements caused by task interference (Wickens, 2002; 2008). For example, it is much easier to talk on a cell phone and walk than it is to text while walking. Talking on a cell phone requires auditory and verbal information processing resources, walking requires visual (focal and ambient), spatial and manual response information processing resources, and texting requires visual (focal), verbal and manual response information processing resources. Therefore, texting while walking is more difficult, will cause an operator more workload, and could result in more errors because both tasks require visual (focal) and manual response information processing resources.

The use of the MRT framework has not been applied to temporal perception and its ability to measure workload before. As discussed above, both MRT and time perception use attention as a resource for performance. MRT provides the framework for choosing tasks that should interfere and cause more workload. Two tasks that interfere and cause more workload should require more attention than two tasks that do not interfere, which would leave less attention for time and cause a longer duration to be produced. One visual and two auditory tasks were chosen. The visual task was performed with each auditory task. The difference between the two auditory tasks was the code of processing that each task used. The visual task used the verbal code of processing and was performed with either an auditory verbal task that also used the verbal code of processing or an auditory tonal task that did not use the verbal code of processing. When performed together the two verbal tasks should interfere with one another's performance because they must share verbal processing resources. The visual verbal and auditory tonal task should not interfere with one another's performance because they do not share the verbal processing resource.

Purpose of the Research

Although previous research (Hart, 1978; Zakay & Shub, 1998) shows duration production to be a good measure of single and dual task difficulty and workload, there has been no previous research on using duration production to measure the workload caused by dual nontemporal task interference. The goal was to see if duration production increased more when two nontemporal tasks used the same codes of processing than when two nontemporal tasks did not use the same codes of processing.

It was hypothesized that longer durations would be produced when two nontemporal tasks used the same processing resources compared with shorter durations produced when tasks did not use the same processing resources. Previous research showed that the concurrent duration production task did not interfere with most nontemporal tasks, and duration productions increased as nontemporal task difficulty increased. Therefore, it was assumed that the concurrent duration production task would not interfere with the dual nontemporal task performances, and would be able to measure workload caused by dual nontemporal task interference.

Experimental Tasks

Participants performed four different tasks. The tasks were a concurrent duration production task (D), a rapid serial visual verbal (word) presentation task (VV), a rapid serial auditory verbal (word) presentation task (AV), and an auditory tone detection task (AT). The D task required participants to press the space bar to start a 15 second duration. After the participant estimated 15 seconds had passed; pressing the space bar again ended the trial. The VV task required participants to watch words presented on a computer screen one after the other and press the ‘z’ key every time they saw an animal word. The AV task required participants to listen to words being spoken one after another and press the ‘/’ key every time they heard the name of an animal. The AT task required participants to listen to tones being played one after another and press the ‘/’ key every time they heard the tone change. The VV, AV, and AT tasks also required the participant to press the space bar to start the task. When there was not a D task, the task would last 15 ± 1 seconds and then stop on it’s own.

Participants performed each of the four tasks individually to get baseline performance. Each participant also performed each nontemporal task concurrently with the D task to get a baseline of duration production with each task and performance on each task when performed concurrently with the D task. As described above, the duration estimates should increase when performed with each nontemporal task, but the nontemporal task performance should not change from baseline.

Participants performed the VV task concurrently with the AV task to get a baseline performance for dual tasks. According to MRT, the AV and VV tasks should interfere with one another. Performance on one or both of these tasks performed simultaneously should be worse than the baseline performance for each task alone because they both use the verbal code of processing. According to MRT the AT and VV task should not interfere with one another because they do not use the same code of processing. Performance on both of these tasks should not be statistically different from their baseline performance.

Participants performed the D, VV, and AT tasks simultaneously to get a measure of duration production with dual nontemporal tasks that should not interfere with one another's performance. The VV and AT performance should not be statistically different from the baseline dual task performance. Duration productions should be longer than the duration production for any of the single task baselines, but shorter than the duration productions for the D, VV, and AV condition.

Participants performed the D, VV, and AV tasks together to get a measure of duration production with dual nontemporal tasks that should interfere with one another's performance. The VV and AV performance should stay the same as the baseline for dual task performance, but the duration productions should be longer that the duration productions for the D, VV, AT condition.

Temporal productions have previously been positively correlated with subjective workload (Zakay & Shub, 1998). It was predicted that the temporal productions would be positively correlated with subjective workload for each condition.

Arousal level can affect temporal perception (Zakay & Block, 1997) and is part of the AGM. It was predicted that arousal levels would not change during the course of this study.

Methods

Participants

Thirteen participants (5 females, 8 males *M*age = 23 years, age range: 19 - 26 years) were recruited from the population of undergraduate and graduate students on the Rochester Institute of Technology campus using posters and e-mails. Participants had normal hearing and normal or corrected to normal visual acuity. All participants had been speaking English for at least five years and were able to recognize the animal names in the study. Participants were entered into a drawing to win one of two $50.00 cash prizes.

Apparatus & Stimuli

There were four different tasks performed during the thesis experiment. The tasks were a concurrent duration production task (D), a rapid serial visual verbal presentation task (VV), a rapid serial auditory verbal presentation task (AV), and an auditory tonal detection task (AT).

The rapid serial tasks were devised using Microsoft Access. A main word list for the VV and AV tasks was developed from word lists found online as well as adding words the experimenter could think of. There were 82 target animal name words and 178 non-target words (see Appendix A). A separate word list was developed using words from the main list for each trial that involved a verbal task. The word lists were randomized for each trial. For example, in a trial that had a VV and AV task there would be two separate word lists developed. The main word list was randomized and the first 60 words were chosen for the VV list. The main list was randomized again and the first 60 words were chosen for the AV list. This was done for every trial. Each participant performed the randomized trials in the same order and saw the same randomized word list developed for each trial. The words were presented in Tahoma font size 20 and instructions for each task were presented in Tahoma font size 12.

Each word in the main word list was recorded individually by a female voice using Audacity software (for digital recording and editing of sounds). Once the word was recorded the blank space before and after the word was spoken was trimmed, leaving only the word being spoken. The duration it took for each word to be spoken was used to determine how long the word would be presented visually, so word duration was identical regardless of presentation modality.

Each trial that did not have a duration task lasted 15 ±1 seconds. The word lists for the trials that did not have a duration task contained 60 words, the participant may not have seen all sixty words, but this ensured that the program did not run out of words before the trial ended. The word lists for trials that did have a duration task had at least 120 words so if the participant did not stop after 15 seconds the program had enough words to go through. Because each participant could produce a different duration on the D task, some participants may have seen or heard more words within a word list.

Two notes, B and F, were recorded using an ocarina (wind instrument, similar to a flute) and Audacity software. Each note had 6 different durations that it would last, .5 s, .75 s, 1s, 1.5s, 2s, and 5s. Ten intervals of each note duration were randomized for trials without the D task. This list of 60 notes was then added to one trial with an AT task without the D task. For trials with the D task, 20 of each note duration were randomized. The list of 120 notes was then added to one trial with an AT task that would be performed with the D task. A list of notes was randomized for each trial that had an AT task. The list of 120 notes would be 1 minute 47.5 seconds and ensured that there would be enough notes for the program to go through if the participant went over 15 seconds. All participants heard the same randomized tone list developed for each of the trials with an AT task.

When the auditory words or notes changed in the program there was a blip sound that was caused by the program moving from one sound file to another. Participants were made aware of this sound during the instructions and were able to experience it during the practice trials. They were told to ignore this sound.

A pre-research questionnaire (see Appendix B) was given that asked participants their date of birth, sex, major, if they had any formal music training (and if so how much in years), and if they had any formal dance training (and if so how much in years). After the pre-research questionnaire was developed it was determined that these data would not be analyzed because if people with music or dance training (that have been trained to count to themselves to keep time) counted to themselves it would not affect their duration estimates as long as they did not count out loud (Hart, 1978).

NASA-TLX.

The NASA-TLX has been used for more than 20 years to measure subjective workload in various multitask environments (Hart, 2006; Rubio, Diaz, Martin, & Puente, 2004). As of 2006, the NASA- TLX had been used or reviewed in more than 550 papers (Hart, 2006). Given the NASA- TLX's extensive use and its ability to successfully measure subjective workload in many situations, it measured subjective workload in this thesis.

The NASA-TLX uses one question about each of six areas of workload. The areas of workload (and questions) are: mental demand (how mentally demanding was the task?); physical demand (how physical demanding was the task?); temporal demand (how hurried or rushed was the pace of the task?); performance (how successful were you at accomplishing what you were asked to do?); effort (how hard did you have to work to accomplish your level of performance?); and frustration level (how insecure, discouraged, irritated, stressed and annoyed were you?). Hart and Staveland (1988) developed the original NASA-TLX. The electronic version used in this thesis was adapted from the paper and pencil version (available online at http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLXScale.pdf). Operators gave each of the six questions (one regarding each area of workload) a rating, usually ranging from 0-100, after performing each task (Hart & Staveland, 1988). These ratings were then added up and divided by the number of workload areas (6) to get an overall measure of workload.

The adapted NASA-TLX used a sliding scale for each question that did not display numbers or ticks, only end anchor points labeled *very low* and *very high* (see Appendix C). The program recorded the placement of each slider, which ranged from 0-

100. The rating scale for performance went in the opposite direction (100-0) from the other tasks, so the higher the slider went, the lower the rating, and had anchor points of *failure* and *perfect.* The performance rating scale went in the opposite direction of the paper and pencil version's performance scale to reduce confusion by allowing the anchor points (low on the left and high on the right) to be consistent across the scales. The rating each participant gave to each of the six questions was summed and divided by six, which gave an overall workload rating out of 100 for each condition.

Arousal.

Heart rate and galvanic skin response have previously been used to measure attention and arousal respectively (Angrilli, Cherubini, Pavese, & Manfredini, 1997; Wickens & Kramer, 1985; Wierwille & Eggemeier, 1993; Thayer, 1967). Thayer (1967) had participants rate activation adjectives on a 4-point rating scale (definitely feel, slightly feel, cannot decide and definitely do not feel). Thayer (1967, 1970) found that the self-report scale that he developed was a valid measure of activation (arousal) and was significantly correlated with heart rate and galvanic skin response measures. Wearden, Pilkington, and Carter (1999) used a modified version of Thayer's (1967) self-report rating scale that used the same 4-point rating scale, but only used two randomly chosen adjectives from each of the original 4 dimensions (general activation, high activation, general deactivation, and deactivation – sleep) to measure arousal. They found that arousal decreased as trials proceeded.

An arousal rating scale adapted from Thayer (1967) and used by Wearden et al. (1999; see Appendix D) was adapted for use in Microsoft Access and used to measure arousal. Each verb was rated on a scale of 1-4, with 1 - *definitely feel*, 2 - *feel slightly*, 3 -

cannot decide and 4 - *definitely do not feel*. The verbs were divided into activation and deactivation. Activation ratings consisted of the mean rating given to lively, jittery, energetic and stirred up. Deactivation ratings consisted of the mean rating given to calm, drowsy, relaxed, and sleepy (Wearden et al., 1999).

A post-research questionnaire was given that asked participants:

At any point did you find yourself counting to keep track of time?

If yes, do you remember when?

Do you think that your arousal level changed at all during this research?

If yes, at what point do you think your arousal level changed?

(see Appendix E for screen shot of questionnaire). After the post-research questionnaire was developed it was determined that the data would not be analyzed because counting to oneself would not have an effect on duration productions (Hart, 1978) and the arousal scale used by Wearden et al. (1999) was determined to be a more sensitive measure of arousal changes throughout the study.

