THE EFFECTS OF MENTAL AND PHYSICAL LOAD IN EXERGAMES ON GAMEPLAY EXPERIENCES

By

Irem Gokce Yildirim

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Media and Information – Master of Arts

2017

ABSTRACT

THE EFFECTS OF MENTAL AND PHYSICAL LOAD IN EXERGAMES ON GAMEPLAY EXPERIENCES

By

Irem Gokce Yildirim

The purpose of this study was to observe how the mental load and physical load affect players' task load, enjoyment, flow, perceived competence, effort, performance and future play preference in exergames. Moreover, the interaction between mental and physical load for their effects on these variables was investigated. In a 3 (Mental Load: low-med-high) x 3 (Physical Load: low-med-high) between subject experimental setting, participants played an exergame including a "multiple object tracking" task. Mental load was manipulated by the number of target objects in the gameplay screen to be tracked and physical load was manipulated by the tension level of the exercise bike. Results showed that an increase in mental and physical load resulted in a significant increase in perceived task load and decrease in performance while performance also mediated these relationships. Moreover, a main effect of mental load on perceived effort was revealed. Although there were no significant differences in motivational experiences, such as enjoyment, flow and competence, between conditions with different levels of workloads, it was found that as mental and physical load increases, players' experience of enjoyment, flow and competence decreases. As being the first experimental study applying dual-task approach into exergames using subjective workload assessments, this study helps pave the way for future exergame research and design for better gameplay experiences by explaining the interactions in exergames.

To my mother (Meziyet AYDIN) and father (Selamettin AYDIN)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

With the increase in the level of capabilities and interactions provided by the technological advancements, people are now expected to perform complex tasks which require both mental and physical resources. While subjective workload (or task load) assessments are generally used for mental or physical activities separately in the domains of effective learning (Kyndt et al., 2011b; Kooiman and Sheehan, 2014; Kyndt et al., 2011a; Pimenta et al., 2015; Rubio-Valdehita et al., 2014) and health promotion (Jakobsen et al., 2014; Maillot et al., 2012; Sun, 2013; Staiano and Calvert, 2011), the overall task load experience during a dual-task performance has not been focused in the previous research (DiDomenico and Nussbaum, 2008).

The popularity of exercising devices as a solution for sedentary behavior in our daily lives has made it essential to find engaging and effective ways of workouts (Owen et al., 2010). Exergames have been proposed as an innovative solution to increase physical activity (Sween et al., 2014) by adding the layer of physical demand to the design of video games in addition to the mental demand they already require (Kalyuga and Plass, 2009). However, the composition of these demands may turn a game into a complex activity for players. There is a lack of research on the effects of both mental and physical demands on gameplay experiences. Since the level of these demands may influence the overall experience and should be balanced to maximize outcomes (Kiili and Perttula, 2013, p. 140), it is very important to assess the subjective workload for mental and physical demands to make exercising fun. To achieve this goal, the following questions are needed to be answered for exergames. What are the contributions of mental and physical load to the overall task load experience? How do mental and physical challenges influence enjoyment, flow, perceived competence and effort? Which one is more effective in predicting performance? And lastly, is there any interaction between mental and physical load for an optimum gameplay experience? Therefore, the purpose of this study is to investigate the relationship between mental and physical load, and their effects on taskload, enjoyment, flow, competence, effort and performance.

CHAPTER 2

LITERATURE REVIEW

2.1 Mental Load in Games

Cognitive Load Theory asserts that cognitive load is the mental energy that is needed to process a given information (as cited in Ang et al., 2007). Mental load is "the aspect of cognitive load that originates from the interaction between task and subject characteristics" (Paas et al., 2003, p. 64). According to Paas and Van Merriënboer (1994), mental load is based on individuals' characteristics and knowledge about the given task. Moreover, cognitive overload may occur when the amount of resources to process the information exceeds one's capacity (Kalyuga et al., 1999). Since "games usually require simultaneous performances of several cognitive and motor activities" (Kalyuga and Plass, 2009, p. 720), people may feel overloaded with the high level of mental demand that a game requires. For example, in previous research, the cognitive resources in games were utilized (e.g. levels of element interactivity, redundant information) for learning and performance by preventing cognitive overload (Annetta et al., 2009; Huang and Johnson, 2009; O'Neil et al., 2005). Therefore, three conditions with different levels of mental load manipulations (low-medium-high) were used in the experiment to examine their effects on the outcomes for this study. To quantify the impact of mental load on the subjects, mental workload, which is the "cost incurred by an individual, given their capacities, while achieving a particular level of performance on a task with specific demands" (as cited in Hart and Staveland, 1988, p. 977), is assessed using subjective ratings (Brunken et al., 2003). Mental demand dimension of NASA Task Load Index (NASA-TLX) is one of these subjective measurement instruments for perceived mental load (Hart and Staveland, 1988), and it was used for this study.

Visual elements as sensory inputs require cognitive processing. Games require visuaspatial skills along with other cognitive executive functions such as multiple object tracking, switching or updating skills (as cited in Nankar, 2016, p. 90). Rooted from Pylyshyn (2004)' theory of multiple object tracking, Allen et al. (2006) expands this by stating that multiple object tracking activity causes cognitive load. Franconeri et al. (2010) showed that the number of objects, the speed of tracking, and closeness of the objects are some determinants of the level of mental load. Similarly, Sweller (1994) claims that the number of elements is one of the main contributing factors for cognitive load and Allen et al. (2006) found that as the number of targets increased, subjects' performances decreased with increased cognitive load. Moreover, processing 7±2 chunks items is suggested as the limited capacity of working memory (Miller, 1956), identical objects are harder to track (Botterill et al., 2011), and "working memory is the key resource in multitasking" (as cited in Ang et al., 2007, p. 170). Therefore, in this study, multiple object tracking task with identical features is used to manipulate mental load as used in other previous studies (Pylyshyn, 2004; Hardy et al., 2015).

