ABSTRACT

BRYANT, MICHELLE RENEE. Determining Challenge from Stress and Workload in a Complex Task. (Under the direction of Dr. Anne Collins McLaughlin)

The experiences of workload and stress are often considered linked to the difficulty in a task and performance on that task- as difficulty increases, workload and stress tend to increase and performance tends to decline at high difficulty levels. These effects are often assumed to be stronger in individuals with less resources available for the task or personality traits that can alter the subjective experience of stress and workload. An experiment was performed to assess the effects of personality, age group, and task difficulty on performance of a complex task. Reported stress during performance and reported mental workload during performance were also collected. Analyses were performed using multi-level modeling and found effects of age and task difficulty. It is concluded that the subjective experiences of stress and workload may differ from each other at different levels of difficulty.

Determining Challenge from Stress and Workload in a Complex Task
North Carolina State University

by
Michelle Renee Bryant

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Arts

Psychology

Raleigh, North Carolina

2013

APPROVED BY:

Dr. Sharolyn Lane

Dr. Shevaun Neupert

Dr. Anne Collins McLaughlin
Chair of Advisory Committee

DEDICATION

This Thesis is dedicated to my mother, my dad, my brother, my husband, his mother, and his brother without whom I would not have been able to pursue my dream.

BIOGRAPHY

Michelle attained her Bachelor of Arts in Psychology from North Carolina State University in 2005. She was a high school social studies teacher for three years before enrolling in graduate school to pursue her Ph.D. Michelle continues to teach and enjoys interacting and mentoring students.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Performance has been at the forefront of human factors research since the inception of the field. The study of subjective factors thought to impact performance has increased in the last 30 years. Subjective workload and stress have been identified as variables that could provide a better understanding of performance. For example, excessive workload negatively affects performance. However, lowering workload too much can result in a vigilance task. This would paradoxically also negatively impact performance by increasing difficulty (Warm, Parasuraman, & Matthews, 2008). Similarly, stress, especially in sustained-attention tasks, has been shown to negatively impact performance (Hancock & Warm, 1989; Lazarus, 1999). In order to verify that considering only stress or workload alone does not fully explain performance, both were measured. This may also lead to an understanding of performance as a function of both subjective workload and stress.

By investigating how workload and stress impact performance it may be possible to predict when performance will decline. Prediction of the point of decline in performance is important for two reasons. First, if performance declines are predictable by measuring workload and stress, tasks can be adjusted for difficulty according to these variables so that performance decline is avoided. Second, it may be possible to find a balance between workload and stress where high performance is maintained despite the fact that task difficulty or workload and stress remain relatively high. The results of this research would allow for a deeper understanding of the performance of individuals who experience high levels of workload and stress and solutions for maintaining high performance.

Individual differences in performance, however, are likely to be factors of consideration. Cognitive, perceptual, and motor differences due to age group or personality should be considered when examining workload and stress. Investigating workload and stress through the lens of individual differences allows for a better understanding of these performance variables and provide for greater external validity. That is, by considering how individual differences affect subjective workload and stress, we may better understand these variables in the greater population.

Challenge

The experience of challenge is a relatively new area in the study of performance. Identifying a point in a task where a participant experiences challenge may be related to subjective feelings of workload and stress. According to Guadignoli and Lee (2004) challenge is a function of two interacting variables: potential available information (PAI), and functional task difficulty. PAI is the information that is gained by performing a task. Functional task difficulty is the level of subjective difficulty experienced by the performer. Subjective experiences that occur during the task could impact the level of PAI as well as the subjective difficulty experienced by the learner. Guadignoli et al. (2004), argue that optimal levels of both PAI and functional task difficulty result in an optimal level of challenge for maintaining performance on a task. However, one possible problem with this definition of challenge lies in the measurement of PAI. Though experts perform in a manner that suggests they better utilize PAI than novices (see Ericsson & Lehmann, 1996 for a review), it is unclear how to quantify knowledge gained from PAI.

Previous researchers have shown that learning is not measurable by performance change alone (e.g. Newell, 1991; Bjork & Bjork, 2011). Bjork et al. (2011) point out that it is difficult to pinpoint when knowledge has been gained because it is not a directly observable event. However, they agree with Guadignoli and Lee (2004) that challenge (or desirable difficulties) does enhance learning (e.g. distributed vs. massed practice). Since measuring knowledge seems to be so important to the challenge point framework proposed by Guadignoli et al. (2004), it is imperative to find representative, alternative variables by which to measure challenge.

It is proposed that the use of three well-established constructs be used to predict when the experience of challenge occurs; task difficulty, subjective workload, and subjective stress. Tulga and Sheridan (1980) looked at a mathematical equation for predicting subjective workload in participants in relation to performance outcomes. Wickens (1999) proposed that individuals will "off-load" when task difficulty is too high to maintain performance in order to reduce workload. Further, Hancock and Szalma (2008) examined stress as an influencing variable that consumes resources interfering with positive performance outcomes. To this end, by examining each of these three variables results will further challenge research as it relates to performance maintenance.

Task Difficulty

Observing performance when difficulty increases does not provide a complete picture of challenge since performance may not change appreciably even when the internal state of the human changes greatly. For example, people can maintain similar levels of performance even when difficulty is greatly increased, but this does not mean they are having the same

experience as at lower difficulty. To understand the factors affecting performance we must examine the effects of subjective measures such as subjective workload and stress to see where a task might become too difficult for any individual.

Workload

Predictive models of the impact of subjective workload on performance are prevalent in the literature. Overload and under-load (or vigilance) are two such examples. Overload occurs when subjective workload exceeds comfortable levels for the participant. This excess in workload is evidenced by performance declines (e.g., Hockey, 1997). Likewise, in vigilance experiments where targets are minimal, under-load resource strain is also evidenced by performance declines (Szalma, Hancock, Dember, & Warm, 2006). These findings are important to consider at either end of the task demand spectrum. However, little of the research examining workload has focused on the limits of stable performance.

Understanding performance under very high workload is crucial to understanding limits of stable performance. There are many examples of the need for examining performance maintenance despite high subjective workload. In a study performed by Sperandio (1978), when air traffic controllers' airplane load exceeded a particular "comfort zone," they changed strategy from individually routing airplanes to a set routine for all. This reduced the strain on their working memory and is argued to be evidence for the pursuit of reduced subjective workload. Wickens (1999) supported this finding when he observed that individuals will "off –load" when task difficulty increases beyond the ability to perform. Tulga and Sheridan (1980) provided a mathematical approach to predicting performance by considering task load and subjective workload. As task load increased, so did subjective

mental workload. However, once subjects reached a point where many simultaneous tasks were required, they pursued strategies where results were immediate rather than delayed, reducing their immediate subjective workload. These observations support the idea that as task load increases, there is a point where it becomes impossible to plan ahead and obviating this need reduces subjective workload. Further, this adds evidence to the claim of a link between the experience of workload and subsequent performance on a task.

Challenge as a Function of Workload

Just as the association of performance and subjective workload is not simple, neither is the relationship between subjective workload and challenge. Research has not yet examined the connection between challenge and subjective workload reports. Wickens (2002) claimed that by measuring subjective workload we can predict when task demands compete for similar resources. Hancock and Warm (1989) found that just before performance declined individuals reported an increase in subjective workload. Thus, subjective workload may indicate strain on resources just before a task becomes too difficult to maintain performance. This may be closer to an estimate of challenge than performance alone (Figure 1). Using subjective workload reports, we can likely monitor experience of challenge and potentially observe where a task has become too challenging once task-shedding/off-loading begins. However, subjective workload likely does not paint a complete picture of challenge by itself – it is also important to consider stress.

Stress

Stress can be examined in a psychological context as an intervening variable where failures between perceived demands and ability to adapt occur (Sanders, 1983). Sanders

(1983) explained that this type of stress exists within the participant and occurs when there is a conflict of desired performance and outcomes. For example, a participant is asked to solve problems and must exert high effort while deciding and implementing a particular strategy. When outcomes are unsuccessful for the individual, a stress response will occur.

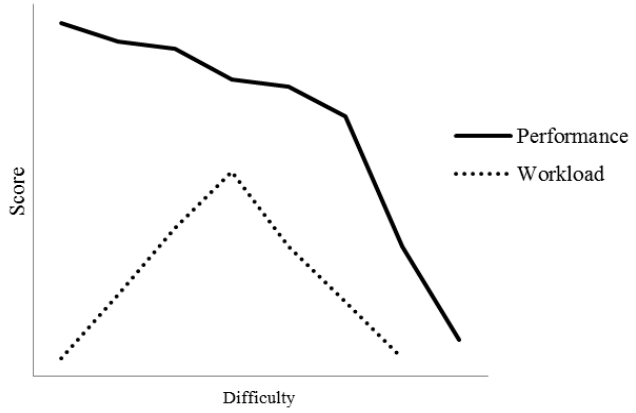


Figure 1. As task difficulty increases, so does subjective workload, shown hypothetically above. Once a task becomes too difficult, subjective workload should begin to decline due to task shedding. This is a result of the individual reverting to strategies that no longer allow planning. The challenge region of an individual lies just beyond their highest subjective workload where an attempt to compromise between performance stability and adaptability occurs.

Investigating stress alongside subjective workload would account for an integral part of performance: adaptability (Hancock & Warm, 2009; Hancock & Szalma, 2008). Similar to research conducted on subjective workload and task difficulty, determinants of stress rely on the amount of information presented to a subject as well as the meaning that the subject applies to the information (Hancock, 2009). Individuals will assign different meanings to information based upon individual differences such as personality (Lazarus, 1999). Lazarus (1999) refers to this assignment of different meanings as appraisal. Therefore, the way an individual appraises the demands, constraints, and opportunities of a task will determine the

reaction to that task. For example, when one appraises the task as threatening his performance goals, higher levels of stress tend to be reported (Lazarus, 1999). Individual differences such as personality may influence the adaptability of an individual (Hancock & Szalma, 2008) and thus the experience of stress.

Challenge as a Function of Stress

There are two competing views on when a stress response may occur with regard to performance on a task. The first arises from a time-on-task perspective where arousal is at its highest at the start of a task and declines as time on task increases (Sanders, 1983). The resulting relationship is more likely to occur when there is no change in task difficulty. The second view of stress is observed when effort fails to keep up with task demands. At this point, physiological equilibrium, a balance between a number of factors including effort, cannot be maintained (see Sanders, 1983 for a full explanation of physiological equilibrium). In this second type of stress, the stress response would be depicted in an inverted U-shape relative to task demands. Therefore, using a model of task difficulty and stress (Figure 2) it is predicted that as task difficulty increases so will subjective stress. Challenge is likely to be experienced just before the highest levels of stress are reported. This would be the case because unlike subjective workload, the individual will report a stress response beyond the experience of challenge resulting from factors such as performance degradation, resource overload and the internal conflict that occurs due to the individuals desire to perform well even while the difficulty makes it impossible to do so.

Individual Differences

The effects and interactions of stress, subjective workload, and performance can likely be used to predict challenge, however individual differences that affect each of these variables must also be considered. Factors such as age (e.g., Parker, McDonald, Rabbitt, & Sutcliffe, 2000) and personality (e.g., Lazarus, 1999; Hancock & Szalma, 2008) each have a role in predicting performance, subjective workload, and stress.

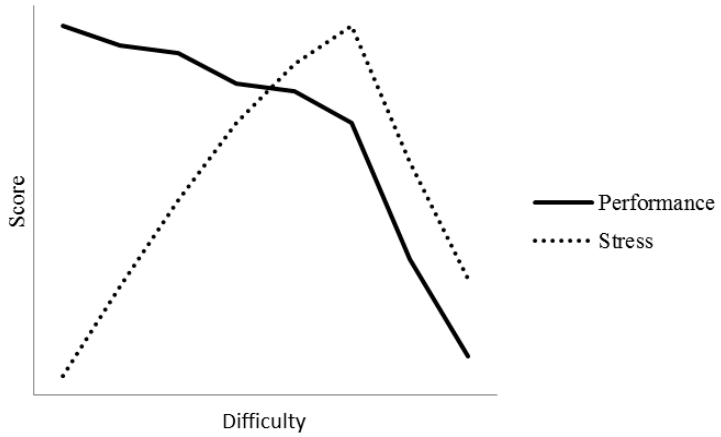


Figure 2. As task difficulty increases, so does subjective stress. Once a task becomes too difficult, stress may continue to increase until an individual can no longer conceive performance recovery. This is when stress will begin to decline.

Age. Differences that exist between younger adults and older adults in perceptual abilities, cognition, and motor cognition give cause to expect that age group is a moderator of the subjective workload- and stress-performance relationship. As adults age, perceptual, cognitive, and movement abilities tend to decline (for a review see Rogers, 1997). Knowledge of age-related decline allows us to predict that portions of a task relying heavily on any one or several of these areas will be differentially challenging to older adults.

Likewise, as task difficulty increases older adults should report their subjective workload as higher than younger adults (Boyer, Pollack, & Eggemeier, 1992; Ho, Wheatley, & Scialfa, 2005). Thus ability differences should result in experience of higher subjective workload than younger adults when task difficulty is the same.

Taking into account age decrements in cognitive processing capability (see Rogers, 1997 for a review) stress is likely to be reported differently between younger and older adults. If stress is an outcome of the knowledge of incompatibility of task demands and task performance, the attentional resources of older adults may limit the awareness of such incongruity (Sharit & Czaja, 1994) and subsequently influence the experience of challenge. In contrast to this, in laboratory experiments older adults often exert more effort (a measure of workload) to achieve similar performance results as younger adults in the same tasks (Hess & Ennis, 2011) and may be more at risk of negative appraisals as a result (Verhaeghen, Steitz, Sliwinski, & Cerella, 2003). Therefore, their report of stress may be lower due to relative differences in their perception of stress but their subjective workload while performing a task may be *higher* than younger adults due to the amount of effort required during the task.