Procedure

A within subjects design was used. Practice effects and arousal were controlled by randomizing the order of tasks performed and the word lists for each trial. Each participant performed the same random order of trials and saw the same word lists. In this study the arousal rating scale (Wearden et al., 1999; adapted from Thayer, 1967) measured arousal and the NASA-TLX (Hart & Staveland, 1988) measured subjective workload .

There were 11 different conditions. Each participant performed one practice for each condition. Each participant then performed 55 trials consisting of five of each condition (see Table 1 for conditions) in the same order.

Table 1
Tasks and Combinations of Tasks Performed in this Thesis With and Without Duration

Tasks	
With	Without
D	
D VV	VV
D AT	AT
D AV	AV
D VV AT	VV AT
D VV AV	VV AV

Each participant signed an informed consent form (see Appendix F) before starting the study. Participants were seated in front of a computer. They were asked to remove all timing devices. An overview of the study was read to each participant (see Appendix G). The participants filled out a pre-research questionnaire. The participants then performed a practice of one of each of the 11 conditions. Each trial had instructions listed. Participants could take a break after each trial if they wanted, because they were in control of when the next trial started. Participants performed all trials in the same random order. The NASA-TLX was administered after the last trial of each condition. The arousal scale was administered after every 7 - 12 trials. Once the participant was finished with the trials they answered a post-research questionnaire. They were then entered into a raffle and were free to leave.

Analysis

Outliers.

The mean data from the duration productions, the visual verbal performance, the auditory tonal performance, the auditory verbal performance, the NASA-TLX, and the arousal scale were each plotted. Box plots were generated, each data point represented the mean of the trials for that task per participant. Outliers were removed by removing the highest or lowest trial for the participant indicated to be an outlier for that task until the participant no longer showed as an outlier. An outlier was any data point lying outside the whiskers.

There were four outliers across the six task conditions with the D task (see Appendix H for more details). There were 10 outliers across the 14 task conditions (VV, AV and AT task performance are separated in the dual task conditions) for nontemporal task performance (see Appendix I for more details).

The NASA-TLX overall workload ratings had one participant who showed as an outlier. Participant 7 was removed from the mean NASA-TLX ratings because of the extremely low overall workload ratings given (see Appendix J). It can only be assumed that participant 7 did not understand what was being asked in the NASA-TLX and did not follow instructions to answer based on the task that was just performed.

There were several outliers in the arousal data. There was no sensible way to remove these outliers because of the high variability of the arousal data (see Appendix K).

Performance Measures.

Microsoft Access recorded performance for duration estimation. System functions calculated time at the processor level with millisecond accuracy. The duration productions were rounded to a tenth of a second for analysis. The performance for the practice trials was not analyzed except for the arousal ratings, which were used as a baseline for participant arousal. Performance for the VV, AV and AT tasks was also recorded by Microsoft Access as the number of targets hit, the number of targets missed, and the total number of targets presented during each trial for each type of task. The number of hits was then divided by the total number of targets presented to get a proportion of correct responses for each type of task. For example, if a VV and AT task were performed concurrently the number of VV target hits would be divided by the total number of VV targets presented during that trial and the number of AT target hits would be divided by the total number of AT targets presented during that same trial. Internal reliability for the five trials in each condition except the duration condition was tested using Cronbach's Alpha (see Appendix L). The duration condition was not tested for internal reliability because it was the same for each of the five trials, while the word lists in each of the other conditions trials were different.

Microsoft Access recorded NASA-TLX and arousal ratings. For each NASA-TLX that participants filled out, the rating given for the six scales were added up and then divided by six to get the mean subjective workload. The arousal scale ratings[1] were divided into activation and deactivation ratings. Activation ratings consisted of the mean ratings given to lively, jittery, energetic and stirred up. Deactivation ratings consisted of

[1] Note: two of the arousal scales were removed because they occurred after the same type of task, and there was no way of knowing which one was performed first. It is important to know the order that the arousal scales were performed because it is assumed that arousal will decrease as time goes on.

the mean ratings given to calm, drowsy, relaxed, and sleepy. Cronbach's alpha was used to test internal reliability for the NASA-TLX rating scales as well as arousal activation and deactivation ratings (see Appendix M).

Order effects.

Order effects were analyzed for each condition. The D, VV, AT condition was the only condition that displayed order effects. As trials progressed from 1-5, the duration productions and auditory tonal performance improved. The more participants performed this task, the better they became on the auditory tonal task, which allowed more attention for time perception, and resulted in durations closer to the requested time over trials.

Results

Hypothesis: Was the Temporal Task Able to Measure Workload caused by Dual Task Interference for Two Nontemporal Tasks That Share Resources?

The D task did not differentiate workload caused by dual task interference as shown in Table 2. Contrary to the hypothesis, there was no significant difference in the duration produced for the D, VV, AT task condition (M=24 s, SD=6.21 s) and the duration produced for the D, VV, AV task condition (M=21.8 s, SD= 6.17 s); $t(12)=1.7$, $p > .05$, $r = .44$.

Did the Concurrent Duration Production Task Affect Nontemporal Performance?

Yes, the D task improved AT task performance, but did not have an effect on VV performance or AV performance (see Table 2). Paired *t*-tests were performed to test for significant differences in nontemporal task performance when performed alone and when performed with the D task.

Table 2

Mean Duration Productions in Seconds, NASA-TLX Subjective Workload Ratings and Performance (SD in parentheses) and [Medians in brackets]

Task	Mean duration produced for 15 second Target	Duration paired *t*-test	NASA-TLX ratings		NASA-TLX workload ratings paired *t*-test	Mean performance in proportion of correct responses		Performance paired *t*-test
			Without duration	With duration		Without duration	With duration	
D	13.8 (4.07)	N/A	N/A	27.5	N/A	N/A	N/A	N/A
AT	19.6 (5.05)	$t(12)=4.76$ $p= < .05$, $r=.81$	27 (12.25)	44.9 (11.97)	$t(11)=-8.93$ $p= < .05$	.66 (.05) [.66]	.89 (.14) [.95]	$t(12)=-7.49$, $p < .05$, $r=.91$
AV	19 (4.71)	$t(12)=4.16$ $p= < .05$, $r=.77$	31.9 (15.55)	42.9 (17.81)	$t(11)=-4.6$ $p= < .05$	.41 (.11) [.40]	.42 (.14) [.37]	$t(12)=-.25$, $p > .05$, $r=.07$
VV	19.4 (5.44)	$t(12)=4.55$ $p= < .05$, $r=.80$	27.3 (12.97)	32 (15.83)	$t(11)=-1.4$ $p= > .05$	.64 (.08) [.63]	.60 (.09) [.60]	$t(12)=1.77$, $p > .05$, $r=.46$
VVAT	24 (6.2)	$t(12)=5.45$ $p= < .05$, $r=.84$	51.8 (14.04)	43.4 (18.65)	$t(11)=2.3$ $p= < .05$	.50 (.14) [.47], .64 (.15) [.72]	.44 (.15) [.38], .65 (.14) [.60]	See text
VVAV	21.8 (6.17)	$t(12)=6.32$ $p= < .05$, $r=.88$	45 (17.47)	58.5 (15.09)	$t(11)=-4.9$ $p= < .05$	.41 (.07) [.43], .30 (.16) [.28]	.52 (.14) [.56], .29 (.04) [.30]	See text

Did the Two Verbal Tasks Interfere with One Another's Performance?

As predicted by MRT, the two verbal tasks interfered with one another's performance (see Figure 2). Paired *t*-tests were performed to see if VV or AV performance was significantly different when performed alone and when performed concurrently.

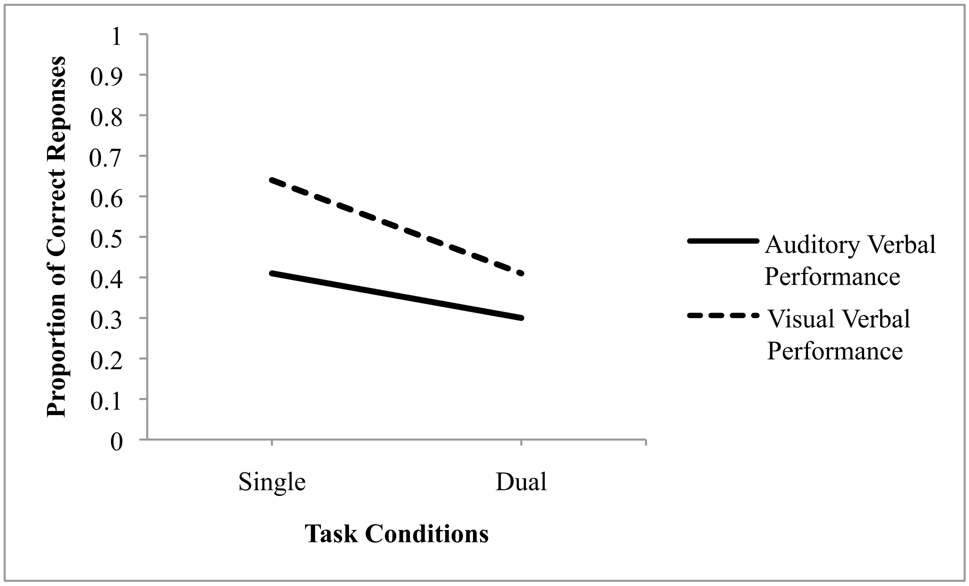


Figure 2. Mean visual verbal and auditory verbal performance in single and dual task conditions. Visual verbal and auditory verbal performances were better in the single task condition compared to the dual task condition.

VV task performance was better when performed alone (M=.64, SD=.08) than when performed concurrently with the AV task (M=.41, SD=.07); $t(12)=10$, $p < .05$, r = .94. AV task performance was also better when performed alone (M=.41, SD=.11) than when performed concurrently with the VV task (M=.3, SD=.16); $t(12)=4.69$, $p < .05$, r = .8.

VV task performance improved when the D task was added to the VV, AV task condition (see Figure 3). VV task performance was worse when performed with the AV task (M=.41, SD=.07) than when performed with the D and AV tasks (M=.52, SD=.14);

t(12)=-4.09, $p < .05$, r =. 76. As predicted, there was not a significant difference between AV task performance when performed with the VV task (*M*=.3, *SD*=.16) and AV task performance when performed with the D and VV tasks (*M*=.29, *SD*=.04); *t*(12)= < 1.

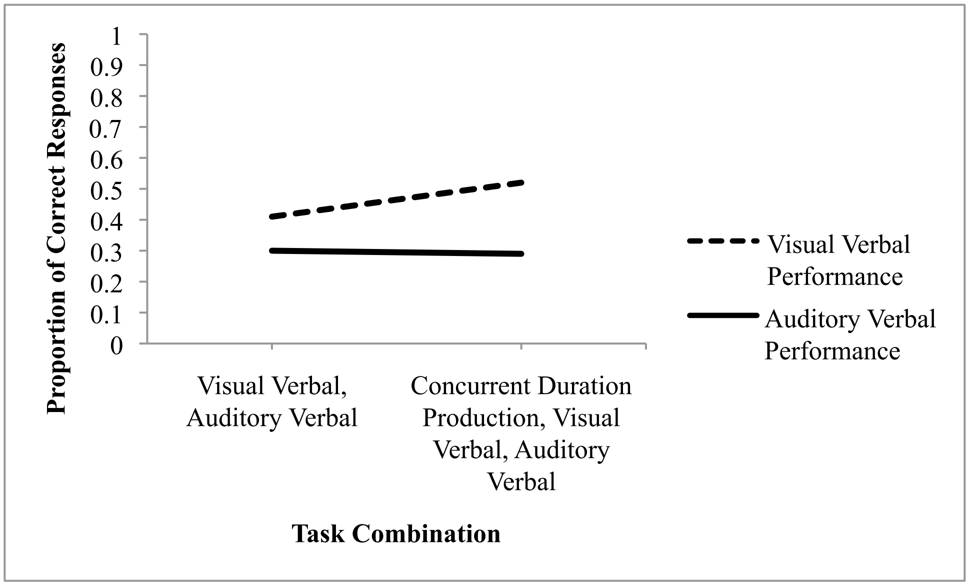


Figure 3. Visual verbal and auditory verbal task performance in the visual verbal, auditory verbal task combination and in the concurrent duration production, visual verbal, and auditory verbal task combination. Visual verbal performance increased when the concurrent duration production task was added.