Based on the previous research, mental load manipulation has an impact on the overall task load ratings. DiDomenico and Nussbaum (2008)'s study showed that as the mental load increased the overall task load (measured by NASA-TLX) ratings also increased. It is reasonable that since the NASA-TLX measurement includes "mental demand" dimension, there is a contribution by assessment of mental load to the overall task load (H1a). Moreover, since effort consists of mental workload (Wickens, 2017) that needs to be exerted to perform a particular task in addition to physical workload, perceived effort is expected to increase with higher levels of perceived mental load (H1b). Research on the direct relationship between mental load and enjoyment while playing video games is limited. Ang et al. (2007) suggests that with a decrease in performance due to the cognitive overload, enjoyment may be retrieved leaving frustration in its place instead (H1c). Moreover, it was found that decreasing performance (number of failures) may increase perceived task load (Hancock, 1989). Therefore, there might be a mediation effect of performance between mental load and overall task load experience (H2). As flow occurs when a player's skills are balanced with challenges, with an increasing mental load, flow decreases (Tozman et al., 2015; Qin et al., 2009) (H1d). Mental load is also measures of task difficulty (Xie et al., 2016) and competence is based on the optimal level of the game's difficulty (Constant et al., 2017); therefore, higher mental load is expected to decrease competence (H1e). Considering the positive relationship between enjoyment and future play preference (Ryan et al., 2006), it is suggested that decreasing enjoyment with an increase in mental load will result in decrease in future play preference (H1f). It was also found that mental load can be assessed by performance (Paas and Van Merriënboer, 1994; Paas et al., 2003; Allen et al., 2006). Due to the negative relationship between mental load and performance, an increase in mental workload may cause failures (Paas and Van Merriënboer, 1993; Ang et al., 2007) (H1g).

Based on these findings, the following hypotheses are proposed:

H1: Higher mental load will result in higher a) perceived task load, b) effort; whereas, it will result in lower in c) enjoyment, d) flow, e) competence, f) future game preference and g) performance in the exergame.

H2: Performance will mediate the relationship between mental load and overall task load in the exergame.

2.2 Physical Load in Exergames

Physical load is "factors relating to biomechanical forces generated in the body" (Wahlström, 2005, p. 168). Perceived exertion, which is the "subjective intensity of effort, strain, discomfort, and/or fatigue that is experienced during physical exercise" (Robertson and Noble, 1997, p. 407), is one of the measurements of physical load in addition to the objective physiological determinants (e.g. heart rate, energy expenditure based on oxygen expenditure) (Capodaglio, 2001). As Borg (1972) found that the subjective ratings of perceived exertion are positively correlated with physiological responses (as cited in Hutchinson and Tenenbaum, 2006, p. 466) and perceived physical and mental load is complementary to the physiological measurements (Borg, 1990, p. 56), BORG's Scale of Perceived Exertion (Borg, 1998) as a subjective assessment tool was used to measure physical load in this study.

Exergames, which are the video games requiring physical interactivity to function, are broadly used by people from any ages. In these games, exercising activity is added as another layer to the top of mental activity that a video game already demands and a strong correlation between exergaming and increasing energy expenditure has been revealed in the previous studies (Sween et al., 2014). Although exergames related studies are mainly focused on health and education domains, the entertaining feature of these games is still prominent and crucial to promote physical activities for people living in a sedentary lifestyle (Lai et al., 2012).

Although there is a little research in the literature about the relationship between players' interaction with exergames and their experience of positive moods, some of the previous studies showed that as physical workload increases, overall perceived taskload increases (DiDomenico and Nussbaum, 2008) (H3a), and perceived effort increases since "workload is an indicator of the mental and/or physical effort required to carry out one or more tasks at a specific performance level" (as cited in Arroyo-Gómez et al., 2017) (H3b). Although there is no study which empirically tested the effects of different levels of physical loads on enjoyment in an exergame, based on the balanced level of challenges for an optimum experience approach in Flow theory (Sinclair et al., 2009; Qin et al., 2009; Lai et al., 2012), high physical load is predicted to result in less enjoyment and flow (H3c, H3d). In their study, Park et al. (2014) suggest that higher level of physical workload will deteriorate perceived competence and performance (H3e, H3g). Due to the positive correlation between enjoyment and future play preference (Ryan et al., 2006), it is predicted that an increase in physical load will reduce the future play preference for the game (H3f).

Based on the evidences and interpretations above, we propose the following hypotheses:

H3: Higher physical load will result in higher a) perceived task load, b) effort; whereas, it will result in lower in c) enjoyment, d) flow, e) competence, f) future game preference and g) performance in the exergame.

H4: Performance will mediate the relationship between physical load and overall task load in the exergame.

Park et al. (2013) found that lower perceived taskload has a positive effect on perceived competence and involvement. Since we used immersion subscale from the GameFlow measure, all the items in our survey were related to involvement. Therefore, we argue that task load will also mediate the relationship between perceived mental and physical load and their effects on competence, performance and flow experiences in our study (H5). Moreover, since "when in flow, a player does not have to invest effort to keep his mind on the task" (Kiili et al., 2013, p. 202), an argument can be made that flow will mediate the relationship between perceived effort and mental load as well as physical load(H6).

H5: Perceived taskload will mediate the relationship between competence and mental and physical load, respectively, as well as the relationship between flow and mental and physical load, respectively, in the exergame.

H6: Flow will mediate the relationship between mental and physical load separately with perceived effort in the exergame.

2.3 Dual-Task in Exergames

According to dual-task methodology, cognitive resources are split for concurrent tasks and attention switching occurs Damos (1991). For example, when a cognitive performance is accompanied by physical exercising, it was found that dual-task effect was significantly related to energy demand by the task, and more attentional resources is necessary to control movements (Brisswalter et al., 2002, p. 555). Similarly, research on executive functions showed that cognitive skills required to plan, monitor and execute actions are associated with physical performance (as cited in Coppin et al., 2006). Although in some health promotion related studies dual-task interventions have been used to ameliorate age related physical and cognitive declines (Nankar, 2016), they may also result in cognitive overload as Strobach et al. (2012) stated that dual tasks require coordination of actions and shifting between these actions. Shifting is one of executive functions, which refers to “switching between attentional sets or task sets” (Van Muijden et al., 2012), and it may affect dual-task performance (Miyake et al., 2000) and increase cognitive load (Ang et al., 2007).

In addition to the physical activity, exergames require cognitive functioning such as sensing stimuli and paying attention (Larsen et al., 2013, p. 6). The mental and physical demand required in exergames "make players experience divided attention and impose continuous demands on the attention, potentially resulting in dual-task interference" (Park et al., 2014, p. 419). The integration of physical and cognitive demands into exergames requires specific attention to their implementations (Lyons, 2015). How some game mechanics (feedback, challenges and rewards) are operationalized can affect physical and psychological reactions to exergames. The "interplay between cognitive load and physical load" (Pisan et al., 2013) is necessary for expected outcomes. Although the dynamics in the intervention of mental effort to physical exertion have been studied in many previous research, there is no exergame focused study observing the interactions between these in a concurrent setting using dual-task(multi-task) approach for gameplay experience related outcomes. For example, the effect of physical exercise on cognitive (or mental) performance was observed during or after (not a dual-task design) in many empirical studies and some contradictory results were revealed (Brisswalter et al., 2002). While in some studies it was found that physical exercise improved cognitive performance, some other showed that physical exertion deteriorated the performance in cognitive tasks (as cited in Krausman et al., 2002, p. 4). In another study, Davey (1973) found an inverted U relationship between the level of physical exertion and after mental performance of short term memory. Therefore, in this study, mental and physical load are manipulated in three different levels for a possible curvilinear relationship with gameplay experiences. These results led researchers to highlight the importance of the type of the cognitive task (e.g. visual perception, arithmetic, decision making), the duration and intensity of the physical exertion on the performance results (Brisswalter et al., 2002; Krausman et al., 2002) and timing of physical exertion and cognitive task (concurrent/dual-task or not) Brisswalter et al. (2002). Moreover, Kiili and Perttula (2013) proposed a framework to balance these two loads for more engaging exertion and learning experiences.