Similarly, it is expected that documented cognitive differences in younger and older adults will play a role in performance and subsequently the experience of challenge. Czaja and Sharit (1993) found age differences in performance by task type even when controlling for individual differences in typing speed and previous computer experience. Therefore, by considering factors associated with aging a more accurate prediction of when a task is challenging for older adults may be determined.

Personality. Researchers have found personality-related differences between individuals in their experience of stress (Bolger & Zuckerman, 1995; Neupert, Almeda, Mroczek, & Spiro, 2006; Salthouse 2004; Szalma, 2009a, b). Young adults who score highly on the neuroticism trait (high-N) have been shown to report significantly higher interpersonal daily stress than those who score low on the neuroticism trait (low-N) (Gunthert, Cohen, & Armeli, 1999). Similarly, research has found that high-N individuals are more vulnerable to stress (Suls, 2001). Lastly, high-N individuals have also been found to be more likely to appraise a situation as threatening (Craske, 1999). Therefore, differences in performance across individuals who are high-N may be more susceptible to negative performance outcomes due to stress.

Study Overview

The current study was designed to examine the experience of challenge by measuring subjective workload and stress in an individual across task difficulties. A novel, attention demanding, visual-perceptual, computer task was employed. The researcher manipulated difficulty of the task by altering the contrast between target elements and the background of the display as well as the demands of a secondary task. Participants responded to subjective workload and stress measures after each trial. Due to likely individual differences in the experience of subjective workload and subjective stress for age group, younger and older adults were measured. Similarly, there should be differences in how neuroticism affects performance. Therefore, a personality inventory was used to measure neuroticism. Difficulty level, age group, and neuroticism were expected to be predictors of each dependent variable;

performance, subjective workload, and subjective stress. Three different models were derived to perform separate analyses for each DV.

Hypotheses:

DV: Performance

1. All participants will perform better on the task at low difficulties compared to high difficulties (Main effect of Difficulty).
2. Younger adults will perform better than older adults on the task in general (Main effect of Age Group).
3. Participants who score highly on the neuroticism trait will perform more poorly than those who score low (Main effect of Neuroticism).
4. Participants who score highly on the neuroticism trait will perform differentially worse at higher levels of difficulty (Interaction of Difficulty and Neuroticism).

DV: Workload

1. All participants will report higher workload at higher difficulty levels (Main effect of Difficulty).
2. Older adults will report higher levels of subjective workload than younger adults (Main effect of Age Group).
3. Participants who score highly on the neuroticism trait will report higher workload than those who score low on the neuroticism trait (Main effect of Neuroticism).
4. Controlling for age group, participants who score highly on the neuroticism trait will report higher levels of workload at higher levels of difficulty (Interaction of Difficulty and Neuroticism).

DV: Stress

1. All participants will report higher subjective stress levels at higher levels of difficulty (Main effect of Difficulty).
2. Older adults will report lower stress levels than younger adults (Main effect of Age Group).
3. Participants who score highly on the neuroticism trait will report higher stress than those who score low on the neuroticism trait (Main effect of Neuroticism).
4. Participants who score highly on the neuroticism trait will report higher levels of stress at higher levels of difficulty (Interaction of Difficulty and Neuroticism).

METHOD

Participants

Younger participants were recruited from undergraduates enrolled in an introductory Psychology course at a large, southeastern university and participants over age 65 were recruited through a database of potential participants maintained by the Learning, Aging, Cognitive, and Ergonomics Lab. A total of fifty-six young adults and twenty-nine older adults were recruited for the study. Ten younger adult participants had to be excluded in the final analysis due to a protocol change with one additional participant excluded for colorblindness. Four older adult participants had to be excluded from the final analysis due to unknown data collection error. For this analysis, forty-six young adults (M_{age} = 20.1, SD_{age} = 1.5) and twenty-six older adults (M_{age} = 75.5, SD_{age} = 5.5) were included. Participants were given course credit (young adults) for their participation or compensated monetarily (older adults) for their time.

Materials

Ability Tests. Near vision and normal color vision were both assessed. Color vision was assessed by selected electronic plates from the Ishihara color test to test for red-green color blindness (Ishihara, 1994). A computerized choice reaction time test was given to highlight age group differences in reaction time.

Participants were administered a computerized Functional Field of View test (FFOV) (Feng, Spence & Pratt, 2007) to measure the range of view outside of a fixation point. Last, participants were given the Mental Rotation Test (MRT) (Peters, Laeng, Latham, Jackson, Zaiyouna, & Richardson, 1995; Vandenberg, & Kuse, 1978) to measure spatial ability. The Ginsburg Functional Acuity Contrast Test (FACT) (Ginsburg, 1984) was administered to measure the peak contrast sensitivity and lowest levels of contrast sensitivity of participants.

Questionnaires. All participants were given a demographics survey (Appendix A) and exit interview (Appendix E). Each was given to collect information about previous experience and usage rates of computers, and general feelings about playing video games.

International Personality Item Pool-NEO (IPIP-NEO). The IPIP-NEO (Neuroticism, Extroversion, Openness, and Conscientiousness) (Goldberg, Johnson, Eber, Hogan, Ashton, Cloninger, & Gough, 2006) is comprised of 50 questions that assess an estimate of the participants standing within the five-factor personality model (See Appendix C).

Life-Shortened Stress State Questionnaire. The SSSQ (Helton, 2004) is a self-report measure created to assess stress via three subscales (e.g., task engagement, distress, and worry) in a shortened form (Appendix D). The measure is a validated shortened form of

the Dundee Stress State Questionnaire (Matthews, et. al., 1999, 2002). This assessment was given as a measure of baseline subjective stress.

Equipment. The experimental task was performed on IBM-compatible computers (1.80 GHz Pentium Dual-Core, 1.96 GM RAM). Screen size was 19" with a resolution of 1280 x 1024 pixels and a refresh rate of 60 Hz.

Experimental Task. The study required participants to play a 2 dimensional computerized game (Figure 3) in a dual-task paradigm. Task 1 consisted of a game with a bucket that could be moved from left to right at the bottom of the screen. Participants moved the bucket using the right and left arrow keys. From the top of the window, two-dimensional balls fell at a rate where participants were unable to 'catch' all of the specified targets. This rate was determined by pilot testing. The goal of the game for each trial was to collect (or 'catch') balls that fell from the top of the screen by moving the bucket underneath the ball as it approached the bottom of the screen.

While 'catching' the targets, one ball changed color throughout the task (e.g., red, blue, or green). This ball appeared on the screen for a brief time before changing back to the same color contrast as the trial the participant was playing. While the ball was on the screen participants had to indicate with their left hand what color the ball was by pressing the correct indication key. They used the computer keys Z, X, and C to indicate whether they saw red (Z), blue (X), or green (C). They performed this task, Task 2, at the same time as Task 1.

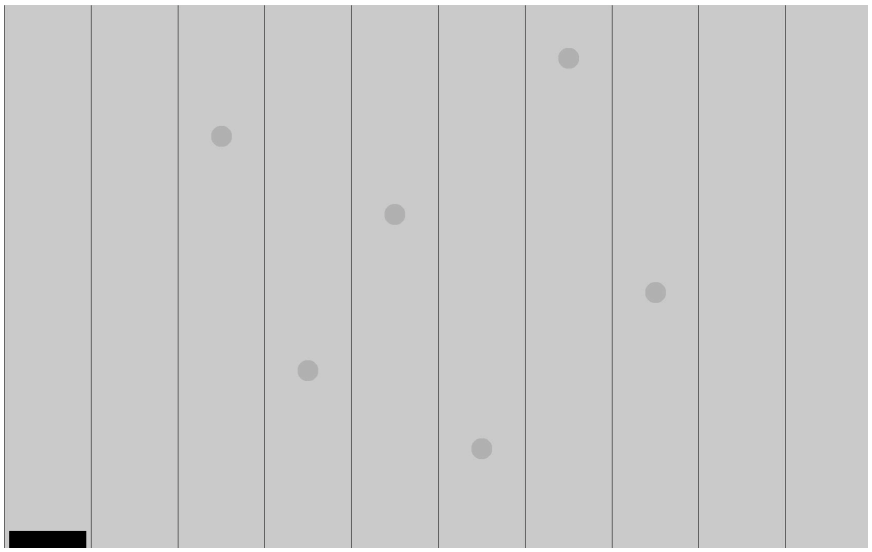


Figure 3. The experimental task was comprised of ten columns from which target balls fall from the top of the screen to the bottom of the screen. The bucket (indicated by a black line at the bottom of the screen) moves both left and right. To 'catch' a ball, the bucket must be aligned below the target as it reaches the bottom of the screen. Feedback was provided for a 'catch' by flashing the bucket green. Participants move the bucket left and right with the corresponding keyboard arrows.

Each participant completed 1 practice trial for each of the 2 tasks. Practice was followed by 5 dual-task trials where both tasks were performed simultaneously under varying conditions of difficulty (see Appendix F). During the first practice trial participants only 'caught' balls and contrast between balls and the background was high to allow participants to concentrate on learning to use the interface rather than experiencing difficulty due to experimental manipulations. In the second practice trial participants only indicated the color of the target using their left hand on the Z, X, and C keys. Once the experimental trials began, task difficulty was manipulated by changing the lightness values of the balls (decreasing contrast and making the targets more difficult to perceive) as well as increasing the number of colored balls. A pilot study was conducted comprised of undergraduate students enrolled in a research course and graduate students in the Psychology department to

determine the appropriate contrast level changes. Contrast was programmed to change using the RGB (x,x,x) units for color value where x is an integer between 0 and 255. The final rate of 2 RBG units of change in contrast level per trial was determined whereby contrast of the targets in experimental trial 5 was the same as the background of the two-dimensional field (RBG level, 199). For task 2, difficulty was manipulated by increasing the number of possible indications from 13 to 37 and from one color to three colors (Figure 4). Each trial lasted approximately 2 minutes. Trials were counterbalanced so that participants did not receive trials from highest contrast to lowest contrast. This ensured that participant's subjective stress and subjective workload reports were indicative of each trial rather than building throughout the task.

NASA-TLX. A modified computerized NASA-TLX, (Appendix B, Sharek, 2011) based on Hart and Staveland's (1988) original assessment, was administered after each trial to assess subjective workload. Participants rated their subjective feelings on 6 dimensions (physical demand, temporal demand, mental demand, performance, effort, frustration). Once participants rated these items they were shown the dimensions in pairs and asked to choose which dimension contributed more to their subjective workload during the preceding trial. Participants took the assessment 8 times throughout testing; once after each practice trial, once between each of the 5 experimental trials and once more following the administration of the FFOV.

Task-SSSQ. This assessment was a modified form of the Life-SSSQ. Modifications were made to the statements to reflect the assessment of task stress (e.g., "I felt dissatisfied during the last trial"). This assessment was administered alongside the NASA-TLX.

			Experimental Trials Counterbalanced across all Participants				
	Practice 1	Practice 2	Experimental Trial 1	Experimental Trial 2	Experimental Trial 3	Experimental Trial 4	Experimental Trial 5
Task	Task 1	Task 2	Dual Task	Dual Task	Dual Task	Dual Task	Dual Task
Task 1 Objectives	144 Possible Catches (High Contrast)		144 Possible Catches RBG(191,191,191)	144 Possible Catches RBG(193,193,193)	144 Possible Catches RBG(195,195,195)	144 Possible Catches RBG(197,197,197)	144 Possible Catches RBG(199,199,199)
Task 2 Objectives		37 Possible Indications (3 colors)	13 Possible Indications (1 color)	25 Possible Indications (2 colors)	25 Possible Indications (2 colors)	37 Possible Indications (3 colors)	37 Possible Indications (3 colors)

Higher Difficulty

Figure 4. Participants had two practice sessions, one for each task completed in the order presented here. The experimental trials were counterbalanced so that no participant experienced the same order of trials. Difficulty was manipulated by both reducing target contrast and increasing indicating demands.

Design

Performance scores were calculated by adding the sum of correct color indications (i_c) and dividing by the total possible indications (i_p) which produced an indicating percentage. Next, the indicating percentage was added to the sum number of balls caught (c_c) divided by the total possible catches (c_p) (equation below). By calculating the scores using this equation we ensured that both tasks were equally weighted.

$$(c_c/c_p) + (i_{ci}/i_p) = \text{Performance score}$$

Difficulty was considered a predictor of performance at the trial level. Age group and neuroticism were considered predictors at the person level.

Procedure

Once informed consent was collected in the laboratory participants were given the demographics form and then assessed for colorblindness and near vision. All participants took the contrast sensitivity assessment and Life SSSQ followed by a series of ability tests. Then, participants were given instructions on how to complete the experimental task and given the opportunity to ask questions. Participants then completed the practice trials. Once finished, questions were addressed and the participants began the experimental task. After every trial, participants completed the computerized NASA-TLX and SSSQ.

Following this participants completed the FFOV assessment and responded to the Task-SSSQ and NASA-TLX. Last, participants responded to a computerized exit interview, were debriefed, and given course credit or compensation for their time. In general, younger adult sessions took 2 hours to complete while older adults sessions varied from 3 to 4 hours.

RESULTS

Participant demographics, ability test information, and general descriptors may be found in Table 1. Task performance, task stress, and task workload data were examined via multilevel models that incorporated personality, age group, and difficulty as predictors.