Did Two Tasks That Do Not Share Codes of Processing Interfere with One Another's Performance?

Yes, the VV task performance was worse when performed with the AT task (see Figure 4). VV task performance was better when performed alone (*M*=.64, *SD*=.08) than when performed concurrently with the AT task (*M*=.50, *SD*=.14); *t*(12)=5.3, $p < .05$, r = .84. There was no such difference for AT task performance (*M*=.66, *SD*=.05; *M*=.64, *SD*=.15); *t*(12)= < 1.

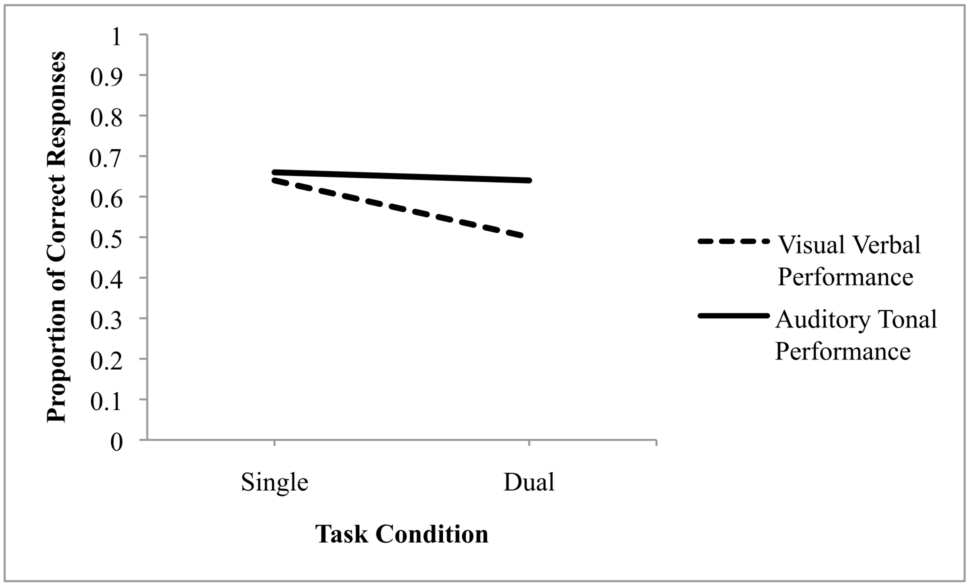


Figure 4. Visual verbal and auditory verbal performance in single and dual task conditions. Visual verbal performance decreased when performed with the auditory tonal task.

Further, VV task performance decreased when the D task was added to the VV, AT task combination (see Figure 5). VV task performance was better when performed concurrently with the AT task (M=.50, SD=.14) than when performed concurrently with the D and AT tasks (M=.44, SD=.15); $t(12)=2.75$, $p < .05$, $r = .62$. There was not a significant difference between AT task performance when performed with the VV task (M=.64, SD=.15) and the AT task performance when performed with the D and VV tasks (M=.65, SD=.14); $t(12)= < 1$.

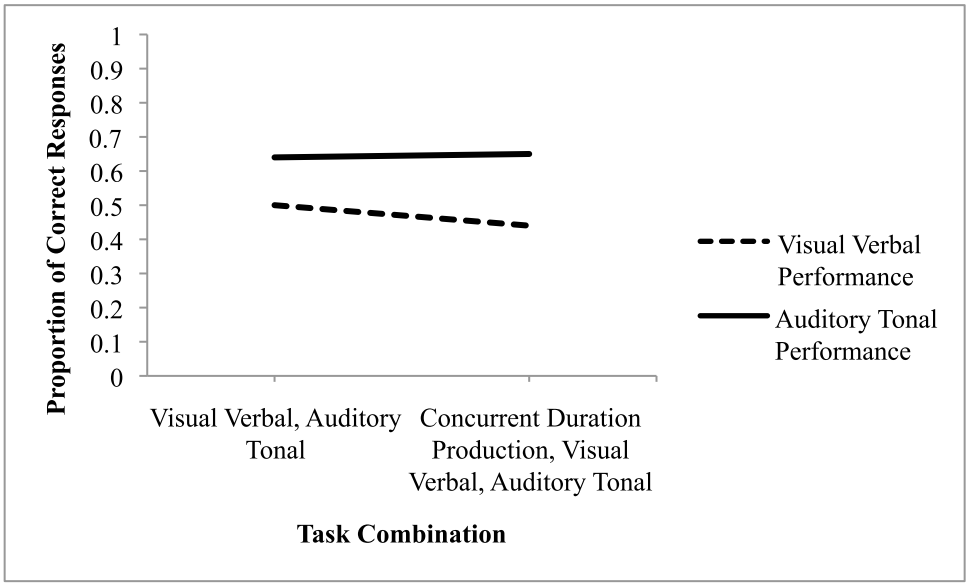


Figure 5. Visual verbal and auditory tonal performance in the visual verbal, auditory tonal and concurrent duration production, visual verbal and auditory tonal task combinations. Visual verbal performance decreased when the concurrent duration production task was added.

Did the Nontemporal Tasks Affect the Concurrent Duration Production Task?

Yes, the more nontemporal tasks that were performed simultaneously with the D task, the longer the produced durations became (See Table 2).

A within-subjects repeated measures ANOVA was performed to test the effect of the baseline temporal task and nontemporal tasks on duration production. The results show significant differences between durations produced for the baseline temporal task and nontemporal tasks $F(5, 60)=15$, $p < .05$. Paired *t*-tests were performed to examine the differences between the baseline duration and duration produced while concurrently performing the different nontemporal task combinations. The mean duration produced for the baseline duration task was significantly shorter than all of the mean durations produced concurrently with any of the nontemporal task conditions (see Table 2).

Do Subjective NASA-TLX Workload Ratings Correlate with the Duration Productions from the Concurrent Duration Production Task?

Yes, NASA-TLX workload ratings did correlate with the durations produced. A Pearson correlation was performed for mean NASA-TLX ratings and mean duration produced for each task per participant. The NASA-TLX subjective workload ratings were significantly positively correlated with mean duration produced for each participant, r =.27, p= < .05.

Post hoc paired *t*-tests were performed to see if subjective workload ratings were influenced by the addition of the D task to the three nontemporal tasks, the addition of the D task to the dual nontemporal task conditions, and dual task interference.

The addition of the D task caused an increase of subjective workload ratings for both auditory tasks (see Table 2 and Figure 6). This indicates some interference that caused more subjective workload with the D task and auditory tasks.

In the VV, AT task condition the addition of the D task caused a decrease in subjective workload, while in the VV, AV task condition the D task caused an increase in subjective workload (see Table 2 and Figure 6). The subjective workload showed that participants experienced more workload on the D, VV, and AV task condition and less workload on the D, VV, and AT task condition.

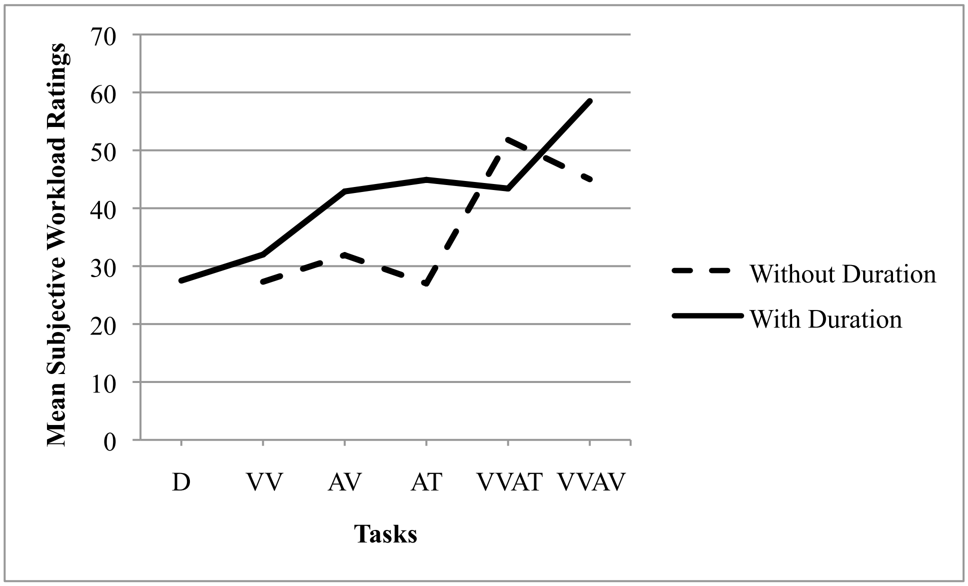


Figure 6. NASA-TLX ratings for conditions with and without duration task. Subjective workload ratings were higher for the VV, AV, AT, and VVAV tasks when performed with the D task, while the VVAT task had more subjective workload when performed without the D task[2].

Did arousal level change during the study?

Arousal levels did not change over time during the study. A repeated measures ANOVA was performed to see if the ratings were significantly different from one another. The Greenhouse-Geisser test was used because Mauchly's test of Sphericity was violated ($x^2(44)=148.4$, $p < .05$). The arousal ratings were not significantly different over time $F(1.57, 18.83) = < 1$.

Discussion

Although durations produced in the D task changed with the number of nontemporal tasks participants performed concurrently, they did not change with the difficulty between tasks as suggested by performance. The durations did not show

[2] Each of the six workload scales were also analyzed individually to determine if any one scale showed more workload than another scale. All of the workload scales except physical demand showed the same pattern as seen in Figure 6. Physical demand was low for all tasks and did not change much from one task to another (see Appendix N)

differences associated with subjective workload, so duration production was not a measure of the workload caused by dual nontemporal task interference. Unlike previous research, the D task was an unpredicted confounding variable for subjective workload, AT task performance, and VV task performance. The AT task performance increased when performed with the D task, while the VV task performance increased or decreased when the D task was performed concurrently in triple task conditions. The VV performance was also affected by the AT task. Subjective workload increased for the auditory tasks when they were performed with the D task. Subjective workload for the VV, AV task was lower than the subjective workload for the VV, AT task, but when the D task was added subjective workload was greater for the VV, AV task than for the VV, AT task.

Mean duration production was very close to the requested 15 seconds when performed alone (Table 2). This suggested that participants were able to get very close to the requested duration when not performing any other tasks. Mean duration production increased when participants concurrently performed a nontemporal task and increased again when participants concurrently performed two nontemporal tasks. These results support Zakay's (1998) previous research that suggested participants naturally allocate resources across temporal and nontemporal tasks when not instructed on which task is more important.

The nontemporal task performance suggests that mean duration production was not sensitive to the difficulty of different nontemporal tasks. AV performance was worse than VV or AT performance, yet the duration produced for each of the nontemporal tasks was approximately 19 seconds and was not significantly different (Table 2). This was interesting because in previous research duration productions were sensitive to the difficulty within and between nontemporal tasks (Zakay & Shub, 1998, experiment 1). This suggested that the amount of workload to perform all three tasks was similar, even though performance was different. The task participants performed and their performance did not affect duration productions, but the number of tasks participants performed did. Previous research showed that the manipulation of task difficulty affected duration productions. This thesis manipulated task difficulty by adding a second nontemporal task that had the same or different code of processing as outlined by MRT. The addition of the second task caused duration productions to increase, but the code of processing did not affect duration productions. One conclusion would be that even though two nontemporal tasks that use the same codes of processing may be more difficult to perform together (due to the use of verbal processing), they may not use more cognitive / perceptual resources than two nontemporal tasks that do not share codes of processing resources, which would be why there was no difference in mean duration between the D, VV, and AV task condition and the D, VV and AT task condition. The above conclusions are speculative, and await further research controlling possible confounding variables.