Based on dual-task methodology[1], to contribute to the literature in the field of exergames, in this study the interaction between mental and physical load is observed using the subjective assessment tools during. There are only a few exergame related studies for observing gameplay

[1] Although dual-task approach was included in this thesis to explain the dynamics in concurrent activities, the exergame that was used in this study might not necessarily be defined as a dual-task but rather one multifaceted task which requires both physical and cognitive resources.

experiences using many different cognitive load manipulation tasks (i.e. choice reaction test, short-term memory/retention test and arithmetic test) and physical load manipulation methods (e.g. target speed) (as cited in Park et al., 2014). In their studies, Hardy et al. (2015) and Park et al. (2014) utilized tension level to manipulate physical load and in this study we used the same method to change physical load for three different levels in three conditions. Considering the contribution of multiple-object tracking task to the perceived mental load as explained in the previous section, it is used to manipulate mental load in three conditions with different levels.

Lastly, the following research question is proposed for this study to explore the question of dual-task interactions in exergames:

RQ1: Will there be any interaction effect of mental and physical load on a) overall task load, b) enjoyment, c) flow, d) competence, e) effort, f) future game preference and g) performance in the exergame?

In this research, the goal is to explore the effects of different levels of mental and physical load on perceived task load, flow, enjoyment, effort, competence, preference for future play and performance, hoping to find an optimum levels of mental and physical load for an optimal gameplay experience. This study may help researchers better understand how different levels of physical and mental load in exergames affect users' interactions. It will also help pave the way for future exergame research around the concepts of personalized and adaptive exercising experiences (Figure 2.1).

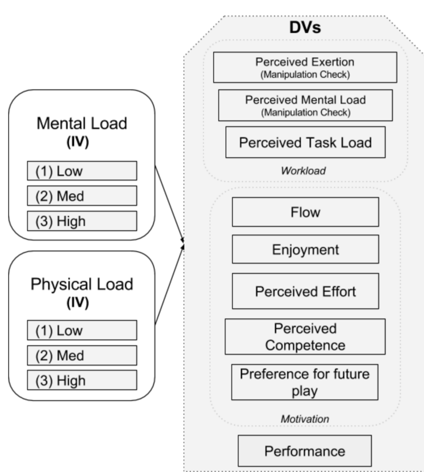

Figure 2.1: Proposed research framework.

CHAPTER 3

METHODS

3.1 Participants

Undergraduate and graduate students from a large Midwest university in the United States were asked to participate in the study. As the game used in the experiment requires physical activity, a screening criteria of physical activity readiness was applied. In total, 114 participants (46 male; 68 female) were eligible to participate in the study. The average age of the participants was 21.

3.2 Stimuli

The target game for this study was an exergame entitled *CogXer*. A 27in *iMac* is used as computer screen for the game and an *FitDesk Pedal Desk* exercise bike was connected to the iMac through an Arduino sensor tracking the speed information. Audio was delivered via built-in speakers on the iMac (Figure 3.1). This exergame was designed following the feedback and challenge design suggestions posit by Lyons (2015).

Figure 3.1: Screenshot of the intro scene in the target game.

The game was designed to provide two main challenges; mental and physical load. Multiple object tracking task was applied in the game to promote cognitive challenge whereas the player controls the avatar in the game by cycling as the physical challenge. This target game was tested with one year of playtesting by undergraduate and graduate students from the target audience and the design was improved based on the feedback collected.

In the game, multiple object tracking task was simulated as an implementation of visual tracking, searching and recalling actions. In a zoo like environment, turtles are used as objects to track and the goal in the game is feeding as many turtles as possible. Player has baits the number of which is equal to the number of all turtles in the screen initially. However, the mental challenge about feeding is that since all the turtles are identical and wandering around, player needs to remember and visually track which turtle she or he already fed to reserve the rest of baits for the unfed ones. This mental challenge was manipulated in the game by changing the "number" of target objects in the game (i.e. decreasing or increasing the number of turtles). To support the feeling of competence, timer mechanics is also implemented. For each turtle, the participant is given 10 seconds to feed it (e.g. if there is 5 turtles in a round, player has 50 seconds to feed all the turtles). In the game, player controls an avatar of an unicycle. The speed information was streamed from an exerbike through an Arduino based connection to the game by mapping the speed data into the position of the unicycle. The physical challenge about controlling of the unicycle is that based on the tension level set on the exerbike, the difficulty level of cycling can be manipulated which in turns makes it harder for player to control the avatar in the game (Figure 3.2).

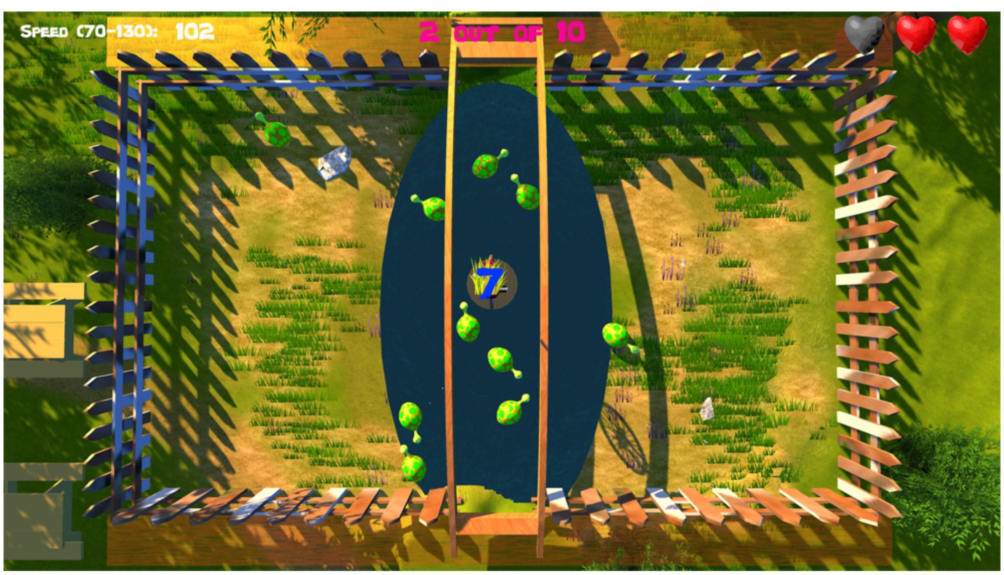

Figure 3.2: Screenshot of the main gameplay scene.