Overview

A multilevel modeling approach was employed to determine individual variation in a 2 level hierarchy. Level 1 variables were examined at the trial level and were measured repeatedly. These measures included performance, workload and stress. Age and neuroticism were Level 2 variables and were measured once. By examining performance, subjective

workload and subjective stress at multiple levels it was possible to see intra-individual change or the change that occurs around a participants own average at various levels of difficulty (Neupert, Miller, & Lachman, 2006) and to make inferences about the effects of age group and neuroticism within-person (Lee & Bryk, 1989).

Table 1. Descriptive statistics for ability tests (and *n*'s) and differences by age group

	N	Younger *M (SD)*	*N*	Older *M (SD)*	*F*
Age	46	20.08(1.47)	26	75.47(5.58)	
Life Stress	45	45.76(12.4)	26	34.41(9.75)	15.6***
Neuroticism	45	28.22(3.54)	26	26.42(2.90)	4.79**
Near Vision	46	27.39(12.4)	25	51.50(30.9)	21.8***
Contrast Sensitivity	46	7.19(.57)	26	4.1(.76)	375.2***
Choice Reaction Time	44	338.64(54.8)	24	416.2(72.3)	24.7***
Mental Rotation	46	30.04(7.89)	25	24.88(5.14)	8.68**
UFOV 30	46	33.13(17.6)	22	12.91(4.62)	28.0***

Note: ** $p < .01$, *** $p < .001$

Note: Age was measured in years; contraast sensitivity via Ginsburg (1984); Mental Rotation via Vandenberg and Kuse (1978); Functional field of view (FFOV) at 30 degrees by Feng, Spence and Pratt (2007); Neuroticism via IPIP-NEO; Performance on game was calculated (see results); Workload measured via the NASA-TLX (Hart & Staveland, 1988); Task stress through the SSSQ (Helton, 2004).

Analyzing subjective stress data where estimates of both levels of variability are made, has been touted by Lazarus (2000) as most meaningful because it is the best substitute for a traditionally longitudinal design.

In multilevel modeling, predictors are typically grand mean centered. However, because age group (young adults = 0, older adults = 1) was dichotomous and difficulty had a meaningful zero, neuroticism was the only predictor grand mean centered for this analysis. The basic equation used for the three models presented can be seen below:

Level 1: $DV_{it} = \beta_{0it} + \beta_{1it}(\text{Difficulty}) + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}\,(\text{Age Group}) + \gamma_{02}\,(\text{Neuroticism}) + u_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}\,(\text{Age Group}) + \gamma_{12}\,(\text{Neuroticism}) + u_{1i}$

In Level 1, the intercept, β_{0it}, was defined as the expected level of the DV for person i. The difficulty slope, β_1, was the expected change in the DV associated with each level of difficulty. The error term, r_{it}, represented a unique effect associated with person i (i.e., how much that individual changed or varied in the DV over difficulty level).

The individual intercepts (β_{0i}) and slopes (β_{1i}) become the outcome variables at Level 2 where the average level of the DV for the sample when difficulty equaled zero was represented (γ_{00}). The independent contribution of Age Group (γ_{01}) and Neuroticism (γ_{02}) on the DV was also represented in the analysis. This showed whether a main effect of Age Group or Neuroticism on the DV existed. The extent to which people vary from the sample average of the DV was represented by u_{0i}. Further, the extent of change in the DV demonstrated in varying levels of difficulty was represented by γ_{10} (main effect of Difficulty). Likewise, the extent of change in the DV a person demonstrated due to Neuroticism was represented by γ_{20} (main effect of Neuroticism). γ_{11} represented whether there was a change in the DV due to an Age Group × Difficulty interaction. γ_{12} represented whether there was a change in the DV due to a Neuroticism × Difficulty interaction. The extent to which people varied from the sample slope of the DV was represented by u_{1i}. In addition, our hypotheses predicted a range of individual differences that might have accounted for differences in the IV's. For this reason, difficulty was allowed to vary from the average in each model.

Model 1: Performance

A fully unconditional model was employed where only performance was entered as the dependent variable. Results from this analysis indicated that 36% of the variability in performance was between people (Level 2) ($\tau_{00} = .03$, z = 4.33, $p < .001$) while 64% of the variability was within-person (Level 1) ($\sigma^2 = .05$, z = 11.97, $p < .001$). Results from this analysis determined that there was enough variability at both levels for further analysis of performance.

Model 1 represented the performance hypotheses. The performance equation can be found below:

Level 1: $\text{Performance}_{it} = \beta_{0it} + \beta_{1it}(\text{Difficulty}) + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}\,(\text{Age Group}) + \gamma_{02}\,(\text{Neuroticism}) + u_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}\,(\text{Age Group}) + \gamma_{12}\,(\text{Neuroticism}) + u_{1i}$

Fixed Effects.

Main Effects. As predicted, younger adults (M = 0.93, SD = 0.29) performed better on the experimental task than older adults (M = 0.66, SD = 0.18), ($\gamma 01 = -.56$, t = -10.30, p <.001) (Table 2). Difficulty level was a significant predictor of performance ($\gamma 10 = -.12$, t = -13.9, p < .001).

Table 2. Unstandardized coefficients (and standard errors) multilevel model of performance

	Fully Unconditional	Model 1
Fixed Effects		
Performance, β0		
Intercept, γ00	.83***(.03)	1.18***(.03)
Age, γ01		-.56***(.054)
Neuroticism, γ02		.009(.007)
Difficulty slope, β1		
Intercept γ10		-.12***(.008)
Difficulty × Age, γ11		.13***(.014)
Difficulty × Neuroticism, γ12		-.003(.001)
Random Effects		
Performance (τ00)	.03***(.006)	.04***(.012)
Covariance of Difficulty(τ10)		-0.006*(0.002)
Variance of Difficulty(τ11)		.0007(.0006)
Covariance of Age Group(τ20)		-.006(.01)
Variance of Age Group(τ21)		.001(.002)
Variance of Slope in Age Group (τ22)		.007***(.0006)
Covariance of Neuroticism(τ30)		.002(.002)
Covariance of Neuroticism and Difficulty(τ31)		-.0006(.0005)
Covariance of Neuroticism and Age Group (τ32)		.0006(.001)
Variance of slope in Neuroticism(τ33)		.00007(.0003)
Within-person fluctuation (σ2)	.05***(.004)	.02***(.002)

Note: *p<.05, **p<.01, ***p<.001

Interactions. Differences between age group performance were further explained by an Age Group × Difficulty interaction (γ_{11} = .13, t = 9.38, p < .001). That is, that while older adults performed significantly worse on the experimental task than younger adults (Low Difficulty; β = .60, p < .001; High Difficulty; β = .97, p < .001) younger adults decreased at a greater rate (β = -.62, p < .001) than older adults (β = -.49, p < .001) across difficulty (Figure

5). When young adults were at lowest difficulty level they performed significantly better than older adults at the same level. Likewise, younger adults out performed older adults significantly at the highest level of difficulty. The rate of change (slope) that younger adults demonstrated from low to high difficulty was significantly different from older adults.

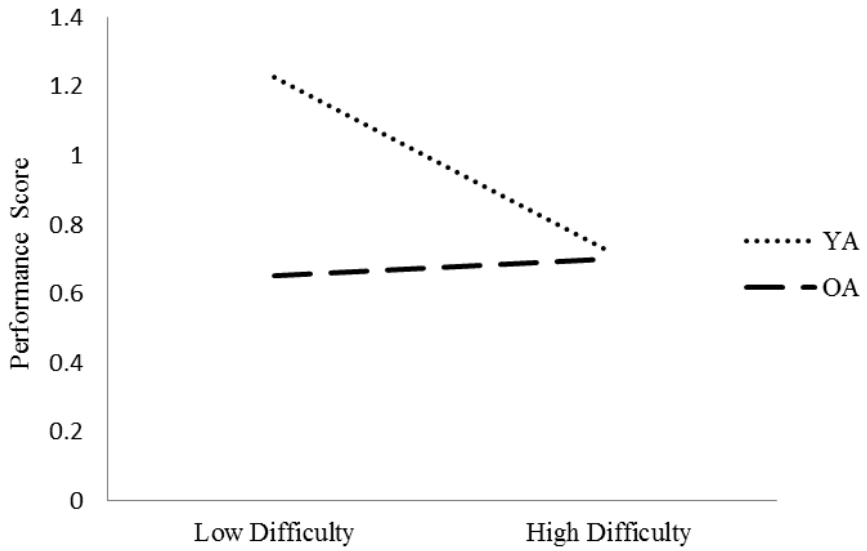


Figure 5. Age Group × Difficulty interaction on Performance. The units on the y-axis represent performance score out of 2. Points were plotted by conducting a slopes and intercepts model where the estimates are calculated using the means and standard deviations of difficulty and age group. A simple slopes test revealed significant differences between younger and older adult performance rate from low to high difficulty. Tests of contrast show that young adults significantly out-performed older adults at both levels of difficulty.

Random Effects. Participants did vary significantly between-person in performance (τ_{00} = .04, z = 3.35, p = .0004). Performance varied significantly within-person (σ^2 = .02, z = 10.26, p < .001). Variance around the slope of performance was also significant by age group (τ_{22} = .007, z = 10.26, p < .001). This suggests that there may be variance in performance that is unaccounted for by age group alone.

This model accounted for 47% of the previously reported 64% within-person variance in performance. Using the Snijers and Bosker (2011) solution for pseudo-R^2 values at level 2, a 1- ratio of reduction of $((\sigma^2/n) + \tau_{00})$ from the unconditional and conditional model with constrained slopes was calculated: 1 – (constrained/unconditional). Therefore, this model accounted for 2% of the previously reported 36% between-person variance in performance.

Non-significant results. Neuroticism was not significantly related to performance. There were also no Difficulty × Neuroticism or Difficulty × Age Group interactions. (Table 2).

Further Inspection of Significant Interaction. To further investigate the significant decrease in performance by younger adults, Performance Score by Difficulty Level was plotted (Figure 6). This showed that older adults did not show variability in performance across dificulty level as expected.

In order to gain a better understanding of older adults' stable performance, a dual-axis scatterplot of older adult performance for both tasks was plotted across difficulty levels (Figure 7). This indicated that at higher levels of difficulty, older adults exhibited performance that suggests they were task switching rather than performing both tasks equally. This is evidenced by indicating scores remaining high despite dramatic declines in performance on the catching task.

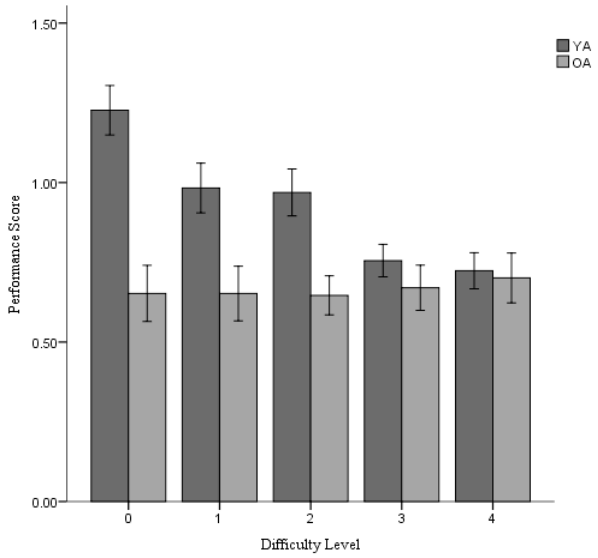


Figure 6. Bar graph of younger and older adult performance score for both tasks out of 2 with error bars. Older adult performance remains relatively stable over difficulty level while younger adult performance decreases at higher levels of difficulty.

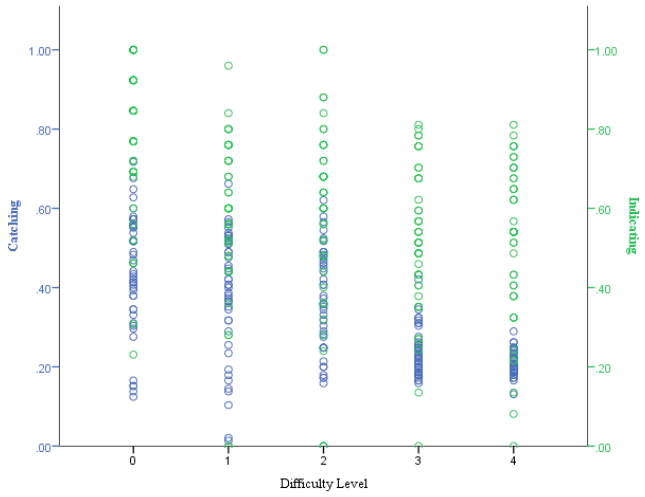


Figure 7. Dual-axis scatterplot display of older adult performance for Task 1 (Catching) and Task 2 (Indicating) over difficulty level. At higher levels of difficulty older adults performed better at the indicating task. However their performance remained relatively poor on the catching task at higher levels of difficulty.

Model 2: Workload

Model 2 was derived to best reflect the hypotheses of task workload:

Level 1: $\text{Workload}_{it} = \beta_{0it} + \beta_{1it}(\text{Difficulty}) + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}$ (Age Group)$+ \gamma_{02}$ (Neuroticism)

$\beta_{1i} = \gamma_{10} + \gamma_{11}$ (Age Group)$+ \gamma_{12}$ (Neuroticism)

Both Difficulty and Neuroticism were constrained to fit the average slope in order to represent the most parsimonious model. This constraint is represented by excluding u_{0i} , and u_{1i} in the β_{1i} outcome variable at level 2.