The addition of the D task caused poorer VV performance in the D, VV, AT task condition. However, the D task improved VV task performance in the D, VV, and AV task condition. These results appear to support Brown's (1997) hypothesis that temporal tasks use the verbal code of processing. However, the improvement or decrement in VV performance was dependent on the other nontemporal task that participants performed with it. This leads one to believe that the improvement or decrement in VV performance could be due to the demands placed on the perceptual / cognitive resources. These results support an energetic model of MRT (Wickens, 1991), which allows more resources to become available across all resources when one resource is in high demand. According to this energetic MRT, more perceptual / cognitive resources than previously available were required in the D, VV and AV task condition, which caused an increase in resources (more effort was put forth). This in turn caused more subjective workload (discussed later), a shorter produced duration, and an improvement in VV performance. The D, VV, AT task condition did not cause enough demand for the perceptual / cognitive resources to increase, which caused less subjective workload, and a decrement in visual verbal performance. Therefore, another conclusion would be that the single nontemporal tasks happened to use the same amount of perceptual / cognitive resources, while the dual word tasks caused more workload that was not seen in duration productions due to an increase in perceptual / cognitive resources for the D, VV, AV task condition.

The D task was not the only task that caused unpredicted effects. The AT task caused a decrement in VV task performance, which was not predicted by MRT. The fact that the AT task did have an effect on VV performance suggests that these two tasks share a common resource. The common resource that they share can only be speculated to be perceptual / cognitive resources.

The NASA-TLX subjective workload ratings showed that when the nontemporal tasks were performed alone the AT and VV task caused the same amount of subjective workload, while the AV task caused slightly more subjective workload. The D task increased subjective workload when paired with an auditory task (Table 2). Subjective workload did not consistently increase from the dual nontemporal task conditions to the triple task conditions, which suggests that the D task does not always cause an increase in subjective workload when paired with auditory tasks.

In Zakay and Shub's (1998) study the D task had no effect on subjective workload ratings. However, Zakay & Shub (1998) used a between subjects design, the Cooper-Harper subjective rating scale (Cooper & Harper, 1969) and did not have participants perform any auditory tasks. Since the NASA-TLX has a temporal demand scale, it is possible that the NASA-TLX is more sensitive than the Cooper-Harper rating scale to measure the workload caused by the D task. More research would be needed to test the sensitivity of the two subjective workload rating scales with duration productions. The within subjects design that was implemented in this thesis may have prompted relative subjective workload ratings to be made. Every participant performed each task individually and concurrently with D. This could have allowed participants to give a higher subjective workload rating to the tasks with D because they were performing two tasks instead of one. More research is needed to determine if the experimental design had an effect on subjective workload ratings.

The NASA-TLX ratings for the triple tasks are complex. In the dual task conditions the ratings for the VV, AT task condition were higher than the VV, AV task condition. However, subjective workload ratings were lower for the D, VV, and AT task condition than for either the VV, AT task condition or the D, VV, AV task condition (Table 2). One might think that the flip-flop in subjective workload is related to the D task affecting the subjective workload of the auditory tasks as before, but performance measures suggest otherwise. Both the NASA-TLX subjective workload ratings and VV performance show a decrease from the VV, AT task to the D, VV, AT task. The NASA-

TLX and VV performance were higher for the D, VV, AV task condition than for the VV, AV task condition. As discussed above, it is possible that more effort was put forth to perform the D, VV, AV task, which caused more resources to be accessible. This in turn caused more subjective workload, but improved the VV task performance and possibly increased the amount of attention given to time, which caused an insignificant but shorter produced duration. An increase in available resources was not seen for the D, VV, AT task condition, which is why the subjective workload decreased (not as much effort was put forth to perform this task), and VV performance decreased.

Recommendations

Zakay and Shub (1998) found that the Cooper-Harper rating scale did not detect an increase in subjective workload when participants performed the concurrent duration production task with nontemporal tasks. Therefore, the Cooper-Harper rating scale should be used in conjunction with the NASA-TLX so that a comparison of the workload ratings to duration productions and auditory tasks can be made. This would determine if the NASA-TLX subjective workload scale detects additional workload when participants perform the concurrent duration production task with an auditory task. This could consist of two separate studies. One would test if either of the subjective workload measures correlated more with concurrent duration productions. Another would test to see if different auditory tasks performed with concurrent duration production caused subjective workload to increase from a baseline. This study would also be able to test for increases in auditory tonal task performance when performed with the concurrent duration production task.

The auditory tonal task should be changed to a visual spatial task, where the words presented are substituted for pictures representing the words. The participant would perform the visual spatial task in combination with the auditory verbal task, the concurrent duration production task, and the concurrent duration production and auditory verbal task. The concurrent duration production task should not interfere with the visual spatial task as it did with the auditory tonal task. Again, as predicted by MRT, the visual spatial task would not interfere with auditory verbal task performance. This would allow for retesting the hypothesis in this thesis with a different set of conflicting tasks.

Limitations

Precautions should be taken when performing further research on this topic. The nontemporal tasks should be tested before final selection of tasks to make sure the D task does not interfere with nontemporal task performance if that is possible. The subjective workload with and without the D task should also be tested to ensure that the D task does not effect subjective workload. Even if MRT suggested that two nontemporal tasks should not interfere with one another, the interference caused by the nontemporal tasks should be tested before examing the effects of duration production. Even when all of these precautions are taken, there is still the risk of the D task having an effect on nontemporal task performance in triple task conditions.

Conclusion

The unpredicted interactions mentioned above, and the speculative conclusions, do not support the use of the concurrent duration production method to measure workload caused by dual nontemporal task interference. There are two conclusions that can be made given these results, both of which are not in favor of using the concurrent duration production method to measure workload caused by dual nontemporal task interference. The conclusion that the perceptual / cognitive resources are not sensitive to when two tasks use the same resources would rule out using the concurrent duration production method as a measure of workload caused by dual nontemporal task interference. The results support this conclusion because the durations produced were not significantly different for the two tasks that shared the codes of processing and the two tasks that did not. The conclusion that more perceptual / cognitive resources were made available due to the concurrent duration production, visual verbal, and auditory verbal task combination demands would also make using the concurrent duration production task unusable as a measure of workload caused by dual nontemporal task interference. The workload would be too difficult to measure because one would not know if lower durations were due to an increase in available resources or due to the task not using as many resources as another task.

As mentioned above, more research would need to be performed to determine if the interactions would cease to exist if a visual spatial task substituted the auditory tonal task and the Cooper-Harper ratings scale substituted the NASA-TLX. As mentioned above, many checks would need to be performed. Even when checks are performed there could still be an interaction between the concurrent duration production task and the nontemporal tasks in the triple task conditions. The method of measuring workload caused by dual nontemporal task interference with the concurrent duration production task would be difficult to implement if the checks would need to be performed before using it and would still not guarantee no interaction between the concurrent duration production method and the nontemporal tasks.

The current results do not support the use of the concurrent duration production method to measure the workload caused by dual nontemporal task interference. More research should be performed due to the unpredicted interactions that appeared in this research.

References

Angrilli, A., Cherubini, P., Pavese, A., & Manfredini, S. (1997) The influence of affective factors on time perception. *Perception & Psychophysics. 59*, 972–982. Retrieved from http://cogprints.org/803/3/**time**perc.pdf

Block, R. A., & Zakay, D. (2006). Prospective remembering involves time estimation and memory processes. In J. Glicksohn, & M. S. Myslobodsky (Eds.), *Timing the future: The case for a time-based prospective memory* (pp. 25–49). London: World Scientific. doi:10.1142/9789812707123_0002

Brown, S. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & psychophysics*, *59(7)*, 1118-1140. Retrieved from http://mc.psychonomic-journals.org/content/34/7/1464.full.pdf

Brown, S. (2006) Timing and executive function: Bidirectional interference between concurrent temporal production and randomization tasks. *Memory & Cognition*, *34*(7), 1464-1471. Retrieved from http://mc.psychonomic-journals.org/content/34/7/1464.full.pdf

Casali, J. G., & Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *25*(6), 623-641. Retrieved from http://www.ingentaconnect.com/content/hfes/hf/1983/00000025/00000006/art00002?token=003c13955096720297d7634257b517b5f5f407b4235476d3f4e4b2547668

Church, R. (1984). Properties of the internal clock. *Annals of the New York Academy of Sciences*, *423*, 566-582. doi:10.1111/j.1749-6632.1984.tb23459.x As cited in Lejeune, H. (1998).

Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioural Processes*, *44*(2), 127-145. doi:10.1016/S0376-6357(98)00045-X

Cooper, G. E. & Harper, R. P. Jr. (1969). The use of a pilot rating in the evaluation of aircraft handling qualities, NASA report TN-D-5143, NASA-Ames Research Center, Moffett Field, CA. Retrieved from http://www.google.com/url?sa=t&source=web&cd=3&ved=0CCIQFjAC&url=http%3A%2F%2Fwww.dtic.mil%2Fcgi-bin%2FGetTRDoc%3FAD%3DAD689722%26Location%3DU2%26doc%3DGetTRDoc.pdf&rct=j&q=The%20use%20of%20a%20pilot%20rating%20in%20the%20evaluation%20of%20aircraft%20handling%20qualities&ei=QFKBTL7IBML38AaixL2xAg&usg=AFQjCNE0SJ-kqgl66bYUhAopTdZYsH3SnA&sig2=YEXwkq4UyjyfdMD0eLXUbg&cad=rja

Gibbon, J. & R. M. Church (1984). Sources of variance in information processing theories of timing. In H. L. Roitblat, T. G. Bever & H. S. Terrace (Eds.), *Animal Cognition* (pp. 465-488). Hillsdale, N J: Erlbaum.

Hart, S. G. (1975a). Time estimation as a secondary task to measure workload. In *Proceedings of the 11th Annual Conference on Manual Control*, NASA-AMES. Moffett Field, CA.

Hart, S. G. (1975b). The influence of distraction on the perception of time. Paper presented at the 55th Annual Meeting of the Western Psychological Association, Sacramento, CA.

Hart, S. G., (1976). A cognitive model of time perception. In R. Wilsoncroft (Chair.), *Time perception*. Symposium conducted at the 56th annual meeting of the Western Psychological Association, Los Angeles, CA.

Hart, S. G. (1978). Subjective time estimation as an index of workload. *Proceedings of the Symposium on Man-System Interface: Advances in Workload Study* (pp.115-131).