3.3 Procedure

Before coming to the lab, participants completed a screening survey about their eligibility to perform a physical activity and interest to play a cognitive exergame. Participants who indicated that they had previous injury records or reasons not to do exercise using the Physical Activity Readiness Questionnaire (PAR-Q) in the screening survey were not allowed to participate in the study. Eligible participants scheduled their visit to the lab using the college's participation pool system.

Participants were randomly assigned to one of the nine conditions in this study: a 3 (mental load: low-medium-high) x 3 (physical load: low-medium-high) between subjects design. Upon arrival to the lab, participants were introduced to the procedure using a well-documented protocol script and signed the consent form. There were three main sessions in the experiment. The first two session were the 'calibration' sessions to identify participants' maximum physical and mental performance levels to gauge their baseline to set the difficulty level variables for the actual gameplay (Table 3.1). Calibration was needed because each individual may have different levels of endurance for the mental and physical load presented in a given task. Since low-med-high values can be different for each participant in this type of mental and physical challenge setting, instead of using a fixed number of target game objects and tension level on bike for each condition, rather, we identified the max values in calibration sessions and setting the number of objects and tension level for the med and high conditions based on that max value to eliminate the individual differences in their respose to different mental and physical demands (DiDomenico and Nussbaum, 2008, p. 978). Following Ryu and Myung (2005)'s suggestion of "in a multitask condition, the task should be broken up into its constituent elements and the demands of each element must be evaluated independently", in these calibration sessions, a slightly different versions of the same target game, CogXer, were used; one without physical challenge, one without mental challenge. The last session was the actual game play which lasted around 5 minutes for each player (participants were let to stop playing if there were tired or exhausted) and it was followed by an online post-test questionnaire.

Table 3.1: Manipulations of Mental and Physical Load for Three Different Levels in the Game

IVs	Levels	Manipulation	Manipulation check item
Mental Load	Low	# of turtles = Min (1)	NASA-TLX1: How mentally demanding was the task?
	Medium	# of turtles = Mean (1, max)	
	High	# of turtles = Max (based on the mental calibration)	
Physical Load	Low	Tension level on bike = Min (1)	BORG: Choose the number that best describes your level of exertion.
	Medium	Tension level on bike = Mean (1, max)	
	High	Tension level on bike = Max (based on the physical calibration)	

- In the mental challenge calibration phase, participants played the game using keyboard (without physical load - no exercise device), while gradually increasing the mental difficulty of game by increasing the number of objects to track on the screen by 2 and giving 10 seconds for each object after each successful round (starting from 2) over time (Figure 3.3). We provided three 'lives' and three replay chance for every round (round refers to gameplay with the same number of turtles), and reduced one for every 'UNSUCCESSFUL' feeding, including: double-feeding and mis-feeding. Therefore, successful round referred to finishing a round with at least one life left (in other words, if there is N turtle in that round, finishing it with at least N-2 successfully fed turtles is successful). Between each increment interval (no rating between consecutive replay trials), participants were asked to rate perceived mental demand to measure mental load (Kalyuga & Plass, 2009; Schrader; as cited in Schrader & Bastiaens, 2012) using the 7-point Likert question from NASA-TLX popping up on the window.

  Maximum mental load for a participant was decided and so the mental calibration stops when one of following conditions happens:

  - If a player cleared the round but rate the round as '7'. (SUBJECTIVE)
  - Participant pressed Give Up button on the screen when s(he) felt it's maximum limit for him or her (after they press give up button, they were asked to rate for the last time. calibration stops no matter they rate it with less than '7').(SUBJECTIVE)
  - If a player failed three replay chance of the round, cognitive calibration stopped there (no matter they rate it with less than '7'). (OBJECTIVE, performance can be used as an indicator of perceived mental load (Paas and Van Merriënboer, 1994; Qin et al., 2009))

When calibration stopped, the number of turtles in the last played round was noted as ‘max’ mental load for this participant. In this way, mental load measurement was mixed with subjective rating and objective performance (Brunken, Plass, & Leutner, 2003) during manipulation in this calibration phase.

Figure 3.3: Screenshot of the gameplay scene in the cognitive calibration session.

- In the physical challenge calibration phase, participants played the same game (without mental load - no object tracking task), but they were asked to just cycle at certain speed (e.g., 100 RPM). The tension level of the exerbike was gradually increased by one level in every 30 seconds (levels are between 1-to-8) (Figure 3.4). Between each increment interval, participants rated perceived physical demand using the BORG Rating of Perceived Exertion (scores were between 6-20). Physical calibration was stopped when participants gave up and identified a tension level as their "max"[1]. This level was noted as ‘max’ physical load for this participant. Then, participants were asked to rest for 5 minutes before the actual gameplay session (Figure 3.5).

[1] During the data collection process in Spring 2017 term, if the participant rated 16 or more in the manipulation check item (BORG scale), we stopped the calibration and took the tension level that he or she rated 16 or less as the max value for that participant. Based on the pilot test results, to increase the difference in manipulation check item between low and medium physical load conditions, we decided to let the participant go as far as they can for the further data collection process which includes Fall 2017 term. Therefore, ’data set’ is used as a covariate in the further analysis in this thesis.

Figure 3.4: Stationary exercise bike (with adjustable tension level) used in the study connected to the computer.

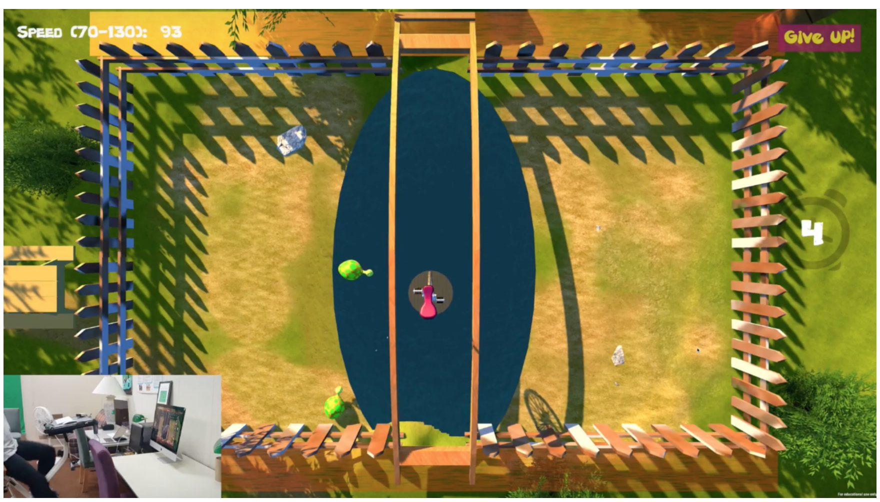

Figure 3.5: Screenshot of the gameplay scene in the physical calibration session.