A fully unconditional model was conducted where no predictors were entered in the model. This concluded that 53% of the variance in workload was between-person (level 2) (τ_{00} = 133.91, z = 4.96, $p < .001$) while 47% of the variance was within-person (level 1)(σ^2 = 120.63, z = 11.83, $p < .001$). This was a sufficient amount of variability to warrant further analysis.

Fixed Effects. As predicted, age group was significantly related to workload (γ_{01} = 10.04, t = 3.06, p = .003) (Table 3). Therefore, on average older adults reported more workload than younger adults (Figure 8).

Random Effects. Participants varied significantly between-person in reports of workload (τ_{00} = 115.13, z = 3.87, $p < .001$). Participants also varied significantly within-person (σ^2 = 116.51, z = 11.68, $p < .001$). There was significant variance around the slope of age group for workload (τ_{11} = 5.68, z = 11.68, $p < .001$). This means that people did vary in their relationship of workload based on their age group (Figure 8).

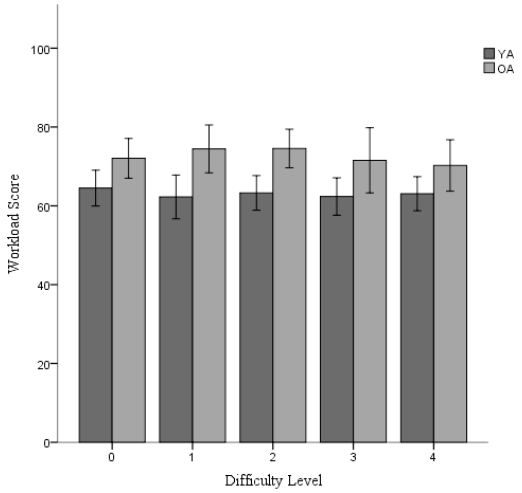


Figure 8. Bar graph of younger and older adult workload scores with error bars across difficulty level. Younger and older adults showed little variance over difficulty level. However, younger adults reported significantly lower workload than older adults at all levels of difficulty. Error bars illustrate the significant within person variance in workload within person. Comparison of error bars between age groups illustrate significant between-person variance.

Non-significant results. Difficulty and Neuroticism were not significantly related to workload. There was also no Difficulty × Neuroticism interaction or Difficulty × Age Group interaction (Table 3).

Table 3. Unstandardized coefficients (and standard errors) multilevel model of workload

	Fully Unconditional	Model 2
Fixed Effects		
Workload, β0		
Intercept, γ00	66.3***(1.5)	64.1***(2.0)
Age, γ01		10**(3.2)
Neuroticism, γ02		.12(.47)
Difficulty slope, β1		
Intercept γ10		-.30(.51)
Difficulty × Age, γ11		-.18(.88)
Difficulty × Neuroticism, γ12		.16(.12)
Random Effects		
Performance (τ00)	133.9***(26.9)	101.66***(21.8)
Covariance of Age Group(τ10)		-22(20.7)
Variance of Age Group(τ11)		5.67***(.48)
Within-person fluctuation (σ2)	120.6***(10.1)	116.51***(9.9)

Note: *p<.05, **p<.01, ***p<.001

Model 3: Stress

Model 2 reflected the hypotheses of stress and is seen below:

Level 1: $\text{Stress}_{it} = \beta_{0it} + \beta_{1it}(\text{Difficulty}) + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}\,(\text{Age Group}) + \gamma_{02}\,(\text{Neuroticism}) + u_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}\,(\text{Age Group}) + \gamma_{12}\,(\text{Neuroticism})$

This model represents the same justifications for constraining Difficulty and Neuroticism as was done in the workload model.

A fully unconditional model was conducted where no predictors were entered. 20%

(τ_{00} = .006, z = 3.24, p < .001) of the variance was between-person, and 80% (σ^2 = .022, z = 11.94, p < .001) of the variance was within-person. This provided evidence of enough variability at both levels to continue analysis.

Fixed Effects. As predicted, Age Group was significantly related to Stress (γ_{01} = -.13, t = -4.26, p <.001) (See Table 4). Therefore, on average older adults reported less stress than younger adults (Figure 9).

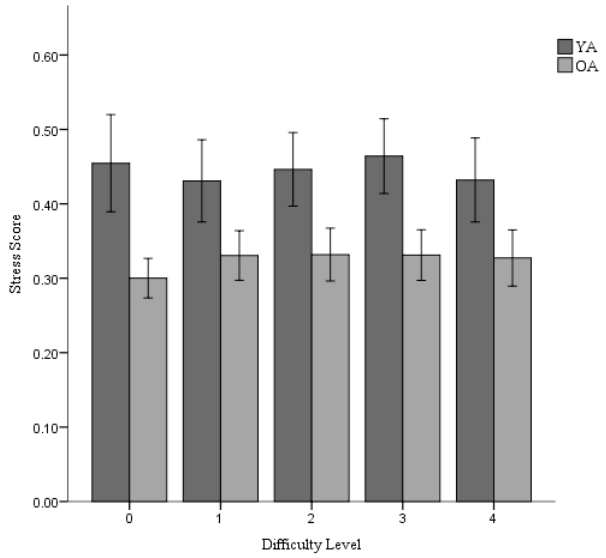


Figure 9. Bar graph of younger and older adult stress scores with error bars across difficulty level. Younger and older adults showed little variance over difficulty level. However, younger adults reported significantly higher stress than older adults at all levels of difficulty.

Random Effects. Participants significantly varied between themselves in their reports of stress (τ_{00} = .003, z = 2.01, p <. 05). Variability within-person was also significant (σ^2 = .02, z = 11.79, p < .001). There was significant variance around the slope of Age Group for

Stress (τ_{11} = .001, z = 11.79, p < .001). Therefore, participants varied in reported stress based on their age group.

Non-Significant results. Neuroticism and Difficulty were not significantly related to stress. There were also no Difficulty × Neuroticism or Difficulty × Age Group interactions (Table 4).

This model accounted for 14% of the previously reported 47% within-person variance in performance. This model accounted for 3% of the previously reported 53% between-person variance in performance.

Table 4. Unstandardized coefficients (and standard errors) multilevel model of stress

	Fully Unconditional	Model 3
Fixed Effects		
Stress, β0		
Intercept, γ00	.40***(.01)	.44***(.01)
Age, γ01		-.13***(.03)
Neuroticism, γ02		.0008(.004)
Difficulty slope, β1		
Intercept γ10		.00004(.007)
Difficulty × Age, γ11		.005(.01)
Difficulty × Neuroticism, γ12		.0002(.001)
Random Effects		
Performance (τ00)	.005***(.001)	.02***(.005)
Covariance of Age Group(τ10)		-.002*(.001)
Variance of Age Group(τ11)		.001***(.00008)
Within-person fluctuation (σ2)	.02***(.001)	.03***(.003)

Note: *p<.05, **p<.01, ***p<.001

Follow-up Analysis

A Pearson's *r* correlation was conducted to determine if there was a significant relationship between subjective workload and subjective stress. Results indicated no significant relationship between workload and stress ($r = -.44$, $n = 348$, $p = .21$).

DISCUSSION

This study measured workload and stress under varying levels of difficulty to determine the role of age group and neuroticism in performance. In general, all participants performed better when difficulty was low. Younger adults performed worse when difficulty was high. Older adult performance remained relatively stable across difficulty. As predicted, older adults reported significantly higher workload across all difficulty levels than younger adults. Also, older adults reported significantly lower stress across all difficulty levels than younger adults. There was no relationship found between the report of workload, stress or performance for those participants who scored highly on the neuroticism trait.

Stress and Workload

This study highlighted the differences between the subjective feeling of stress and workload. Stress and workload were not significantly correlated. This finding reveals that participants' experience of workload and stress were distinct from one another while performing the task. This is an important finding because it supports the hypothesis that workload and stress are fundamentally different constructs that measure different aspects of the subjective and cognitive aspects of performance.

In general, it was expected that participants would report higher workload and stress at higher levels of difficulty. This was not found to be true in this study. Rather, participants reported relatively stable experiences of workload and stress for the duration of the experimental task. A possible explanation of this finding is that task demands did not change enough by difficulty level to result in a change in the experience of workload and stress for the sample. For example, workload definitions typically include some aspect of task demands versus mental resource capacity (e.g., Eggermeier, Wilson, et al., 1991). If the change in difficulty level did not result in a noticeable difference of resource capacity requirements, participants would feel the same mental strain and report similar feelings of workload across all difficulty levels. A similar concept may also apply to stress. That is, since there was no noticeable change in goals or ability to attain those goals (Sanders, 1983), participants would not experience any change in the feelings of stress across difficulty levels.

Individual Differences

In an effort to accurately measure the influence of stress and subjective workload on an individual when performing a difficult task, individual differences that may mitigate or otherwise influence performance were considered. In this study neuroticism did not have a significant relationship to performance, workload, or stress. However, mixed results have been found with regard to neuroticism and its relation to cognitive processing (Mathews & Mackintosh, 1998; Revelle, 1993; DeYoung, Peterson, & Higgins, 2005). Evidence suggests that other personality traits, such as extroversion or conscientiousness, may influence the way in which one experiences stress (e.g., Matthews, 2001; Szalma, et. al., 2006, Teo et. al., 2011). Continuing to examine results in the light of personality differences allows for further

analysis into possible coping mechanisms employed (Matthews, 2001; Teo, et. al., 2011) that may explain differences found in performance. If personality and coping differences arise in stress research, we expect to also see differences in how subjective workload is experienced. Though this was not an outcome found here, it is important to continue to research this area so that we might better understand how personality traits may influence individuals who are managing high stress and subjective workload to compensate for task difficulty.

Age group did significantly influence performance in this study. As noted in the introduction, aging influences in cognitive and motor ability were highlighted by our results in two ways. First, the performance rate of younger and older adults at the lowest difficulty level were significantly different where younger adults outperformed older adults. Second, at the highest difficulty level, older adults exhibited task switching behavior. This would support subjective workload literature (Wickens, 2002) where offloading of tasks occurs when task demand limits are reached. . While age group was a significant predictor of performance, there remained a significant amount of performance variance to be explained within subjects. This supports the examination of repeated measures such as subjective workload and subjective stress.

Older adult subjective workload reports were significantly higher than younger adults suggesting that older adults required more effort than younger adults in the task. What was of interest was the extent to which these groups were different. As noted earlier, older adults' performance did not change as much as younger adult performance at high levels of difficulty even though effort was reported significantly higher for older adults. This may indicate that older adults' task demands were exceeding resources even at lowest levels of difficulty.

Stress was also reported differently for younger and older adults. Previous research has shown that older adults frequently rate stressful events as less stressful than younger adults (Almeida & Horn, 2004). These findings were supported by this study underscoring the hypothesis that subjective stress should be measured independently of subjective workload as it was experienced differently by age group.

The Experience of Challenge

It was expected that by measuring subjective workload, it would have been possible to predict when the task was too difficult to maintain performance once task demands exceeded resource capability. Although subjective workload varied significantly within-person, between-person and as a function of age group, it did not change significantly over difficulty level. By understanding each of these findings it may be possible to inform the use of subjective workload as a proxy for challenge.

First, the significant within-person variability in subjective workload could indicate that the NASA-TLX was not a good measure of subjective workload for this task. This is because workload should change as a function of difficulty, rather than only as a function of the individual taking the assessment. Second, the significant between-person variability results indicate that individuals were different from one another in the way they reported workload. This supports individual differences research that states differences would influence the way in which workload is perceived, interpreted and its outcomes on performance. Third, that age group played a significant role in the report of subjective workload is further evidence that individual differences in the feeling of task demand is an important consideration in performance research. This is especially true for aging research

since older adults' subjective workload was reported significantly higher than younger adults. Taken together, these findings reflect two primary requirements of using workload as a proxy for challenge; 1) workload must vary as a function of difficulty, and 2) workload must vary between-persons at the person level. If both of these conditions are satisfied, we may come closer to using subjective workload as one side of the experience of challenge.

The second side of challenge was to measure subjective stress. It was expected that as task difficulty increased, so would task stress. This relationship was not observed. Moreover, stress did not change significantly over difficulty level for either age group. However, just as was found for workload, significant within-person, between-person and between group variance was observed. Each of these findings underscores that the requirements outlined for workload also apply to stress as a proxy for challenge: 1) stress must vary as a function of difficulty, and 2) stress must vary between persons at the person level. By meeting both of these requirements, stress may be used as a way to observe when an individual is experiencing challenge.

Limitations

A basic limitation of this study was the unequal sample sizes. Second, an unintended consequence of the study was the differential amount of time spent in the appointment session for older and younger adults. Although time spent in the experimental task was very similar, future studies should control for total time spent in session.

A third limitation for consideration is the relatively stable reports of subjective workload and stress. There are several ways to correct for this in future studies. First, task demands may need to be more noticeably different by difficulty level so that the degree of

mental strain is more overt. This change in task demands may allow for more fluctuation in feelings of subjective workload as difficulty is manipulated. Second, it may be necessary to ensure individuals are stable within themselves in their reports of both subjective stress and workload. That is, subjective stress and workload reports should change as a function of difficulty.

Finally, it would be advantageous to include pre-established performance goals at each difficulty level. This protocol would clearly present performance goals at the start of the task. This approach may allow stress to be more fully realized at difficulty levels where participants are unable to reach those pre-established goals.