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. Santa Monica, CA: Human Factors & Ergonomics Society. Retrieved from http://www.ingentaconnect.com/content/hfes/hfproc/2006/00000050/00000009/art00009?token=00481d47fd171d4862c9f7e2a46762c6b675d58763c447b3a76457b673f582f4711ad576

Hart, S.G., & Bird, K. L. (1980). Effect of counting and tracking on verbal and production methods of time estimation. *Proceedings of the 16th Annual Conference on Manual Control*, 189-210.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Advances in Psychology: Vol. 52. Human Mental Workload*, (pp. 139–183). doi:10.1016/S0166-4115(08)62386-9

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*(3), 214-55. doi:10.1037/0033-295X.86.3.214

Raby, Mireille and Wickens, C. D. (1994). Strategic Workload Management and Decision Biases in Aviation, *International Journal of Aviation Psychology*, *4*(3), 211 – 240. doi:10.1207/s15327108ijap0403_2

Rubio, S., Diaz, E., Martin, J., Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology: An International Review*, *53*(1), 61-86. doi:10.1111/j.1464-0597.2004.00161.x

Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports*, *20*, 663-678. Retrieved from www.garfield.library.upenn.edu/.../A1984TV51100001.pdf

Thayer, R. E. (1970). Activation states as assessed by verbal report and four psychophysiological variables. *Psychophysiology*, *7*, 86-94. doi:10.1111/j.1469-8986.1970.tb02278.x

Wearden, J.H., Pilkington, R., & Carter, E. (1999). 'Subjective lengthening' during repeated testing of a simple temporal discrimination. *Behavioural Processes*, *46*, 25-38. doi:10.1016/S0376-6357(98)00059-X

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII*, 239-257. Retrieved from http://books.google.com/books?hl=en&lr=&id=oPGHfhVX6lEC&oi=fnd&pg=PA239&dq=The+structure+of+attentional+resources&ots=LAAOx2W3aa&sig=4cxqniFsGoLBm-70AG_RHu9W82w#v=onepage&q=The%20structure%20of%20attentional%20resources&f=false

Wickens, C.D. (1981). Processing resources in attention, dual-task performance, and workload assessment. Technical report EPL-81-3/ONR-81-3.

Wickens, C. D., & Kramer, A. (1985). Engineering Psychology. *Annual Reviews in Psychology*, *36*(1), 307-348. doi:10.1146/annurev.ps.36.020185.001515

Wickens, C. D. (1991) Processing resources and attention. In. D. L. Damos (Ed.). *Multiple –task Performance*. Retrieved from http://books.google.com/books?hl=en&lr=&id=_KfaN4gKP8MC&oi=fnd&pg=PA3&dq=Processing+Resources+and+Attention&ots=SZF1Kk_9lv&sig=hiYC2WzQqIAQrEmNhOSNnqDqbU4#v=onepage&q=Processing%20Resources%20and%20Attention&f=false

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159-177. doi:10.1080/14639220210123806

Wickens, C. D. (2007). Attention to attention and its applications: A concluding view. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.). *Attention: From theory to practice*. Oxford University Press, USA.

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

Wierwille, W. W., & Eggemeier, T. F. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(2), 263-281. Retrieved from http://www.ingentaconnect.com/content/hfes/hf/1993/00000035/00000002/art00005?token=0044167acf44ed39412f415d7676253e7b2a4a2f425a425e2530332948fec36b97e3

Zakay, D. (1998). Attention Allocation Policy Influences Prospective Timing. *Psychonomic Bulletin & Review*, *5*(1), 114-118.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science*, *6*(1), 12-16. doi:10.1111/1467-8721.ep11512604

Zakay, D., & Shub, J. (1998). Concurrent duration production as a workload measure. *Ergonomics*, *41*(8), 1115-1128. doi:10.1080/001401398186423

Appendix A

List of all non-target and target words with durations

Non-target Words	Milliseconds	Target Words	Milliseconds
Accurate	620	Alligator	800
Alaska	650	Antelope	670
Also	630	Ape	550
Am	580	Baboon	670
Another	550	Bat	570
Arizona	660	Bear	700
Away	550	Beaver	640
Back	540	Beetle	440
Ball	490	Bird	630
Because	810	Butterfly	690
Best	600	Camel	520
Better	440	Cat	710
Beyond	700	Caterpillar	700
Big	490	Cheetah	450
Black	550	Chicken	500
Blender	590	Chipmunk	930
Book	600	Cougar	760
Both	550	Cow	840
Bottle	440	Coyote	840
Bowl	540	Crab	650
Box	740	Crocodile	1000
Bread	510	Deer	700
Bring	630	Dog	580
Calendar	660	Dolphin	600
California	770	Donkey	700
Call	510	Dragonfly	850
Came	630	Duck	490
Car	540	Eagle	460
Case	630	Elephant	890
Chair	530	Fish	740
Chocolate	660	Flamingo	700
Cinnamon	800	Fox	670
College	670	Frog	700
Color	490	Giraffe	850
Computer	610	Goat	810
Couch	580	Goose	610
Could	490	Gorilla	810
Cup	350	Grasshopper	760
Degree	590	Hamster	850
Delaware	630	Horse	850
Desk	630	Jaguar	1050
Distance	800	Kangaroo	740
Door	550	Lamb	630
Each	580	Leopard	560

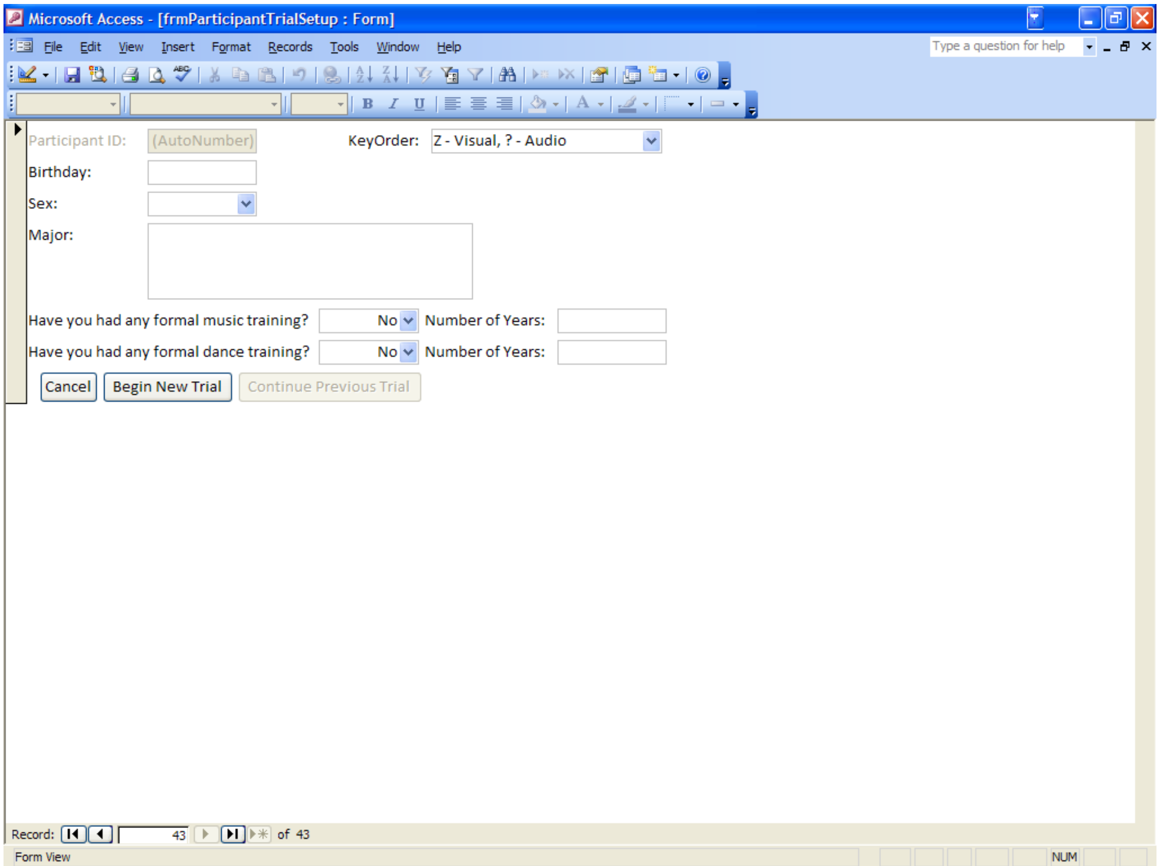
Non-target Words	Milliseconds	Target Words	Milliseconds
Ear	440	Lion	650
Effort	660	Lizard	510
End	600	Lobster	630
Explanation	890	Monkey	740
Far	550	Moose	580
Find	610	Mouse	720
First	690	Octopus	940
Five	600	Owl	540
Floor	700	Panda	540
Florida	560	Parrot	460
Fork	550	Penguin	640
Found	600	Pig	670
Four	740	Rabbit	630
Gas	670	Raccoon	710
Got	540	Rat	600
Governor	490	Rooster	630
Hair	650	Shark	670
Hamburger	720	Sheep	600
Hand	670	Skunk	650
Hat	600	Snake	650
Herbs	670	Spider	800
High	630	Squirrel	650
Home	490	Starfish	1000
House	650	Swan	710
Identical	650	Tiger	650
Into	550	Toad	490
Jacket	520	Turkey	510
Jeans	700	Turtle	510
Jump	640	Whale	540
Kind	580	Wolf	510
Knife	600	Zebra	600
Last	670	Bison	550
Leave	540	Cobra	700
Left	720	Elk	690
Let	700	Falcon	670
Live	550	Ferret	600
Look	600	Hawk	700
Made	630	Ostrich	760
May	550		
Medium	620		
Microwave	790		
Mississippi	810		
More	630		
Morning	710		
Most	740		
Music	630		
Must	670		

Non-target Words	Milliseconds	Target Words	Milliseconds
Nail	570		
Name	670		
Near	600		
Never	580		
Next	670		
Night	700		
Ohio	650		
Only	580		
Open	550		
Opportunity	840		
Orlando	740		
Oven	470		
Over	740		
Own	630		
Page	570		
Paper	560		
Park	640		
Phone	740		
Pittsford	600		
Pizza	540		
Plate	560		
Play	650		
Please	670		
Plug	500		
Present	660		
President	610		
Pretty	550		
Previous	750		
Printer	550		
Rain	640		
Ran	720		
Read	630		
Red	550		
Remain	690		
Right	680		
Ring	550		
Rochester	750		
Run	550		
Saw	580		
Say	600		
School	740		
Seem	720		
Shall	650		
Shirt	760		
Shoes	650		
Should	690		
Show	560		

Non-target Words	Milliseconds	Target Words	Milliseconds
Sing	620		
Sky	750		
Sleet	700		
Snow	740		
Sock	540		
Songs	770		
Soon	690		
Spice	870		
Spoon	740		
Stand	850		
Stereo	760		
Stove	760		
Such	710		
Sugar	630		
Sun	650		
Sure	570		
Table	540		
Television	690		
Tell	580		
Than	740		
These	840		
Think	630		
Too	460		
Trash	740		
Tree	540		
Under	550		
Until	670		
Upon	740		
Use	670		
Want	630		
Way	650		
Where	580		
While	580		
White	600		
Why	630		
Window	650		
Wish	690		
Year	580		
Tiny	640		
Large	750		
Average Duration	**632**		**674**