After the calibration sessions (see Appendix L), the actual gameplay session was conducted. Based on the assigned condition and the maximum values marked during the calibration sessions, mental load was manipulated by the number of target game object in the game to be tracked, and physical load was manipulated by the resistance/tension of the exercise bike. Physical and mental load for low-med-high conditions were set based on the min-mean-max values. "Min" values for low conditions were always fixed as '1' for cognitive load and '1' for physical load. "Mean" and "Max" are set based on the maximum value identified in the calibration phases (Table 3.1):

- If the participant is in high condition, we set the number of turtles and tension level variables to the max value revealed in the calibration sessions,
- If the participant is in med condition, we set the number of turtles and tension level variables to the mean of max value revealed in the calibration sessions and fixed min values,
- If the participant is in low condition, we set the turtles and tension level variables to fixed min values which is '1' as number of target objects for mental load[2] and '1' as tension level for physical load.

In this way, participants in med and high conditions were assigned values based on their personal med and max values that they could endure.

After setting the manipulation variables for the assigned condition, participant played the game with the same number of turtles and same tension level over a series of rounds for about 5 minutes and they were let to stop playing anytime due to fatigue. During this session, gameplay data was recorded in a log file. After playing, participants completed a post survey including questions about their gameplay experiences. The entire procedure took around 45 to 60 minutes to complete. Participants received partial course credits for their participation.

3.4 Measures

3.4.1 Manipulation Check

Since the Mental Demand dimension from The National Aeronautical and Space Administration Task Load Index (NASA-TLX) and BORG Perceived Exertion Scale scale were used during the

[2] During the data collection process in Spring 2017 term, we used '2' as the number of turtles for low mental load conditions. Based on the pilot test results, to increase the difference in manipulation check item between low and medium mental load conditions, we decided to set it to '1' for the further data collection process which includes Fall 2017 term. Therefore, 'data set' is used as a covariate in the further analysis in this thesis.

calibration phases to determine the maximum mental and physical load respectively for each participant, they were used as manipulation check items in the post-survey.

Mental Demand dimension item from NASA-TLX ("How mentally demanding was the task?") was used as an indicator whether the mental load manipulation was successful (see Appendix C). It is an 7-point scale anchored by 1 (*very low*) and 7 (*very high*). In the mental load calibration session, the question of "How much difficult was remembering, tracking and searching?" was used to elaborate this item for participants.

BORG scale (Borg, 1998) was used as an indicator whether the physical load manipulation was successful (see Appendix D). It is an numerical scale ranging from 6 to 20 with several short descriptions provided to represent the level of exertion (e.g. "No exertion at all", "Extremely light", "Very light", "Somewhat hard") (see Appendix D).

3.4.2 Dependent Variables

Overall task load was measured by 7-point scale NASA-TLX (see Appendix C). This scale has six dimensions-mental demand, physical demand, temporal demand, performance, effort and frustration (Hart and Staveland, 1988). Although physical demand dimension is included in this measure, this dimension is not intended to be used to assess perceived physical load. Cronbach's Alpha of the scale was above .7 and reliability was secured (α = .71).

Game enjoyment was measured using the scale developed by Song et al. (2011) (see Appendix E). It is an 7-point scale on some adjectives to describe the game including "boring"(R), "exciting", "enjoyable", "entertaining", "fun", "interesting", "pleasant" (α = .90).

To measure flow, Immersion dimension and two items from Challenge dimension of game flow scale developed by Fu et al. (2009) based on Sweetser and Wyeth (2005)'s framework were used (see Appendix F). Since the immersion dimension has the most factor loading, all the items from this dimension were included in the scale. Sample items includes "I forget about time passing while playing the game","I can become involved in the game" . Moreover, as the balance between skills and challenges is the key to experience flow (Csikszentmihalyi, 1990) and players feel bored or frustrated when they are out of the flow zone (Sweetser and Wyeth, 2005), two items from Challenge dimension "I enjoy the game without feeling bored or anxious" and "The challenge is adequate, neither too difficult nor too easy" were also included to measure flow in this study (α = .84).

Competence and effort were measured using the Perceived Competence and Effort measures (McAuley et al., 1989) from Intrinsic Motivation Inventory (IMI) (α = .88 and α = .77, respectively) (see Appendix G and Appendix H). Although effort dimension is included in NASA-TLX measure, that dimension is not intended to be used to assess perceived effort.

Future play preference measure included two items from Ryan et al. (2006) and additional two items including "Given the chance I would play this game in my free time" and "I would be interested in having my own personal copy of this game" (α = .97) (see Appendix I).

3.4.3 Mediators

Performance was proposed as potential mediator in this study. Using the gameplay log file recorded during actual gameplay session, performance was measured as the success in the visual tracking task which is successful feeding of the turtles in the game (see Appendix J). Since successful feeding requires a certain amount of physical effort (to control the position of the avatar in the game), we can assert that the performance measure is a representation of a mixture of mental and physical effort exerted by the player.

In addition, total game play time (active and resting), average time spent per round, average performance per round, average speed, average speed per round and minimum speed violation per round were also recorded in that log file (see Appendix K).

Gender, weekly gaming hours and weekly exercising hours were reported by the participants in the prescreening survey considering these as potential covariates (see Appendix A).

CHAPTER 4

RESULTS

4.1 Preliminary Analysis

We examined the relationship between gender, weekly gaming hours and weekly exercising hours with dependent variables to control in the subsequent analysis as potential covariates for homogeneity of regression. Gender was found to have main effects on only perceived exertion, competence and performance. Males perceived less exertion [$F(1{,}112) = 5.21$, $p < .05$], perceived more competence [$F(1{,}112) = 7.22$, $p < .01$] and performed better [$F(1{,}106) = 4.54$, $p < .05$] than did females. A Pearson correlation coefficient was computed to assess the relationship between weekly gaming hours and weekly exercising hours and dependent variables. There was significant negative correlation between weekly gaming hours and perceived exertion [$r = -0.331$, $n = 106$, $p = 0.001$] and it had a main effect on perceived exertion [$F(15, 90) = 1.96$, $p < .05$]. Moreover, a significant positive correlation was found between weekly gaming hours and competence [$r = 0.227$, $n = 106$, $p = 0.019$]. Weekly exercising hours variable was only negatively correlated with performance [$r = -0.218$, $n = 106$, $p = 0.029$].