Theoretical Implications

This study was the first to attempt to measure subjective workload and stress where individuals experience a point in a task where their performance was no longer maintained due to a change in difficulty. It was expected that this would result in high levels of subjective workload and stress and that these variables would be able to tell us important information about challenge that was not necessarily explained by examining performance alone. This did not happen. However, examining stress and subjective workload as separate variables did reveal that these variables were experienced differently based on age group differences and should be examined as such in the future. Commensurate with findings in stress research (Hancock & Warm, 1989; Hockey, 1997; Matthews, et. al., 2006) it is expected that workload will be included as an important predictor of challenge where performance is an important outcome in future studies.

Practical Contributions

The results of this study inform researchers of important considerations in the use of subjective workload and subjective stress as proxies for the experience of challenge. Group differences found in the report of each construct indicate a need for special attention in the area of stress and workload toward the creation of programs intended for these specific audiences. Likewise, the results direct researchers to examine current performance conditions that require a serious consideration of workload and stress levels as they impact individual performance outcomes in the face of unacceptable consequences. This research has far reaching applications including future endeavors in the areas of design, training, and human-computer interaction.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, *102*(1), 3-27. doi:10.1037//0033-2909.102.1.3

Almeida, D. M. (2005). Resilience and vulnerability to daily stressors assessed via diary methods. *Current Directions in Psychological Science*, *14*(2), 64-68. doi:10.1111/j.0963-7214.2005.00336.x

Almeida, D. M., & Horn, M. C. (2004). Is daily life more stressful during middle adulthood? In R. C. Brim, O.G., Ryff, C.D., Kessler (Eds.), *How healthy are we? A national study of well-being at midlife* (pp. 425-451). Chicago, IL: University of Chicago Press.

Ball, K., Beard, B., Roenker, D., Miller, R., & Griggs, D. (1988). Age and visual search: expanding the useful field of view. *Journal of the Optical Society of America*, *5*(12), 2210-2219.

Boyer, D. L., Pollack, J. G., & Eggemeier, F. T. (1992). Effects of aging on subjective workload and performance. *Human Factors Society 36th Annual Meeting* (pp. 156-160).

Charles, S. T., & Almeida, D. M. (2007). Genetic and environmental effects on daily life stressors: more evidence for greater variation in later life. *Psychology and aging*, *22*(2), 331-40. doi:10.1037/0882-7974.22.2.331

Craske, M. G. (1999). Worry, Anxiety, and Fear and Imminence of Threat. In *Anxiety disorders: psychological approaches to theory and treatment* (pp.1-29). Boulder, Colorado: Westview Press.

Czaja, S., & Sharit, J. (1993). Age differences in the performance of computer-based work. *Psychology and Aging 8*(1), 59-67. doi: 10.1037/0882-7974.8.1.59.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect cognitive and neuropsychological correlates of the fifth factor. *Journal of Personality, 73*(4), pp. 825-58. doi:10.1111/j.1467-6494.2005.00330.x

Ericsson, K. A., & Lehmann (1996). Expert and exceptional performance : Evidence of maximal adaptation to task constraints. *Annual Review of Psychology 47*, 273-305.

Eggermeier, F. T., Wilson, G. F., et al. (1991). Workload assessment in multi-task environments. In Damos, D. L. (Ed.), *Multiple task performance,* (pp. 329-360). London : Taylor and Francis.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology*, *56*(3), 164-76. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12271747

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*(10), 850-55.

Fozard, J. L., & Gordan-Salant, S. (2001). Changes in vision and hearing with aging. In Birren, J. E., and Schaie, K. W. (Eds.), *Handbook of the Psychology of Aging, (5th Edition,* pp. 241-260). San Diego: Academic Press. Inc.

Ginsburg, A. P. (1984). A new contrast sensitivity vision test chart. *American Journal of Optometry and Physiological Optics*, *61*(6), 403-407.

Goldberg, L., Johnson, J., Eber, H., Hogan, R., Ashton, M., Cloninger, C., & Gough, H. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84-96. doi:10.1016/j.jrp.2005.08.007

Guadagnoli, M.A., & Lee, T. D. (2004). Challenge Point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, *36*(2), 212-224.

Gunthert, K. C., Cohen, L. H., & Armeli, S. (1999). The role of neuroticism in daily stress and coping. *Journal of Personality and Social Psychology, 77*(5), 1087-100. doi: 10.1037/0022-3514.77.5.1087.

Haga, S., Shinoda, H., & Kokubun, M. (2002). Effects of task difficulty and time-on-task on mental workload. *Japanese Psychological Research*, *44*(3), 134-143. doi:10.1111/1468-5884.00016

Hancock, P. A. (1986). The effect of skill on performance under an environmental stressor. *Aviation, Space, and Environmental Medicine*, *57*(1), 59-64.

Hancock, P. A. (2009). Performance on the very edge. *Military Psychology*, *21*(sup1), S68-S74. doi:10.1080/08995600802554680

Hancock, P. A., & Szalma, J. (2008). *Performance Under Stress*. Hancock, P. A., Szalma, J., (Eds.). Burlington, VT.: Ashgate Publishing Company.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, *31*(5), 519-37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14552346

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Hancock, P. A. (Ed.), *Human and Mental Workload.* Amsterdam: North-Holland.

Helton, W. S. (2004). Validation of a short stress state questionnaire. *Human Factors*, *48*(11), 1238-1242. doi:10.1177/154193120404801107

Ho, G., Wheatley, D., & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers*, *17*(6), 690-710. doi:10.1016/j.intcom.2005.09.007

Hockey, G. R. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetical framework. *Biological Psychology*, *45*(1-3), 73-93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9083645

Ishihara, S. (1994). Ishihara's test for colour blindness (38 plates ed.). Tokyo, Japan: Kanehara Shuppan Co.

Kang, N. E., & Yoon, W. C. (2008). Age and experience-related user behavior differences in the use of complicated electronic devices. *Human-Computer Studies, 66*. pp. 425-37. doi:10.1016/j.ijhcs.2007.12.003

Lazarus, R. S. (1999). *Stress and Emotion: A new synthesis*. New York: Springer.

Lazarus, R. S. (2000). Toward better research on stress and coping. *American Psychologist*, *55*(6), 665-673. doi: 10.1037/0003-066X.55.6.665.

Lazarus, R. S., & DeLongis, A. (1983). Psychological stress and coping in aging. *The American Psychologist*, *38*(3), 245-54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6870040.

Lee, V. E. , & Bryk, A. S. (1989). A multi-level model of the social distribution of high school achievement. *Sociology of Education, 62*(3), 172-92.

Loft, S., Sanderson, P., Neal, A., & Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications. *Human Factors*, *49*(3), 376-399. doi:10.1518/001872007X197017.

Mathews, A., & Mackintosh, B. (1998). A cognitive-motivational analysis of anxiety. *Cognitive Therapy and Research*, *122*, 539-560.

Matthews, G., Emo, A. K., Funke, G., Zeidner, M., Roberts, R. D., Costa, P. T., & Schulze, R. (2006). Emotional intelligence, personality, and task-induced stress. *Journal of Experimental Psychology: Applied*, *12*(2), 96-107. doi:10.1037/1076-898X.12.2.96

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Huggins, J., & Falconer, S. (1999). *Validation of a Comprehensive Stress State Questionnaire: Towards a state "big three"?* (F. Mervielde, I., Deary, I.J., DeFruyt, F., Ostendorf, Eds.) *Personality psychology in Europe* (Vol. 7, pp. 335-350). Tilburg, Netherlands: Tilburg University Press.

Neupert, S. D., Almeida, D. M., Mroczek, D. K., & Spiro, A. (2006). Daily stressors and memory failures in a naturalistic setting: Findings from the VA normative aging study. *Psychology and Aging*, *21*(2), 424-9. doi:10.1037/0882-7974.21.2.424

Neupert, S. D., Miller, L. M., & Lachman, M. E. (2006). Physiological reactivity to cognitive stressors: Variations by age and socioeconomic status. *International Journal of Aging and Huan Development*, *62*(3), 221-35.

Newell, K. M. (1991). Motor Skill Acquisition. *Annual Review of Psychology, 42,* 213-37.

Parker, D., McDonald, L., Rabbitt, P., & Sutcliffe, P. (2000). Elderly drivers and their accidents: The aging driver questionnaire. *Accident Analysis and Prevention*, *32*(6), 751-9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10994602

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition*, *28*, 39-58.

Revelle, W. (1993). *Attention: Selection, awareness and control: A tribute to Donald Broadbent*. Baddely, A., Weiskrantz, L. (Eds.), (pp. 346-373). Oxford, England: Oxford University Press.

Rogers, W. A. (1997). Individual differences, aging, and human factors: An overview. In Fisk, A.D., Rogers, W.A. (Eds.), *Handbook of Human Factors and the Older Adult* (pp. 151- 170). San Diego: Academic Press. Inc.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140-144. doi:10.1111/j.0963-7214.2004.00293.x

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, *53*(1), 61-97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6869047

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science*, *27*(3), 525-559. doi:10.1016/S0364-0213(03)00011-9

Sharit, J., & Czaja, S. J. (1994). Ageing, computer-based task performance, and stress: issues and challenges. *Ergonomics 37*(4), 559-77. doi: 10.1080/00140139408963674.

Sharit, J., Czaja, S. J., Nair, S. N., Hoag, D. W., Leonard, D. C., & Dilsen, E. K. (1998). Subjective experiences of stress, workload, and bodily discomfort as a function of age and type of computer work. *Work & Stress, 12*(2), 135-44. doi:10.1080/02678379808256855

Snijers, T. A. B. & Bosker, R. J. (2011). How much does the model explain? *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling 2nd Edition.* (pp. 109-118). Los Angeles: Sage.

Sperandio, J. C. (1978). The regulation of working methods as a function of work-load among air traffic controllers. *Ergonomics*, *21*(3), 195-202. doi:10.1080/00140137808931713.

Suls, J. (2001). Affect, Stress, and Personality. In Forgas J. P. (Ed.), *Handbook of affect and social cognition*, (pp. 392-409). Mahwah, NJ: Lawrence Elrbaum Associates.

Suss, J., & Ward, P. (2010). Skill-based differences in the cognitive mechanisms underlying failure under stress. *Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 1062-1066).

Szalma, J. L. (2009a). Individual differences in human-technology interaction: incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science*, *10*(5), 381-397. doi:10.1080/14639220902893613

Szalma, J. L. (2009b). Individual differences in performance, workload, and stress in sustained attention: Optimism and pessimism. *Personality and Individual Differences*, *47*(5), 444-451. doi:10.1016/j.paid.2009.04.019

Szalma, J. L., Hancock, P. A., Dember, W. N., & Warm, J. S. (2006). Training for vigilance: the effect of knowledge of results format and dispositional optimism and pessimism on performance and stress. *British Journal of Psychology*, *97*, 115-35. doi:10.1348/000712605X62768

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46*(2), 219-33. doi:10.1518/hfes.46.2.219.37334

Teo, G., Szalma, J. L., & Schmidt, T. (2011) Predicting performance, subjective states and coping strategy in a vigilance task: The role of individual differences. *Human Factors and Ergonomics Society 55th Annual Meeting,* (pp.1467-71).

Tugla, M. K., & Sheridan, T. B. (1980). Dynamic decisions and work load in multitask supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-10*(5), 217-232.

Vandenberg, S.G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*, 599-604.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *The Journal of the Human Factors and Ergonomics Society, 50*(3). pp. 433-41. doi:10.1518/001872008X312152

Wickens, C. D. (1999). *Automation technology and human performance (current research and trends)*. Scerbo, M. W., Mouloua, M. (Eds.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

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

Wu, C., & Liu, Y. (2006). Queuing network modeling of age differences in driver mental workload and performance. *Human Factors and Ergonomics Society 50th Annual Meeting,* (pp. 1117-21).

Young, M. S., & Stanton, N. A. (2001). Mental workload: theory, measurement, and application. In Karwowski, W. (Ed.), *International Encyclopedia of Ergonomics and Human Factors,* (Volume 1 ,pp. 507-09). London: Taylor & Francis.