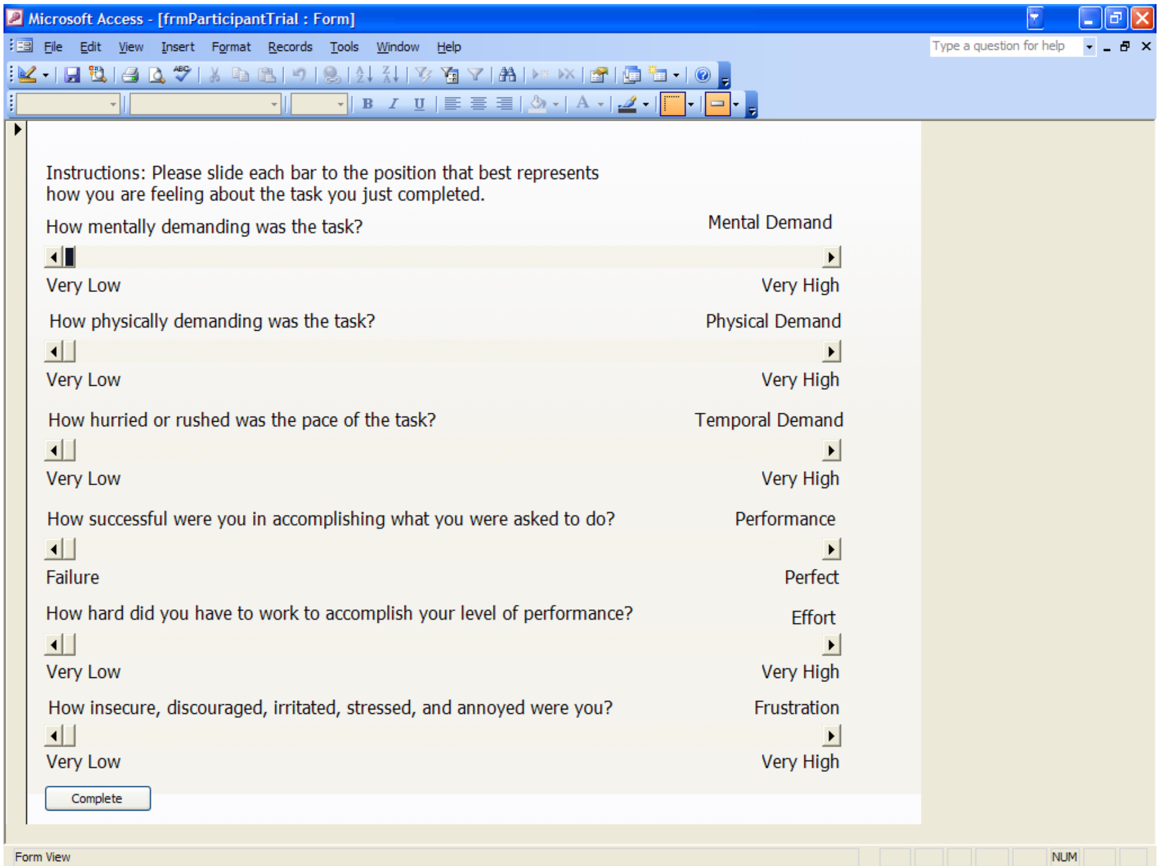
Appendix B

Screen shot of Pre-research questionnaire



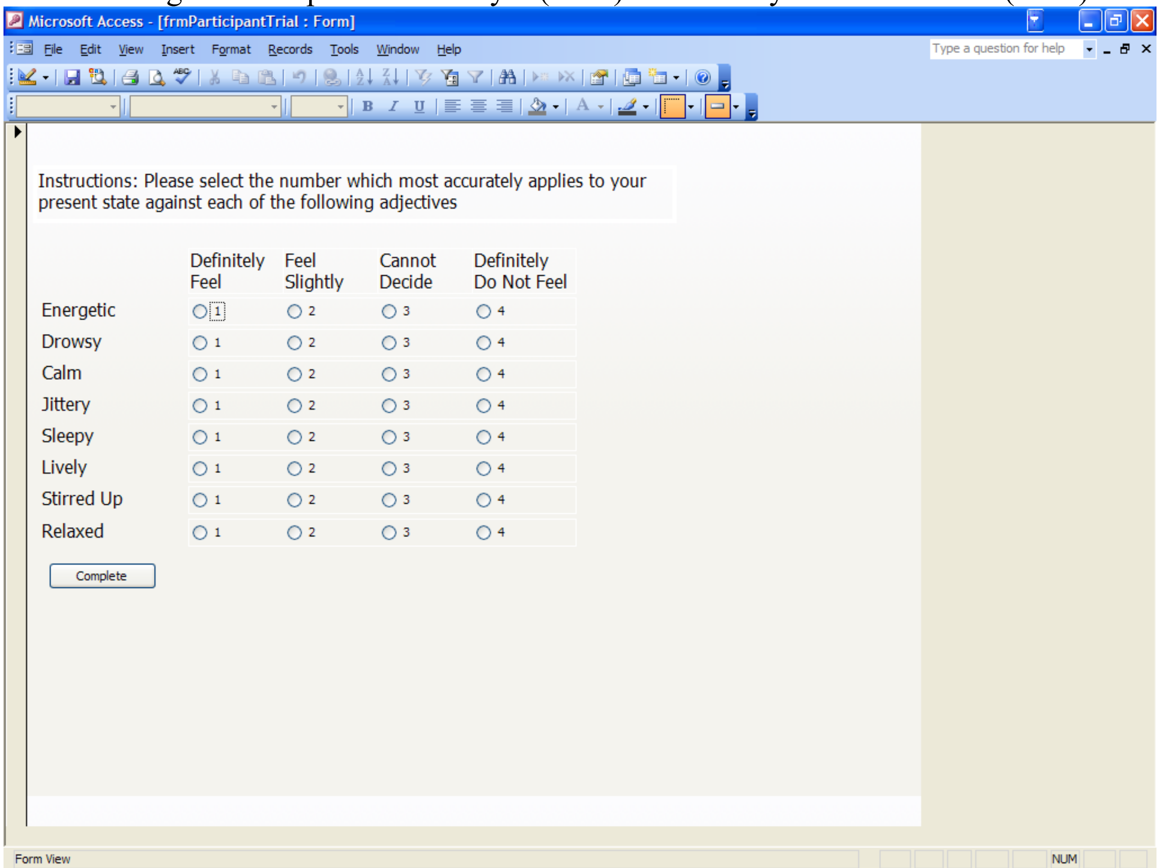
Appendix C

NASA-TLX



Appendix D

Arousal rating scale adapted from Thayer (1967) and used by Wearden et. Al. (1999).



Instructions: Please select the number which most accurately applies to your present state against each of the following adjectives

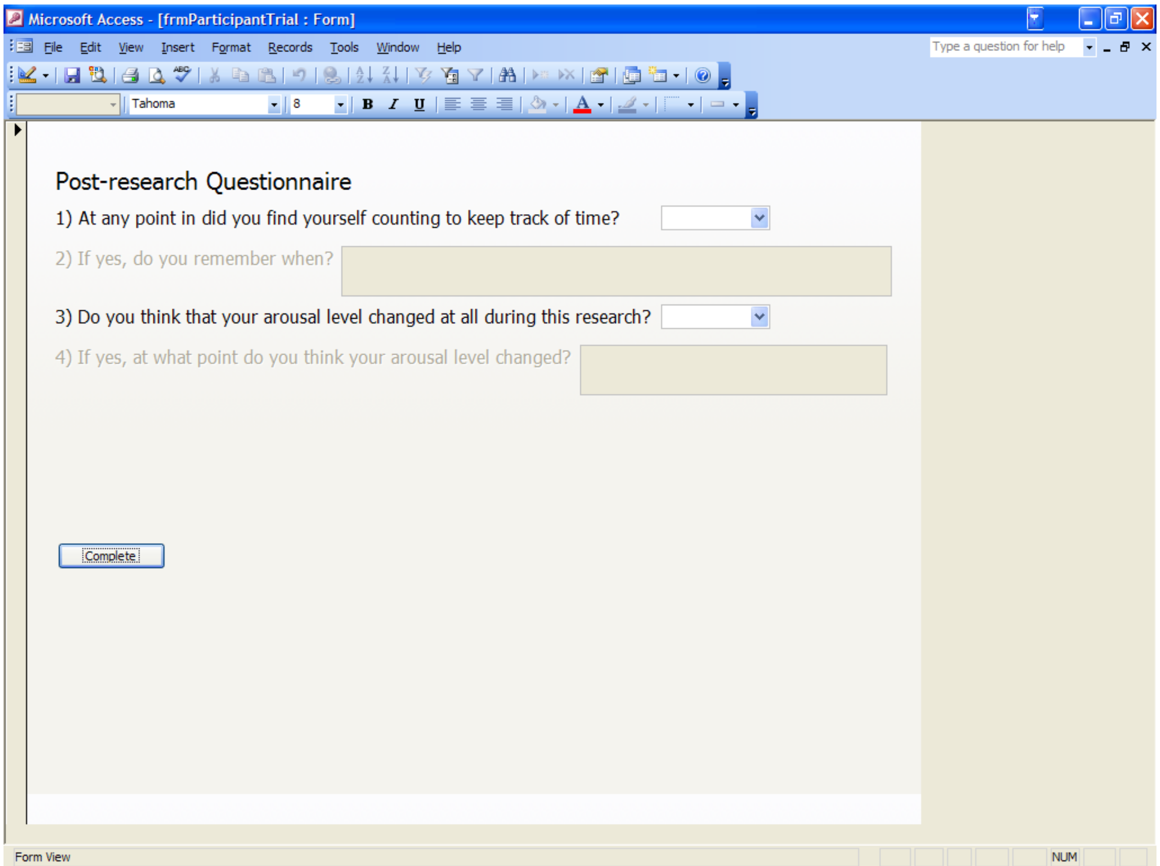
	Definitely Feel	Feel Slightly	Cannot Decide	Definitely Do Not Feel
Energetic	1	2	3	4
Drowsy	1	2	3	4
Calm	1	2	3	4
Jittery	1	2	3	4
Sleepy	1	2	3	4
Lively	1	2	3	4
Stirred Up	1	2	3	4
Relaxed	1	2	3	4

Complete

Form View NUM

Appendix E

Post-research Questionnaire



Appendix F

INFORMED CONSENT FORM FOR BEHAVIORAL RESEARCH STUDY
Rochester Institute of Technology

Title of Project: **Perceived Duration Production as a Measure of Dual Task Attentional Resource Demand**

Investigators in Charge:

Mrs. Mary Merlau
MS Candidate
Dept. of Psychology.
Rochester Inst. of Technology
Tel. (585) 613-5230
Email: mxm8079@rit.edu

Dr. Andrew Herbert
Associate Professor
Dept. of Psychology
Rochester Inst. of Technology
01-3158 Eastman Bldg.
Tel. (585) 475-4554
Email: amhgss@rit.edu

Explanation of the Project.

1. You are being asked to participate in a research study that is looking at the attentional resource demands of dual task performance. The results of this study will be applicable to many fields where mental workload could cause errors, stress, and fatigue. In the future a mental workload scale and conflict matrix could be developed that would enable the prediction of mental workload in dual and single task situations. This type of scale and conflict matrix would produce the greatest benefits for the military, medical field, aviation, and any other field where mental workload could cause increased errors, stress and fatigue.

2. The goal of this work is to evaluate humans' dual task attentional resource demand.

3. This study requires you to engage in up to three simultaneous tasks. You will be responsible for (1) a visual detection task, which entails detecting words that are animal names, (2) an auditory detection task which will require you to detect the names of animals, (3) an auditory detection task, which will require you to detect when a tone changes, and (4) a time production task, which will require you to produce a requested duration of time. Please do not count while producing the requested times. You will also be asked to fill out a mental workload assessment scale called the NASA Task Load Index after every task. You will also be asked to fill out a questionnaire about your arousal levels during the study. At the end of the study you will be asked to fill out a small questionnaire. Please be as accurate as possible in your answers.

4. The only risks to you from participating in the experiment are the slight mental workload and fatigue associated with any detection task.

5. Results of this research will be used to further enhance our understanding of the role of time perception in human mental workload.

Your rights as a research participant

1. We will be happy to answer any questions you have about the study at any time. Mrs. Merlau and Prof. Herbert may be contacted at the telephone numbers and e-mail addresses shown

above. If you have questions about your rights as a research subject, you can call collect the Rochester Institute of Technology Institutional Review Board at (585) 475-7673, or e-mail hmfsrs@rit.edu.

2. No subsequently published results will contain any information that could be associated with individual participants. No information identifying individual subjects will be ever associated with the data collected. All data will be stored and secured only on the investigator's computer after being retrieved from the program.