In the subsequent analysis, gender, weekly gaming hours, weekly exercising hours and data set (Spring 2017 vs Fall 2017, see Section 3.3) variables were controlled. All the data were examined for homogeneity of variance using Levene Tests.

4.1.1 Manipulation Check

Two sets of one-way analysis of variance (ANOVA) tests were conducted to check the manipulation of the independent variables, mental load and physical load.

Shapiro-Wilk values as well as were obtained to observe the distributions within nine conditions for manipulation check items (NASA-TLX1:Perceived Mental Demand for mental load manipulation and BORG: Perceived Exertion for physical load manipulation). The distributions for perceived mental demand in the data of mental load low - physical load medium and mental load medium - physical load high conditions failed the assumption of normality. Similarly, the data in mental load high - physical load medium and mental load high - physical load high conditions failed the assumption of normality for perceived exertion. After applying log10 transformation for the corresponding manipulation item data, normality still failed in these conditions. NASA-TLX1 and BORG contained six outliers in total. To determine whether the outliers influenced manipulation results, the inferential statistical tests (ANOVA) were conducted twice for independent variables respectively, with outliers and without outliers.

With outliers, there was a significant effect of mental load on perceived mental demand for the three conditions [$F(2, 111) = 10.39, p < 0.001$]. A Bonferroni post hoc test revealed that compared to the high mental load condition ($M = 4.98, SD = 1.52$) the perceived mental demand was significantly lower in low ($M = 3.41, SD = 1.59, p < 0.001$) and medium mental load conditions ($M = 4.11, SD = 1.49, p = 0.034$). However, the mean scores were not significantly different between low and medium mental load conditions ($p = 0.158$). One-way ANOVA results for physical load showed that the means of the levels of the perceived exertion significantly different between three physical load conditions [$F(2, 111) = 37.64, p < 0.001$] (Low: $M = 12.66, SD = 2.35$; Medium: $M = 14.72, SD = 2.25$; High: $M = 16.95, SD = 2.87$) suggesting that the manipulation was successful (Table 4.1).

Table 4.1: One-Way ANOVA Results for Manipulation Check (with outliers)

Mental Load	Low (n = 37) *M (SD)*	Medium (n = 35) *M (SD)*	High (n = 42) *M (SD)*	F (2, 111)	η^2
NASA-TLX1 (Perceived Mental Demand) - How mentally demanding was the task? (i.e. How much difficult was remembering and tracking?)	3.41 (1.59)	4.11 (1.49)	4.98 (1.52)	10.39**	0.16
Physical Load	**Low (n = 38) *M (SD)***	**Medium (n = 39) *M (SD)***	**High (n = 37) *M (SD)***	**F (2, 111)**	**η^2**
BORG (Perceived Exertion) - Please choose the number from below that best describes your level of exertion.	12.66 (2.35)	14.72 (2.25)	16.95 (2.87)	37.64**	0.37

Note. $^{*}p$ < .05, $^{**}p$ < .01, $^{***}p$ < .001.

When we excluded the outliers from the analysis, there was a still significant effect of mental load on perceived mental demand for the three conditions [$F(2, 105) = 10.72, p < 0.001$]. However, Bonferroni post hoc test revealed that there was only a significant difference in perceived mental demand between low and high mental load conditions. Although, the difference approached significance between low-medium ($p = 0.099$) with the exclusion of outliers, it was non-significant between medium-high mental load conditions ($p = 0.051$). Without outliers, one-way ANOVA results for perceived exertion were nearly identical between three physical load conditions [$F(2, 105) = 28.43, p < 0.001$]. Pairwise Bonferroni post hoc comparisons still revealed significant differences for all three physical load conditions (Table 4.2).

Table 4.2: One-Way ANOVA Results for Manipulation Check (without outliers)

Mental Load	Low (n = 33) *M (SD)*	Medium (n = 35) *M (SD)*	High (n = 40) *M (SD)*	F (2, 105)	η^2
NASA-TLX1 (Perceived Mental Demand) - How mentally demanding was the task? (i.e. How much difficult was remembering and tracking?)	3.33 (1.41)	4.11 (1.49)	4.95 (1.62)	10.72**	0.17

Physical Load	Low (n = 35) *M (SD)*	Medium (n = 39) *M (SD)*	High (n = 34) *M (SD)*	F (2, 105)	η^2
BORG (Perceived Exertion) - Please choose the number from below that best describes your level of exertion.	12.86 (2.23)	14.72 (2.25)	17.00 (2.39)	28.43**	0.35

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Both one-way ANOVA analyses with and without outliers concluded there was no significant difference between average perceived mental demand scores for participants who were in low mental load condition and participants who were in medium mental load condition. Therefore, we decided to collapse low and medium mental load conditions into one "low" condition without removing the outliers ($M = 3.75, SD = 1.57$) and keep the high condition as it was for the subsequent analysis and applied t-test analysis on this new low and high conditions; $t(112) = 4.06, p < 0.001$. As a result, the design changed to 2 (mental load: low-high) x 3 (physical load: low-medium-high) from 3 x 3.

4.2 Primary Analysis

4.2.1 Main Effects on Dependent Variables

Three sets of two-way analysis of covariance (ANCOVA) were conducted to examine the effects of mental load and physical load on competence (controlling only for data set, after dropping other covariates due to their non-significant effects), effort (controlling only for data set, after dropping other covariates due to their non-significant effects) and performance (controlling only for weekly exercising hours, after dropping other covariates due to their non-significant effects). Considering the non-significance in their effects on the other dependent variables, all the covariates were dropped and a series of two-way analysis of variance (ANOVA) were conducted to examine the main effects of mental load (low-high) and physical load (low-med-high) on the other dependent variables and to test whether there were interactions.