Young, M.S., & Stanton, N.A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics, 50*(8), 1324-39. doi:10.1080/00140130701318855

APPENDICIES

APPENDIX A

Demographics Questionnaire

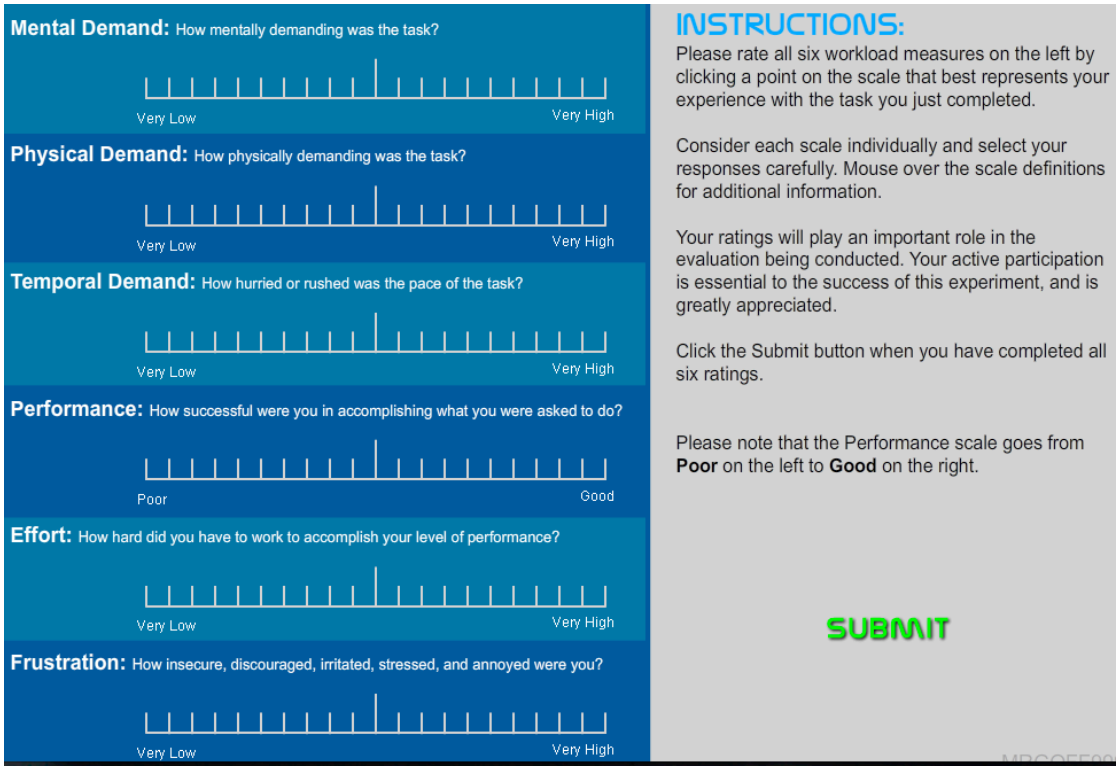
1. Please enter your participant ID
2. Please enter your date of birth
3. What is your sex?
    a. Male
    b. Female
4. Indicate the highest level of completed education
    a. Less than high school graduate
    b. High school graduate or G.E.D.
    c. 2 year technical degree or technical college
    d. Certificate program
    e. 2 years or less of undergraduate status at a 4 year university
    f. 2 years or less of undergraduate status at a 4 year university without a degree.
    g. Bachelor of Arts or Bachelor of Science degree
    h. Some Graduate school
    i. Master's degree
    j. M.D., J.D., Ph.D., or other advanced degree
5. Indicate the race/ethnicity that best describes you
    a. American Indian/Alaskan Native
    b. Asian American/Pacific Islander
    c. Black/African American
    d. Hispanic/Latino
    e. White/Caucasian
    f. Other _____
    g. Multiracial
6. Indicate your occupational status
    a. Full time student
    b. Full time student/part time work
    c. Part time student/part time work
    d. Part time student/full time work
    e. Full time work
    f. Part time work
    g. Retired/not working
    h. Retired/part time working
7. What is your occupation?
8. If you are a student, what college are you currently enrolled in?
    a. Agriculture & Life Sciences
    b. Design
    c. Education
    d. Engineering
    e. Humanities & Social Sciences
    f. Management

g. Natural Resources
h. Physical & Mathematical Sciences
i. Textiles
j. Veterinary Medicine
k. First Year College
l. Distance Education
m. Other _____

9. Is English your first language?
    a. Yes
    b. No
10. Do you play video games?
    a. Yes
    b. No
11. How would you describe your feelings about playing video games?
    a. I do not generally like to play video games
    b. I do like video games
    c. I only like to play one specific video game
    d. I only like certain types of video games and only play types that are in the same genre
12. Indicate about how much time you spend playing video games.
    a. A few times a week
    b. A few times a day
    c. Only once or twice a day but for long periods of time
    d. Only once or twice a week but for long periods of time.
    e. More than any of the above options
    f. Less than any of the above options
13. What would you say are the best reasons why you play video games (choose all that apply)?
    a. For enjoyment
    b. To practice skills you deem important
    c. To procrastinate
    d. To escape reality
    e. To socialize
    f. To relax
    g. To sharpen your mind
    h. To have some "me time"
    i. To do something that requires little effort
    j. Other _____

APPENDIX B

NASA-Task Load Index (TLX)



APPENDIX C

International Personality Item Pool- Neuroticism Extroversion Openness (IPIP-NEO)

In the following survey, there are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes ***you***. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then click the bubble that corresponds to your agreement or disagreement with the statement.

	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Often feel blue.	○	○	○	○	○
Feel comfortable around people.	○	○	○	○	○
Believe in the importance of art.	○	○	○	○	○
Have a good word for everyone.	○	○	○	○	○
Am always prepared.	○	○	○	○	○
Rarely get irritated.	○	○	○	○	○
Have little to say.	○	○	○	○	○
Am not interested in abstract ideas.	○	○	○	○	○
Have a sharp tongue.	○	○	○	○	○

	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Waste my time.	○	○	○	○	○
Dislike myself.	○	○	○	○	○
Make friends easily.	○	○	○	○	○
Have a vivid imagination.	○	○	○	○	○
Believe others have good intentions.	○	○	○	○	○
Pay attention to details.	○	○	○	○	○
Seldom feel blue.	○	○	○	○	○
Keep in the background.	○	○	○	○	○
Do not like art.	○	○	○	○	○

	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Cut others to pieces.	○	○	○	○	○
Find it difficult to get down to work.	○	○	○	○	○
Am often down in the dumps.	○	○	○	○	○
Am skilled at handling social situations.	○	○	○	○	○
Tend to vote for liberal political candidates.	○	○	○	○	○
Respect others.	○	○	○	○	○
Get chores done right away.	○	○	○	○	○
Feel comfortable with myself.	○	○	○	○	○
Would describe my experiences as somewhat dull.	○	○	○	○	○

	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Avoid philosophical discussions.	○	○	○	○	○
Suspect hidden motives in thers.	○	○	○	○	○
Do just enough work to get by.	○	○	○	○	○
Have frequent mood swings.	○	○	○	○	○
Am the life of the party.	○	○	○	○	○
Carry the conversation to a higher level.	○	○	○	○	○
Accept people as they are.	○	○	○	○	○
Carry out my plans.	○	○	○	○	○
Am not easily bothered by things.	○	○	○	○	○

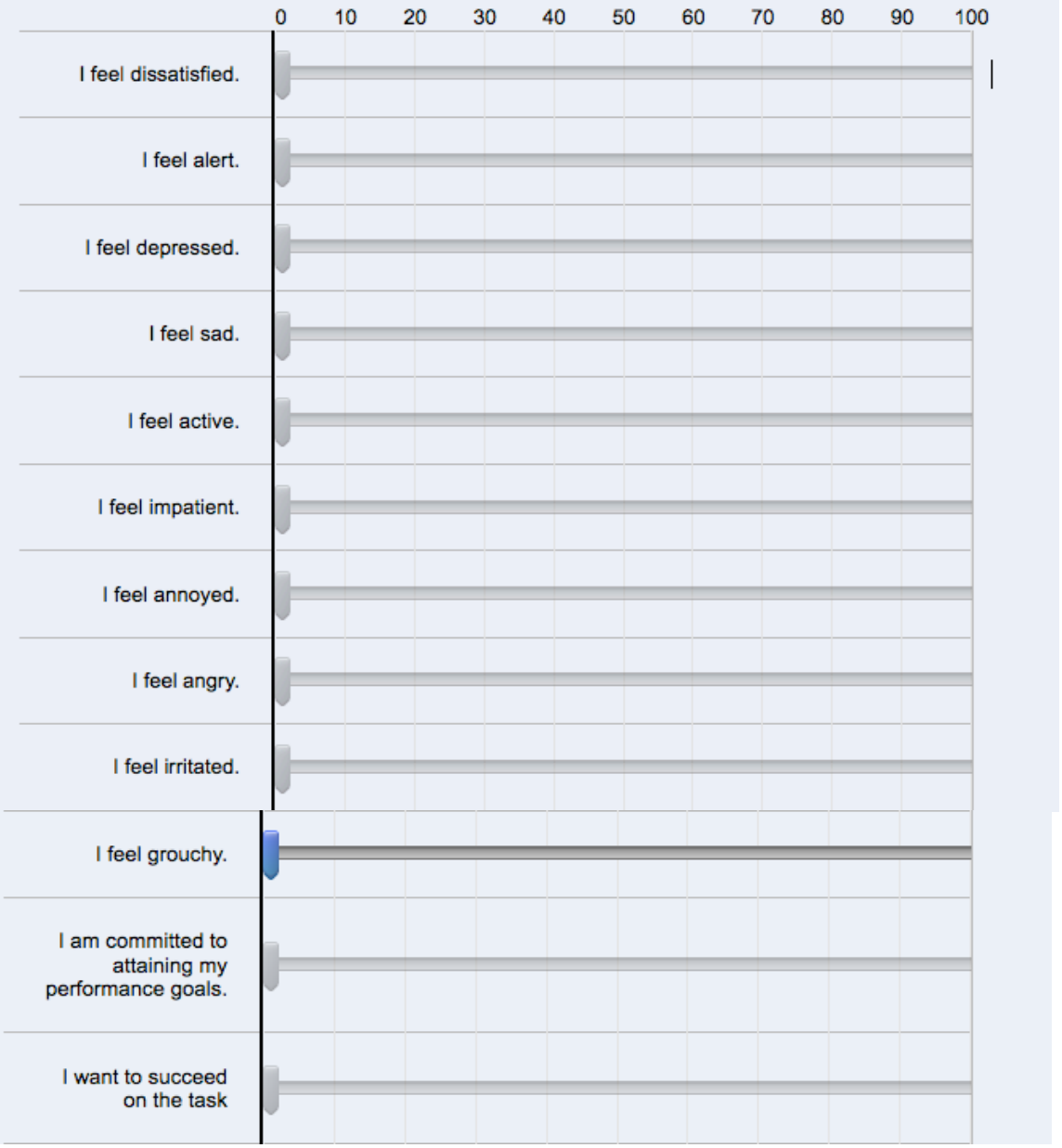
	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Don't like to draw attention to myself.	○	○	○	○	○
Do not enjoy going to art museums.	○	○	○	○	○
Get back at others.	○	○	○	○	○
Don't see things through.	○	○	○	○	○
Panic easily.	○	○	○	○	○
Know how to captivate people.	○	○	○	○	○
Enjoy hearing new ideas.	○	○	○	○	○
Make people feel at ease.	○	○	○	○	○
Make plans and stick to them.	○	○	○	○	○

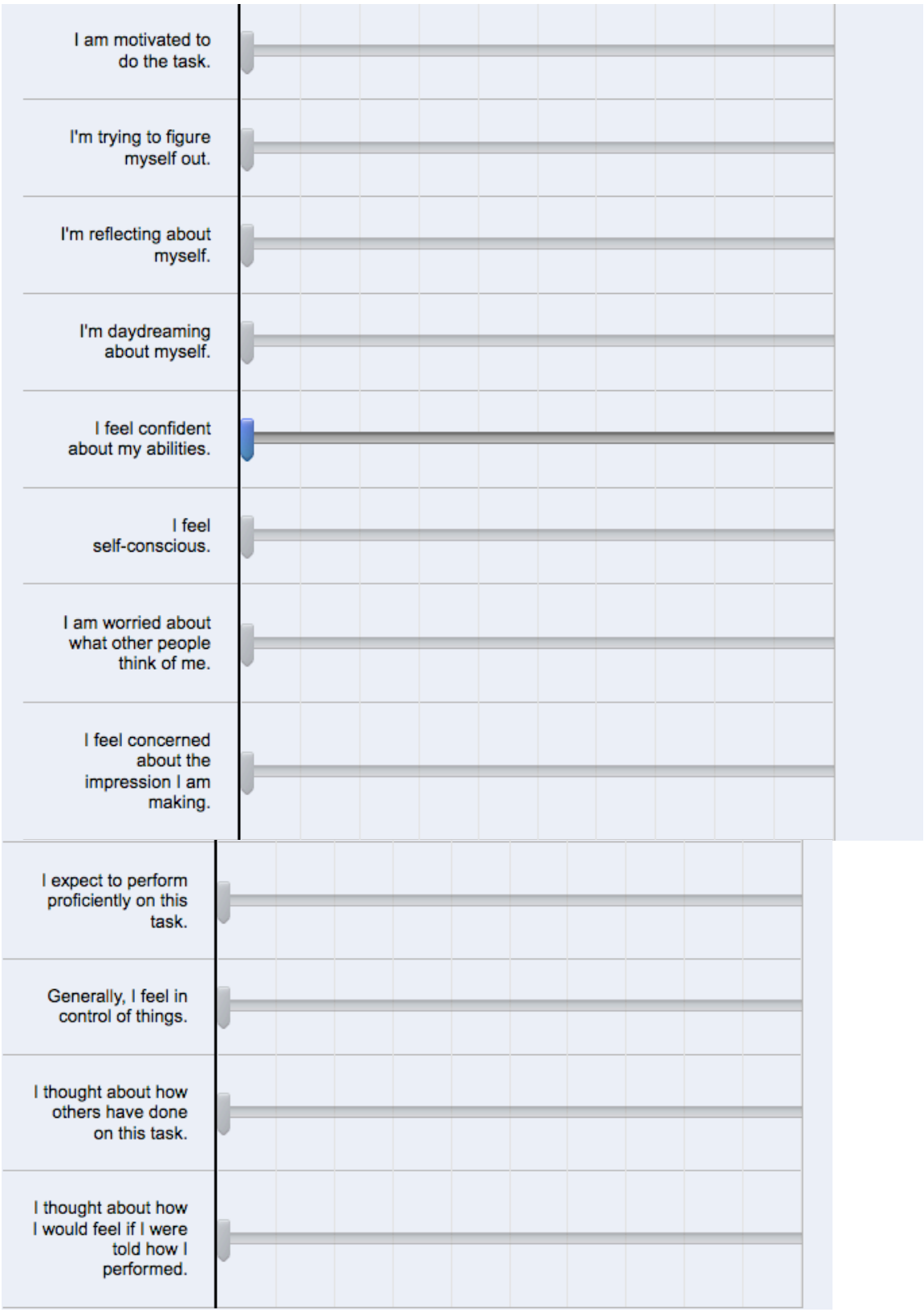
	Very Inaccurate	Moderately Inaccurate	Neither Inaccurate Nor Accurate	Moderately Accurate	Very Accurate
Am very pleased with myself.	○	○	○	○	○
Don't talk a lot.	○	○	○	○	○
Tend to vote for conservative political candidates.	○	○	○	○	○
Insult people.	○	○	○	○	○
Shirk my duties.	○	○	○	○	○

APPENDIX D

Shortened State Stress Questionnaire (SSSQ)

This questionnaire is designed to measure the way you feel about the amount of stress you just experienc from the task. It is not a test, so there are no right or wrong answers. Answer each item as carefully and as accurately as you can by placing the slider on the scale under the number that best describes how accura this phrase is for you. Please note that "0" indicates you do not feel the phrase describes you right now, a "100" indicates you feel the phrase completely describes you right now.