3. Your participation is wholly voluntary. Your decision to participate, or to not participate, or to withdraw from the study during the experiment will in no way influence your relationship with the researcher or your professor(s).

4. You may refuse to participate or may discontinue participation at any time during the project without penalty or loss of benefits to which you are otherwise entitled.

5. Results of the proposed research will be used to further guide our understanding of temporal awareness.

6. The results of this research will be submitted to peer-reviewed journal articles and perhaps presented at a human factors-related conference. No information allowing for identification of individual participants will be included in these reports.

Statement of consent

Participant:

I agree to participate in this study, which seeks to guide development and testing of the measurement of human mental workload. I understand the information given to me, and I have received answers to any questions I may have had about the research procedure. I understand and agree to the conditions of this study as described on this form.

I understand that I am volunteering to participate in this study, that I will be not be compensated for participating apart from the chances of winning a raffle, and that I may withdraw from this study at any time without penalty to me.

I certify that I am at least 18 years old.

I understand that I will be given a signed copy of this consent form.

____________________________________ _______________________

Signature Date

Researcher:

I certify that the informed consent procedure has been followed, and that I have answered any questions from the participant above as fully as possible.

____________________________________ _______________________

Signature Date

Appendix G

Instructions:

This study requires you to engage in up to three simultaneous tasks. You will be responsible for

(1) a visual detection task, which entails detecting words that are animal names. Animal names and other words will be presented on the screen. You must press the Z key as soon as you see an animal name. You must press the button as fast as possible as it will only register as a correct response while the animal name is still up on the screen.

(2) an auditory verbal detection task which will require you to detect the names of animals. Animal names and other words, which have been pre-recorded, will be spoken. You must press the ? key as soon as you hear an animal name. You must press the button as fast as possible as it will only register a correct response while the animal name is being spoken.

(3) an auditory tonal detection task, which will require you to detect when a tone changes. Tones, which have been pre-recorded, will be played through the computer speakers. You must press the ? key as soon as you hear the tone change. (You will hear a blip when the tones are changing from one to another. This sound does not mean that the tone will change, it is just the sound files changing.)

(4) a time production task, which will require you to produce a requested duration of time. You will be asked to produce a duration of 15 seconds. You will press the space bar to start the duration. When you believe that 15 seconds has passed you will press the space bar to stop the duration. Please do not count while producing the requested times.

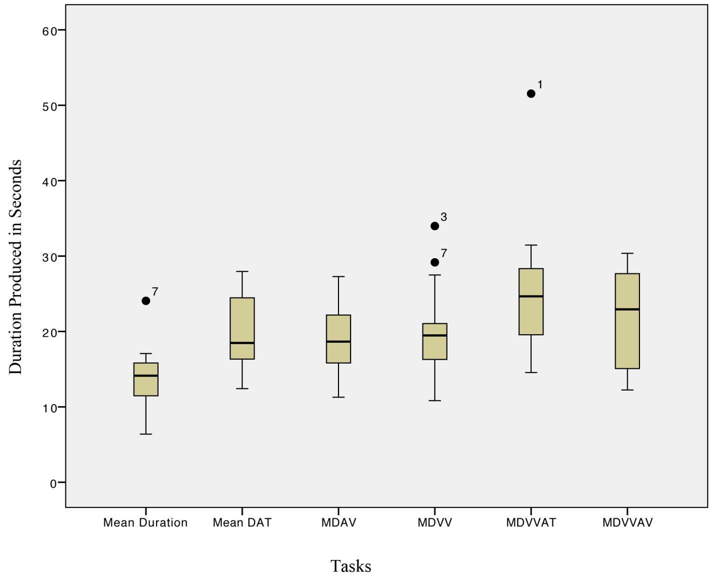
During the detection tasks you must press the appropriate button while the word is on the screen or being spoken in order for it to register as a correct response. Once the word is off the screen or done being spoken it will register as an incorrect response.

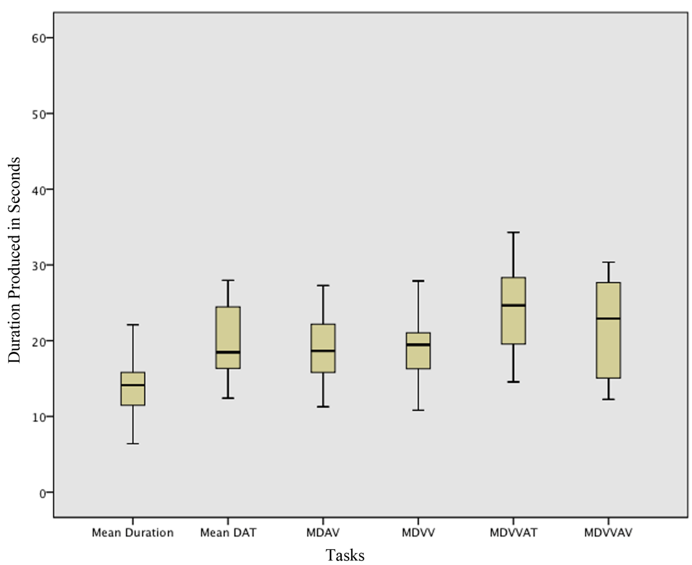
You will also be asked to fill out a mental workload assessment scale called the NASA Task Load Index once after every task. You will also be asked to fill out a questionnaire about your arousal levels at different times during the study. At the end of the study you will be asked to fill out a small questionnaire.

When done with practice trials please stop and take a moment to let me remind you of all the buttons and instructions and answer any questions you may have.

Appendix H

Duration performance box plots with and without outliers

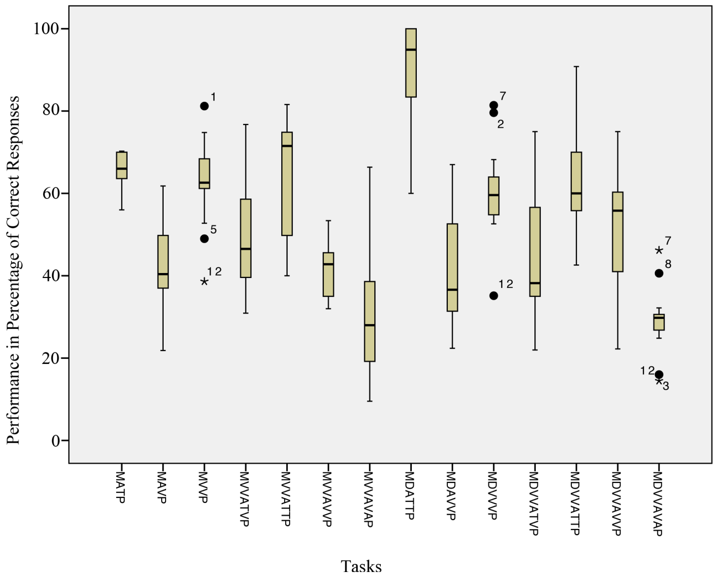


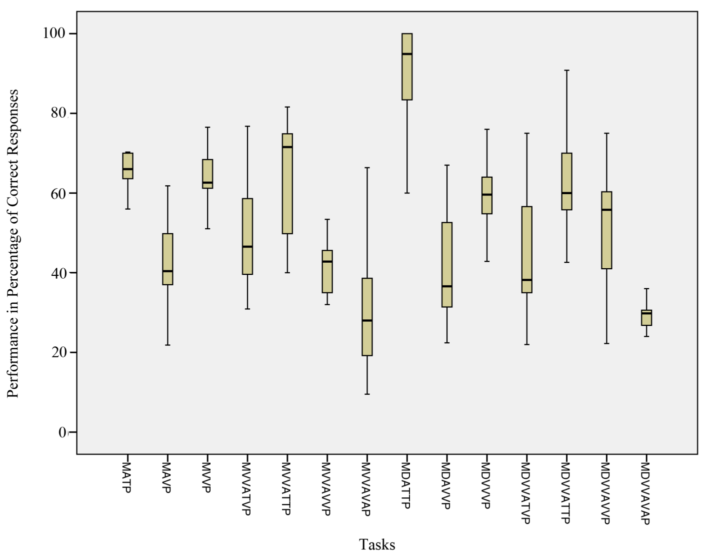


Average durations for each participant that showed to be an outlier in the first figure led to the removal of data points more than 1.5 box lengths away from the top or bottom of each box. Eight trials were removed from three participants to produce the lower box plot.

Appendix I

Nontemporal performance box plots with and without outliers.

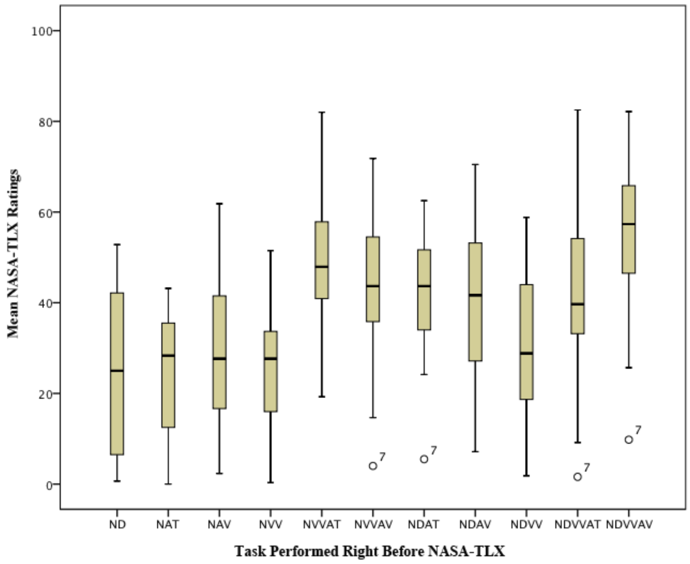


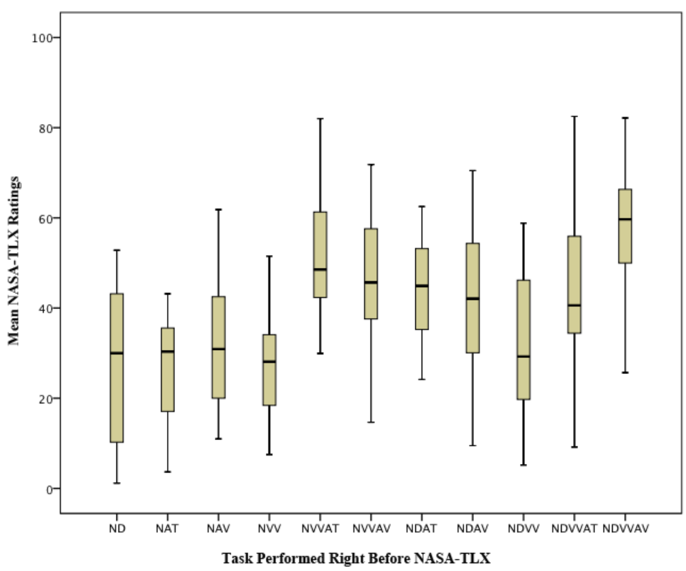


Average performance for each participant that showed to be an outlier in the first figure led to the removal of data points more than 1.5 box lengths away from the top or bottom of each box. Nineteen trials were removed from seven participants to produce the lower box plot.