Both mental load and physical load had statistically significant main effects on NASA-TLX task load ratings. As the mental load increases across low-high conditions, perceived task load increased, $F(1, 108) = 12.99, p < .001$. Similarly, higher physical load resulted in higher perceived task load $F(2, 108) = 5.34, p < .01$. The main effect of mental load on effort was also statistically significant $F(1, 107) = 4.24, p < .05$. Participants in high mental load condition reported higher effort than those in low mental load condition. Lastly, due to the unequal number of subjects across conditions and the violation of Levene's Test of homogeneity of variances revealed in the two-way ANCOVA analysis for performance dependent variable within each of the mental and physical load conditions, two sets of one-way between subjects ANOVA were conducted for each of the independent variables using Welch's test (with Games-Howell post-hoc comparison for three physical load conditions). Results showed that subjects in high mental load condition performed significantly better $F(1, 106) = 20.48, p < .001$ and subjects low physical load condition performed significantly better than those in high physical load condition $F(2, 105) = 8.12, p < .01$. However, no significance was found in the performance difference between low-medium and medium-high conditions (Table 4.3). Although the means are supporting the arguments in H1 and H3, only H1a,

Table 4.3: Two-Way ANOVA and Two-Way ANCOVA Results Showing the Main Effects of Mental and Physical Load on Dependent Variables

Dependent Variables	Mental Load: Low (n = 72) $M(SD)$		High (n = 42) $M(SD)$	$F(1, 108)$	η_p^2	p
NASA-TLX$^{two-way-anova}$	4.23 (0.96)		4.87 (0.90)	12.99**	0.11	0.000
Enjoyment$^{two-way-anova}$	4.81 (1.02)		4.74 (1.12)	0.05	0.000	0.82
Flow$^{two-way-anova}$	4.64 (1.07)		4.55 (0.95)	0.129	0.001	0.72
Competence$^{two-way-ancova}$	3.99 (1.25)		3.46 (1.19)	2.90	0.03	0.09
Effort$^{two-way-ancova}$	4.84 (0.96)		5.26 (0.81)	4.24*	0.05	0.042
Future Play Preference$^{two-way-anova}$	3.49 (1.77)		3.40 (1.73)	0.09	0.001	0.76
Performance$^{one-way-anova\ b1}$	0.59 (0.20)		0.43 (0.11)	20.48**	0.16	0.000
Dependent Variables	**Physical Load: Low (n = 38) $M(SD)$**	**Medium (n = 39) $M(SD)$**	**High (n = 37) $M(SD)$**	**$F(2, 108)$**	**η_p^2**	**p**
NASA-TLX$^{two-way-anova}$	4.15 (0.95)	4.46 (0.85)	4.81 (1.04)	5.34**	0.09	0.01
Enjoyment$^{two-way-anova}$	4.97 (1.05)	4.79 (0.99)	4.60 (1.11)	0.88	0.02	0.42
Flow$^{two-way-anova}$	4.89 (0.68)	4.66 (1.04)	4.26 (1.21)	2.07	0.04	0.13
Competence$^{two-way-ancova}$	3.97 (1.28)	3.82 (1.22)	3.59 (1.26)	1.25	0.003	0.30
Effort$^{two-way-ancova}$	4.87 (0.99)	5.02 (0.93)	5.10 (0.87)	0.91	0.01	0.41
Future Play Preference$^{two-way-anova}$	3.59 (1.64)	3.60 (1.82)	3.18 (1.78)	0.39	0.01	0.68
Performance$^{one-way-anova\ b2}$	0.58 (0.14)	0.58 (0.15)	0.43 (0.24)	8.12**	0.15	0.001

Note. $^*p < .05$, $^{**}p < .01$.
$^{two-way-anova}$ Two-way between subjects ANOVA analysis was conducted.
$^{two-way-ancova}$ Two-way between subjects ANCOVA analysis was conducted.
$^{one-way-anova}$ One-way between subjects ANOVA analysis was conducted.
^{b1} $Low(n = 67)$, $High(n = 41)$, $F(1, 106)$, 6 data were missing due to a technical error of gameplay logging.
^{b2} $Low(n = 35)$, $Medium(n = 39)$, $High(n = 34)$, $F(2, 105)$, 6 data were missing due to a technical error of gameplay logging.

H1b, H1g and H3a and H3g are supported with a statistical significance.

No interaction effects were found (RQ1 is answered).

4.2.2 Mediation Analysis

Before testing the mediation, a Pearson correlation was run to gauge the relationship between manipulation check variables and other dependent variables (Table 4.4). There was a statistically significant, negative correlation between performance and perceived task load (r = -.37, n = 108, p < .001), whereas there was a positive correlation between performance and competence (r = .29, n = 108, p = .002). Moreover, perceived task load was positively correlated with effort (r = .40, n = 114, p < .001) while was negatively correlated with competence (r = .-29, n = 114, p = .002). Interestingly, it was revealed that there was a significant positive correlation between perceived mental load and enjoyment and flow (r = .19, n = 114, p = .048 and r = .20, n = 114, p = .031, respectively).

Multiple mediation tests were conducted after taking the correlations and main effects in the previous analysis into account to examine whether the proposed mediators had effects on the other dependent variables. Since there were no main effects of mental and physical load on competence and flow, we did not conduct further mediation analysis for these variables (H5 and H6 are not

Table 4.4: Pearson Correlations of Manipulation Check and Dependent Variables

	Perceived Mental Demand	Perceived Exertion	NASA-TLX	Enjoyment	Flow	Competence	Effort	Future Play Preference	Performance
	M = 4.20 *SD* = 1.66	*M* = 14.75 *SD* = 2.87	*M* = 4.47 *SD* = 0.98	*M* = 4.79 *SD* = 1.05	*M* = 4.61 *SD* = 1.02	*M* = 3.80 *SD* = 1.25	*M* = 5.00 *SD* = 0.93	*M* = 3.46 *SD* = 1.75	*M* = 0.53 *SD* = .19
Perceived Mental Demand	-								
Perceived Exertion	.25**	-							
NASA-TLX	.62**	.62**	-						
Enjoyment	.19*	.03	.15	-					
Flow	.20*	-.02	.15	.55**	-				
Competence	-.01	-.32**	-.29**	.31**	.30**	-			
Effort	.20*	.34**	.40**	.48**	.37**	-.05	-		
Future Play Preference	.03	-.02	.05	.52**	.47**	.16	.27**	-	
Performance[a]	-.29**	-.40**	-.37**	.08	.19	.29**	-.05	.05	-

Note. $n = 114$.
$^{*}p < .05$, $^{**}p < .01$.
[a] $n = 108$, 6 data were missing due to a technical error of gameplay logging.

supported).

An SPSS macro entitled PROCESS developed by Hayes (2012)[1] was run and bootstrapping method was used to determine the significance of indirect effects (resample was set to 1,000). Three conditions should be established to determine whether mediation has occurred. First criteria is that the total effect of IV on DV (path c) should be statistically significant. Second, IV should predict the mediator M (path a). Lastly, the indirect effect of IV and M together should be statistically significant (confirming M significantly predicts DV (path b) and a drop in the significance (non-significance for full mediation) of direct effect (path c')). For example, in the first analysis of performance as a mediator between mental load and perceived task load by controlling gender; the total effect of mental load on task load was significant .71 ($SE = .18$), $p < 0.001$ and when the mediator included in the model, the significance direct effect reduced to .49 ($SE = .20$), $p = 0.015$ (partial mediation). The biased corrected confidence interval of the indirect effect (.009 and .50), which is the difference between total and direct effect, did not contain zero which showed that the indirect effect was significant. The same procedure was used to test the other mediation hypothesis about physical load. Results showed that performance mediated the relationship between perceived mental and physical load and task load, therefore H2 and H4 were supported (Table 4.5).