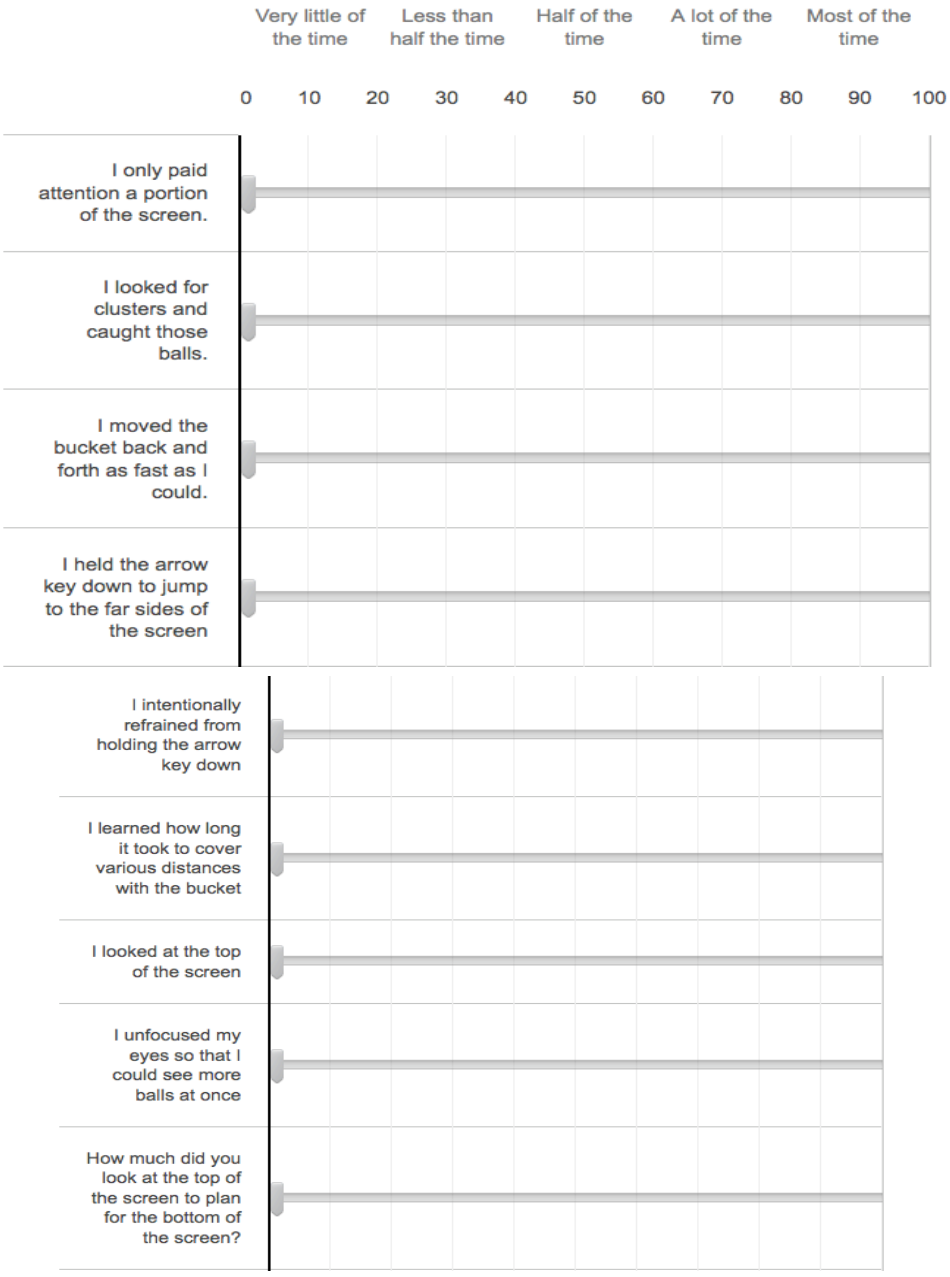


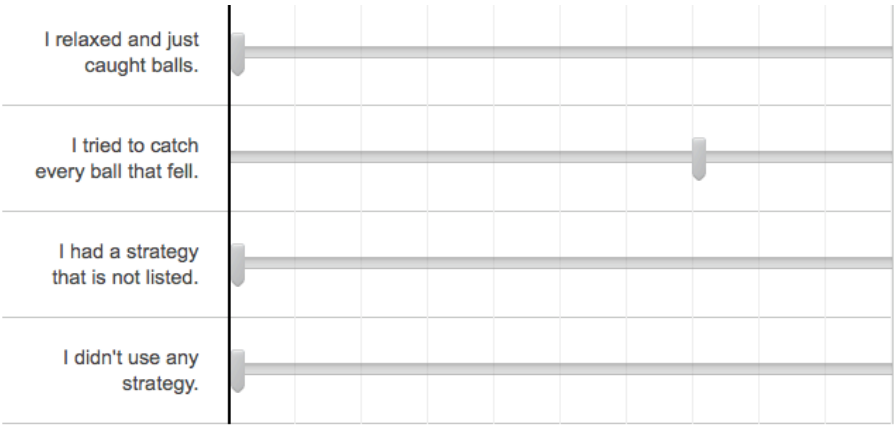

I am motivated to do the task.
I'm trying to figure myself out.
I'm reflecting about myself.
I'm daydreaming about myself.
I feel confident about my abilities.
I feel self-conscious.
I am worried about what other people think of me.
I feel concerned about the impression I am making.
I expect to perform proficiently on this task.
Generally, I feel in control of things.
I thought about how others have done on this task.
I thought about how I would feel if I were told how I performed.

APPENDIX E

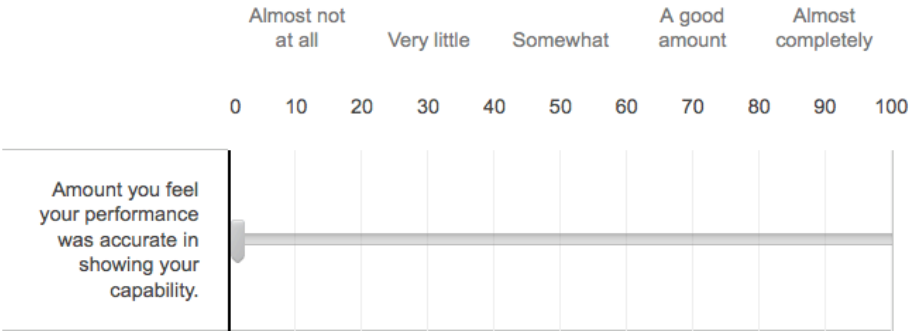
Exit Interview

Did you use any of these strategies to play the game?





Do you feel like your performance on the video game showed your capability?



What did you think about during the experiment?

The contrast changed throughout the task.

- True
- False

The speed at which the balls fell changed throughout the task.

- True
- False

Did you ever give up during the task?

- Yes
- No

Describe a point in the task when you gave up and why.

Did you care about your performance during this task?

	Almost not at all		Very little		Somewhat		A good amount		Almost completely	
	0	10	20	30	40	50	60	70	80	90	100
Amount of care you had toward your performance.											

About how long do you feel like it took you to complete this experiment?

Please answer the following statements:

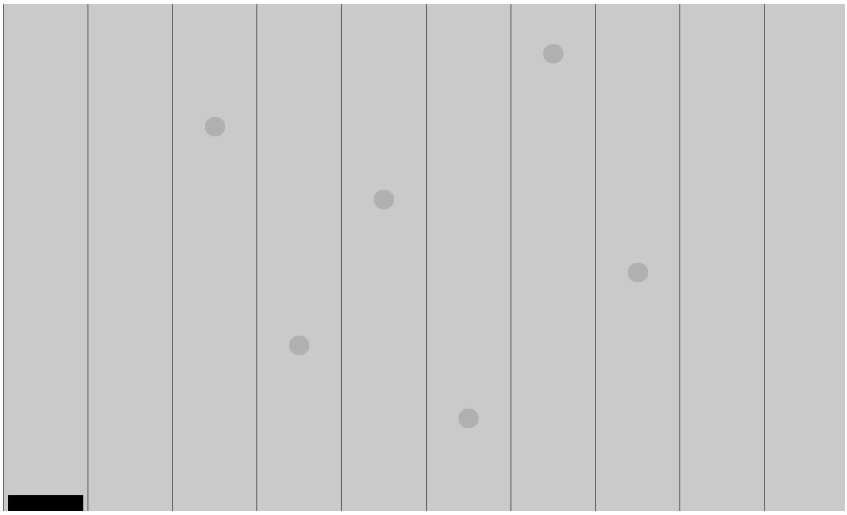
	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I was challenged, but I believed my skills would allow me to meet the challenge.	○	○	○	○	○
My abilities matched the high challenge of the situation.	○	○	○	○	○
I felt I was competent enough to meet the high demands of the situation.	○	○	○	○	○
The challenge and my skills were at an equally high level.	○	○	○	○	○