Appendix J

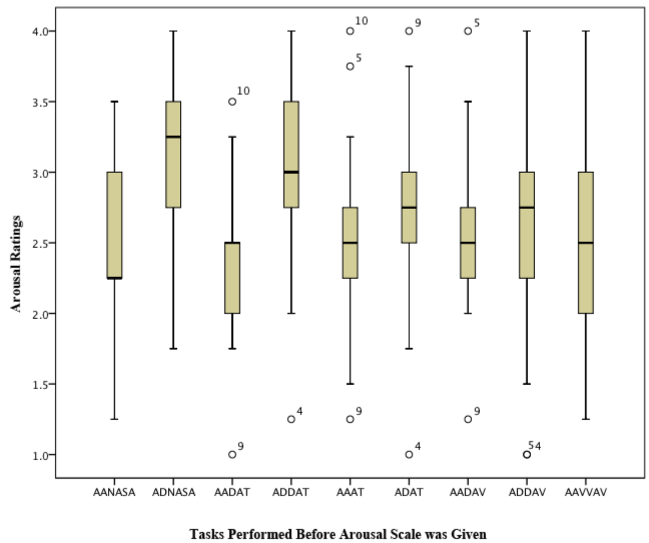
NASA-TLX box plots with and without outliers.





Appendix K

Arousal Scale Rating Box plot with Outliers, AA=Arousal Activation, AD=Arousal Deactivation



Appendix L

Cronbach's Alpha for conditions with word lists.

Cronbach's Alpha for conditions with word lists							
	VV	AT	AV	VVAT		VVAV	
				VV	AT	VV	AV
Without D	0.35	-0.089	0.73	0.76	0.73	0.4	0.81
With D	0.65	0.53	0.82	0.8	0.69	0.82	0.48

Appendix M

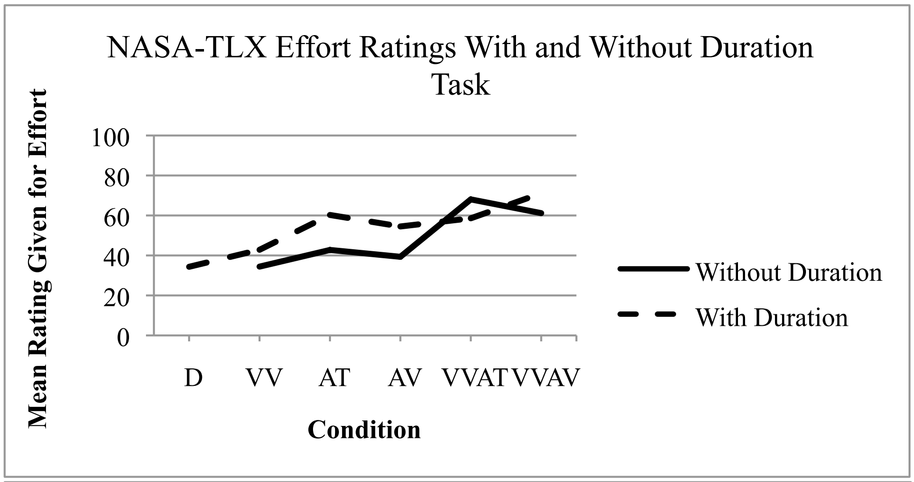
Cronbach's Alpha for each condition's NASA-TLX **and arousal** ratings

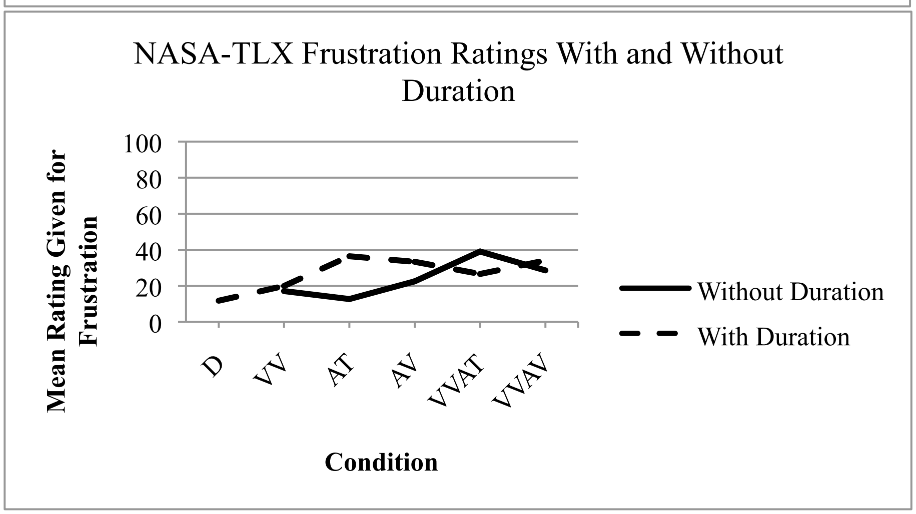
Cronbach's Alpha for NASA-TLX						
	D	VV	AT	AV	VVAT	VVAV
Without D		0.59	0.66	0.74	0.8	0.84
With D	0.76	0.86	0.7	0.85	0.85	0.74

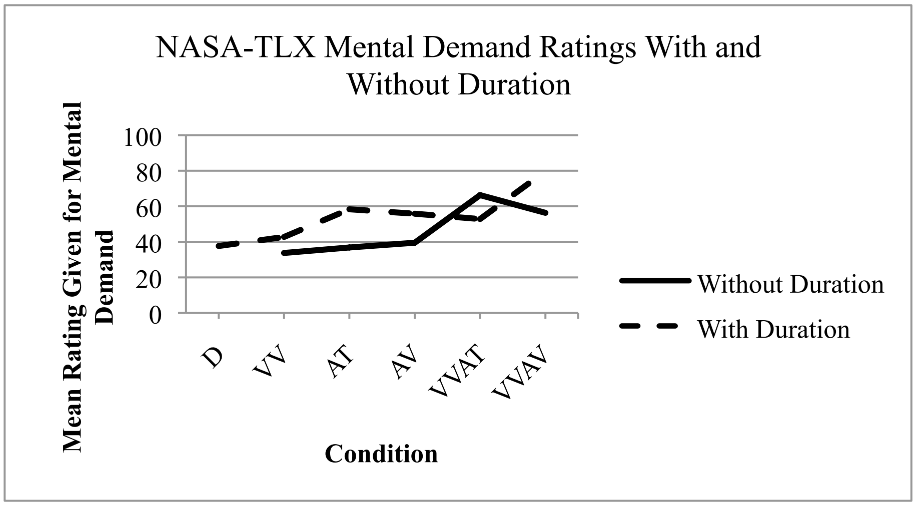
Cronbach's Alpha for Arousal	
Activation	0.66
Deactivation	0.68

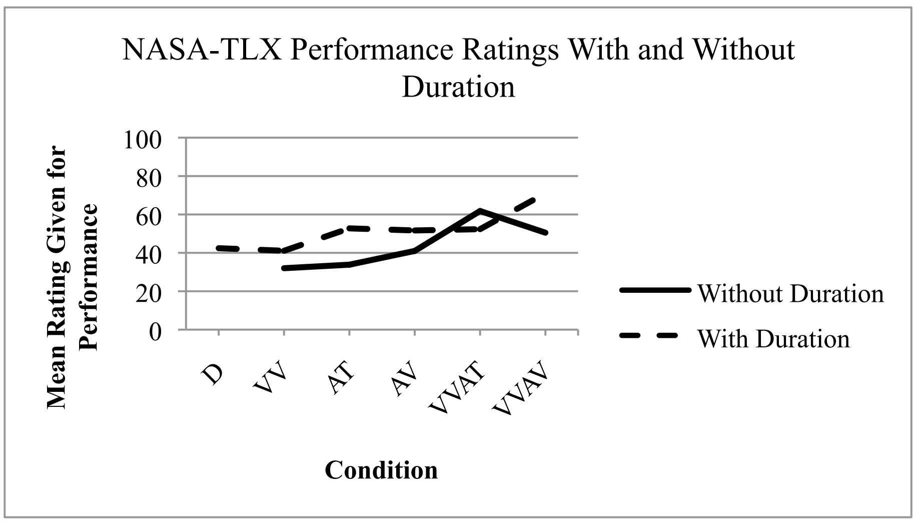
Appendix N

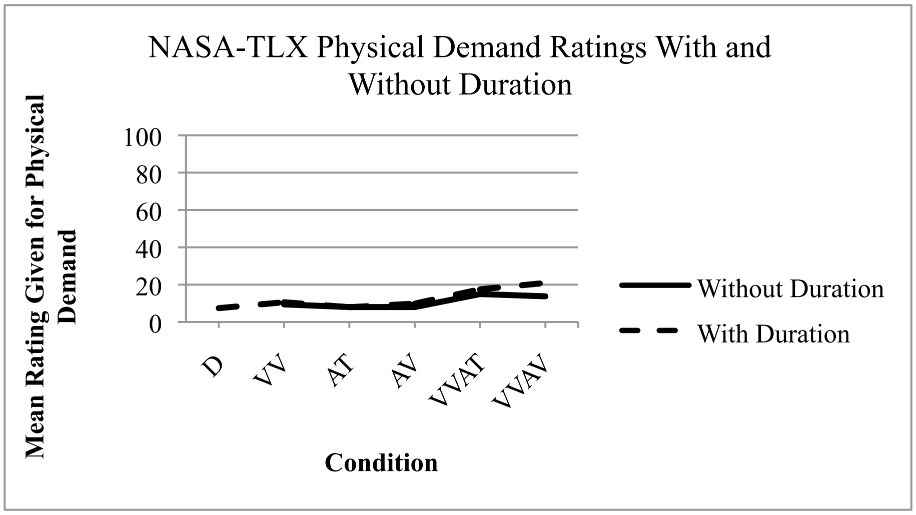
NASA-TLX individual workload scale ratings for each condition

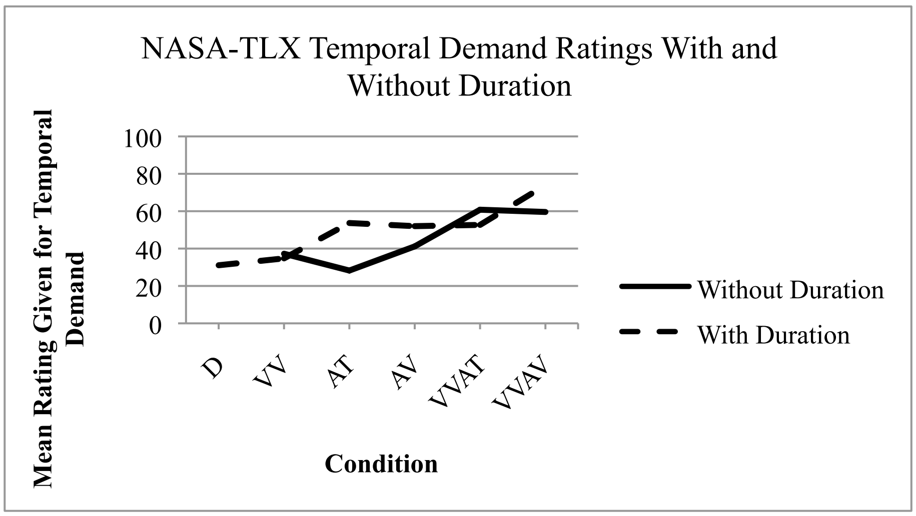





NASA-TLX Mental Demand Ratings With and Without Duration
Mean Rating Given for Mental Demand
100
80
60
40
20
0
D
VV
AT
AV
VVAT
VVAV
Condition
Without Duration
With Duration


NASA-TLX Performance Ratings With and Without Duration
Mean Rating Given for Performance
100
80
60
40
20
0
D
VV
AT
AV
VVAT
VVAV
Condition
Without Duration
With Duration


NASA-TLX Physical Demand Ratings With and Without Duration
Mean Rating Given for Physical Demand
100
80
60
40
20
0
D
VV
AT
AV
VVAT
VVAV
Condition
Without Duration
With Duration


NASA-TLX Temporal Demand Ratings With and Without Duration
Mean Rating Given for Temporal Demand
100
80
60
40
20
0
D
VV
AT
AV
VVAT
VVAV
Condition
Without Duration
With Duration