>>

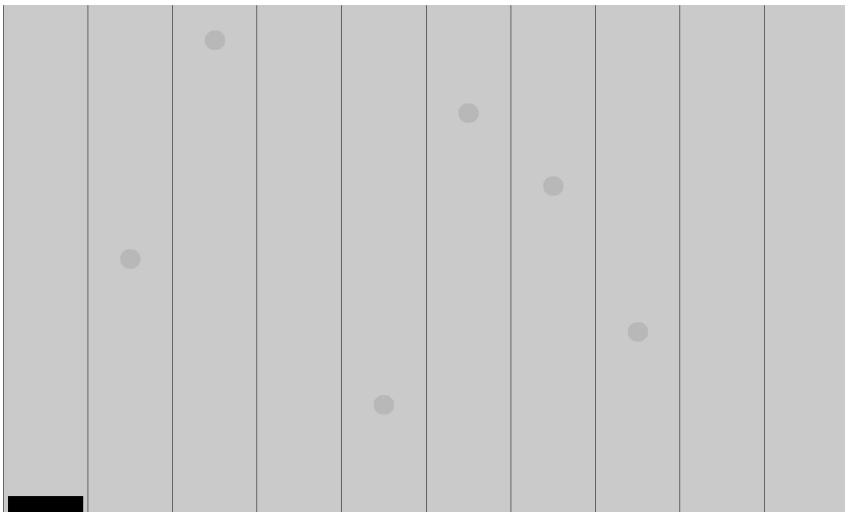
APPENDIX F

Experimental Trials

Practice Trial 1

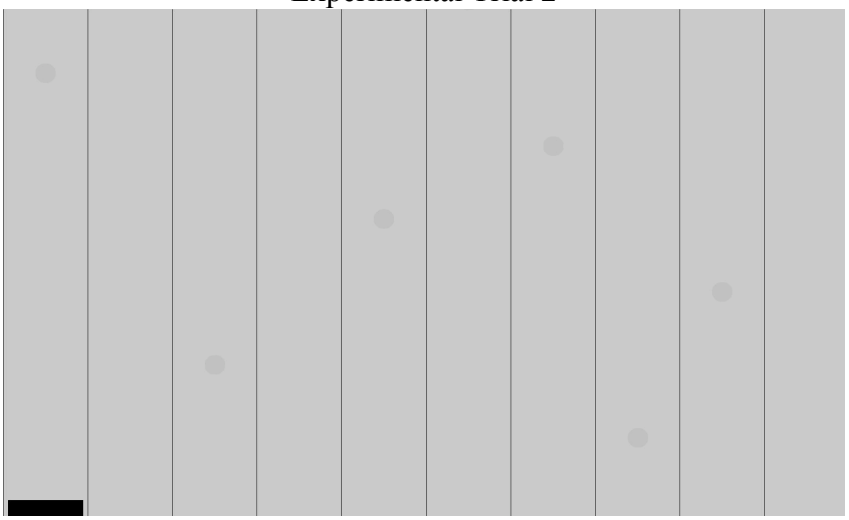


Practice Trial 2

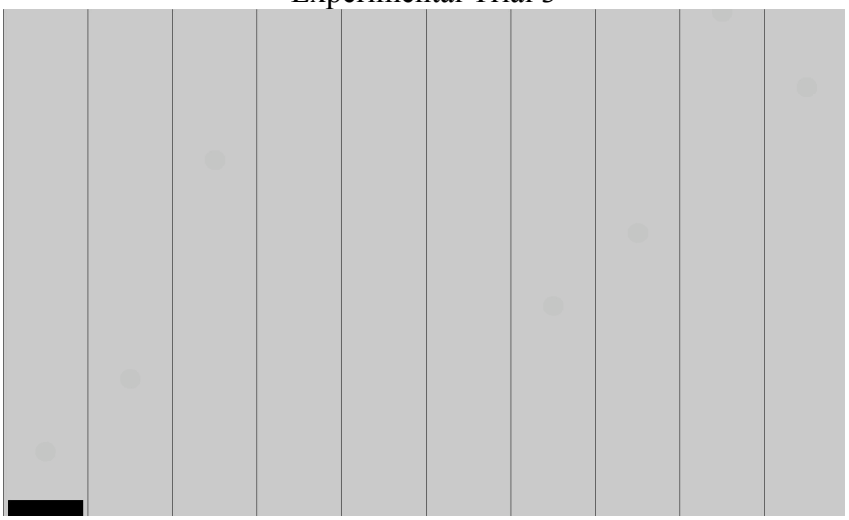


Experimental Trial 1

Experimental Trial 2



Experimental Trial 3



Experimental Trial 4

APPENDIX G

SSSQ Subscales Output by Age and Trial

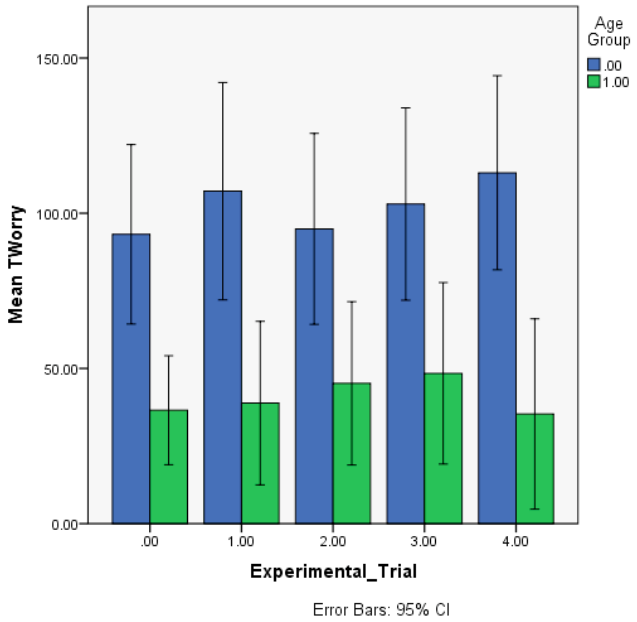


Figure 10. Mean total worry across experimental trial by age group.

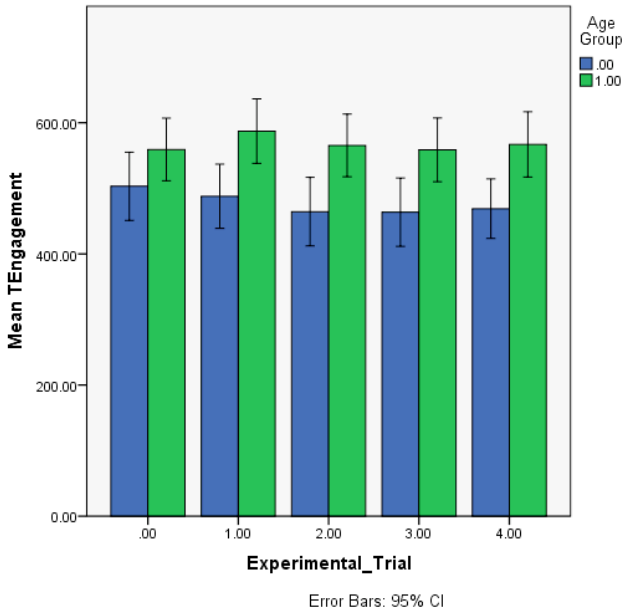


Figure 11. Mean total engagement across experimental trial by age group.

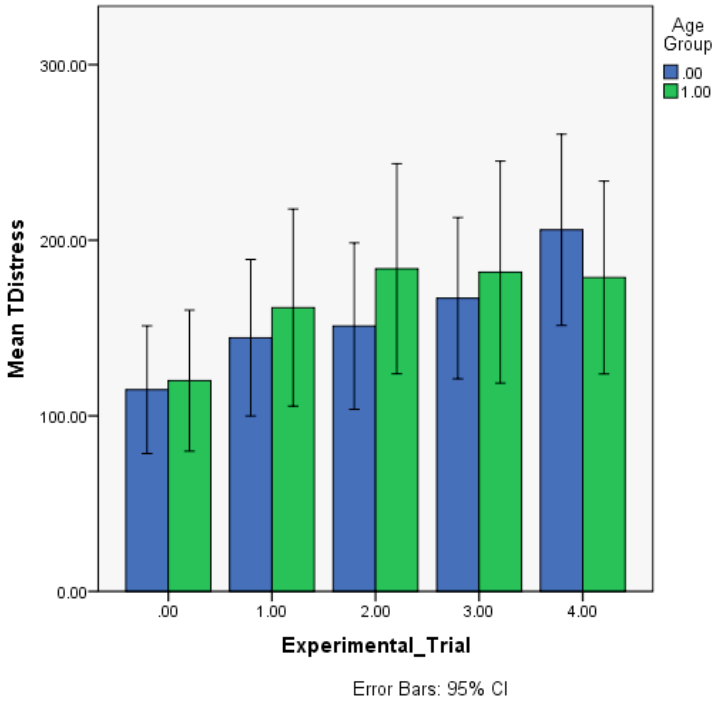


Figure 12. Mean total distress across experimental trial by age group.

APPENDIX H

Output of NASA-TLX by Age and Trial

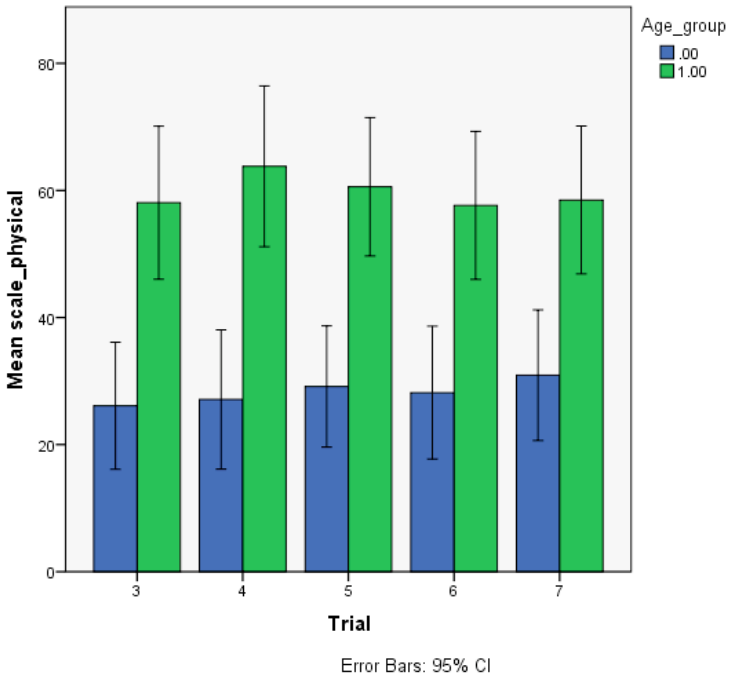


Figure 13. Mean total physical workload across experimental trial by age group.

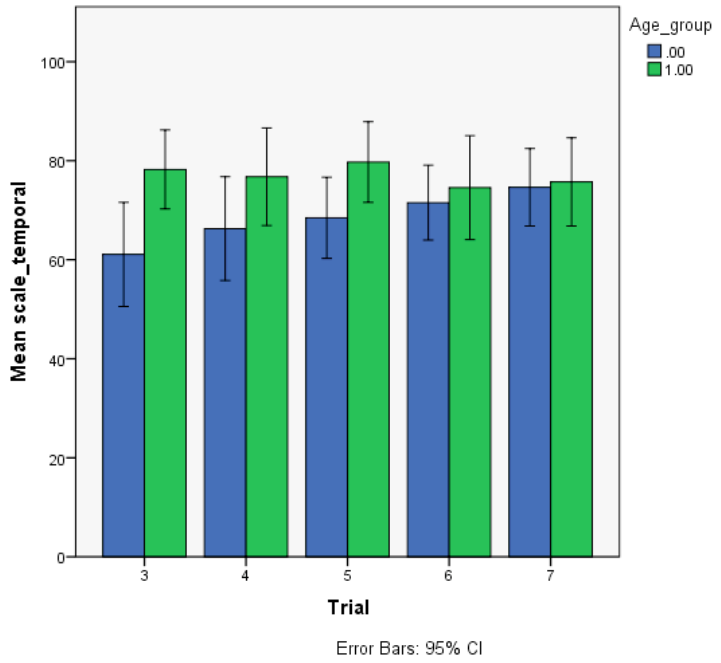


Figure 14. Mean total temporal workload across experimental trial by age group.

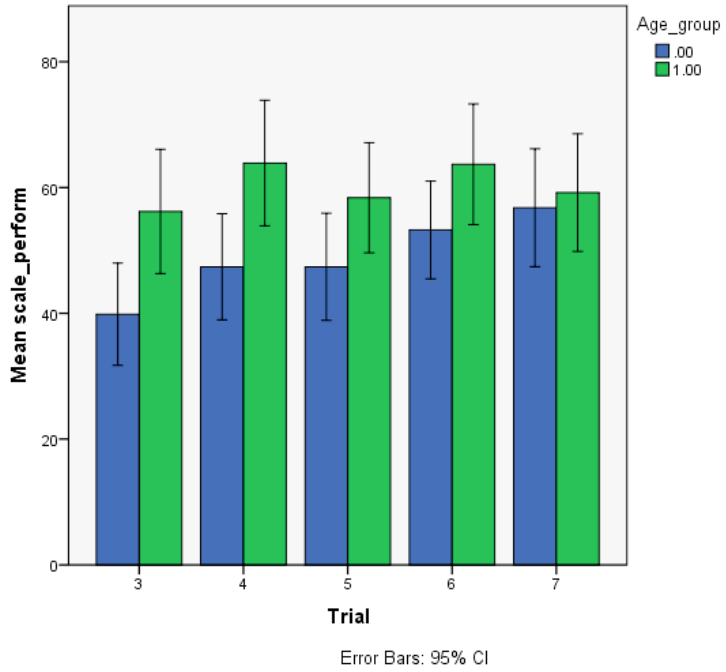


Figure 15. Mean total performance workload across experimental trial by age group.

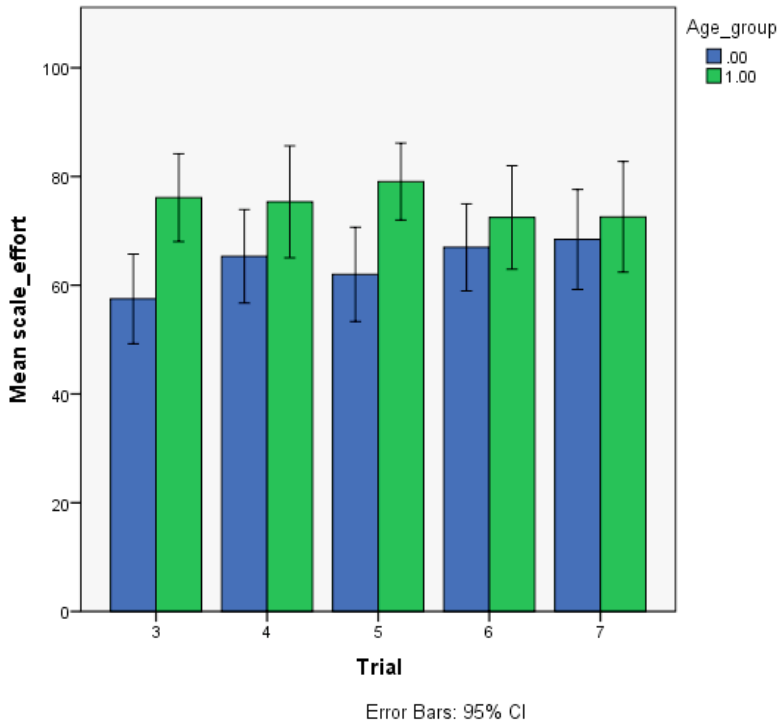


Figure 16. Mean total effort workload across experimental trial for each age group.

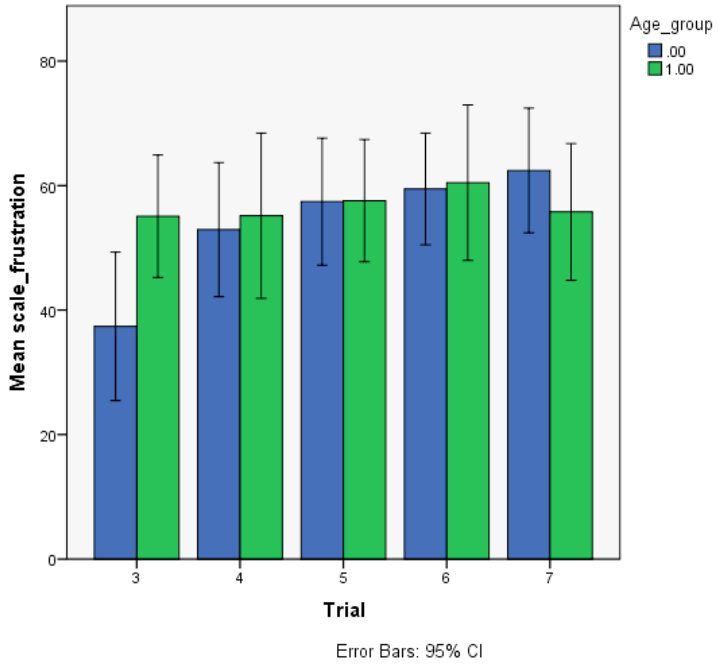


Figure 17. Mean total frustration workload across experimental trial for each age group.