[1] The PROCESS macro for SPSS: `http://www.processmacro.org/`

Table 4.5: Test of Indirect Effects of Independent Variables (IVs) on Task Load (DV) Through Performance (M)

IV (Mental Load), M (Performance), DV (Perceived Taskload) ^{p}	b^*	SE	p
IV to M (path a)	-0.17	0.03	0.000
M to DV (path b)	-1.28	0.51	0.014
Total effect of IV on DV (path c)	0.71	0.18	0.002
Direct effect of IV on DV (path c')	0.49	0.20	0.015
Indirect effects of IV on DV through M (path a*b)	0.22	0.13	
Biased corrected 95% CI: .0.009 - .50			
IV (Physical Load), M (Performance), DV (Perceived Taskload)	b^*	SE	p
IV to M (path a)	-0.07	0.02	0.001
M to DV (path b)	-1.60	0.49	0.002
Total effect of IV on DV (path c)	0.29	0.12	0.013
Direct effect of IV on DV (path c')	0.18	0.12	0.13
Indirect effects of IV on DV through M (path a*b)	0.12	0.05	
Biased corrected 95% CI: 0.02 - 0.22			

Note.n = 114.*p* = partial mediation. b^* = standardized regression coefficient, SE = standard error, CI = confidence interval

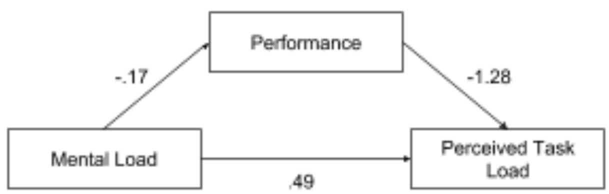

Figure 4.1: Indirect effect of mental load on perceived task load through performance (partial mediation).

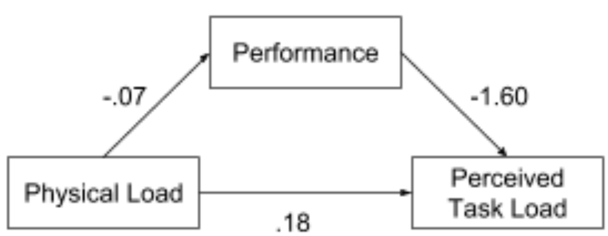

Figure 4.2: Indirect effect of physical load on perceived task load through performance (full mediation).

CHAPTER 5

DISCUSSION AND CONCLUSION

The purpose of this study was to observe how the perceived physical load and mental load affect players' gameplay experiences in exergames in an experimental setting. Using dual-task methodology approach, this research is the first attempt to explain the contributions and interactions between mental and physical demand required by exergames for gameplay experiences including enjoyment, flow, competence, effort, future play preferences and performance.

5.1 Discussion

Supporting the previous findings about the positive relationship between mental and physical workload and overall taskload (DiDomenico and Nussbaum, 2008; Hancock, 1989), we found that as perceived mental load and physical load increased in an exergame, players' perceived task load also increased significantly. As DiDomenico and Nussbaum (2008) showed and Park et al. (2013) suggested, higher mental and physical load resulted in lower performance in this study (Paas and Van Merriënboer, 1994; Paas et al., 2003). The mediating effects of performance between mental (partial) and physical (full) load and perceived task load were found which suggests that performance can also explain the impacts of these load on how much we perceive the overall task load. Moreover, subjects rated more perceived effort with an increase in mental load as it was proposed by Wickens (2017).

Although there was no significance in the results regarding other dependent variables of enjoyment, flow, competence and future play preference, the changes in the patterns with higher mental and physical load (see Appendix M) are supported by other research. As mental and physical load increased, enjoyment and flow decreased as proposed in the previous studies (Ang et al., 2007; Sinclair et al., 2009; Tozman et al., 2015; Qin et al., 2009). When mental and physical load are observed individually, the results highlight that players enjoy playing the exergame when the physical load and mental load are low. However, when we consider the combinations of these two load, such an approach can mislead us since the interactions are ignored (Kiili et al., 2013, pp. 203-204). One important aspect in the results is that following the perspective of dual-task interference (Park et al., 2013), the combinations of different levels of mental and physical load in exergame design may result in different gameplay experiences. For example, in mental and physical load medium condition, subjects rated their enjoyment as highest, whereas in mental load medium and physical load high condition, subjects rated their enjoyment as lowest across all conditions.

The main challenge in this study was to implement dual-task methodology by combining the physical and mental challenge effectively. In exergames, "players frequently perform multiple tasks simultaneously" (Park et al., 2013, p. 419). Since, in exergame design, there is a continuous demand by two channels of tasks (physical performance and mental attention), it becomes crucial to find a way of manipulating these two for an optimum gameplay experience. This might underline the reason why there were no significant differences between different levels of loads in their effects on motivational outcomes such as enjoyment and flow. More specifically, as physical performance also requires a certain level of cognitive effort (e.g. precision, endurance, balance) (Lyons, 2015), the target game in this study might be more mental load focused hindering the physical load. Moreover, "exercise intensity impacts the focus of attention" (as cited in Kiili et al., 2013), and "during high physical workload it is hard to concentrate on problem-solving and game stimuli" (Kiili et al., 2013, p. 204). This increase in "allocation of attention" during dual-task performance (Brisswalter et al., 2002; DiDomenico and Nussbaum, 2011) to coordinate physical and perceptional resources might intensify the level of mental load required by the game and perceived by the subjects. Thus, trying to implement and manipulate physical and mental challenges seperately in the game design might be useless. Considering these interactions, the dual-flow model for exergames highlights that "optimal exergaming experience can be achieved when both the psychological (cognitive) and physiological challenges are in balance" (Sinclair et al., 2009). Therefore, assessing exergames as dual-task interplay, the challenges in the game should be designed in a way that the game mechanics requiring cognitive and physiological aspects should be adjusted and utilized for a continuous challenge-skill balance (Martin-Niedecken and Götz, 2017; Kiili et al., 2013; Sinclair et al., 2009; Velazquez et al., 2013). One approach is suggested by Kiili et al. (2013, p. 205) is "sequence the cognitive gameplay and physical gameplay" such that the tasks are alternated over time (i.e. one challenge is introduced while controlling the other). Therefore, instead of using a design which requires concurrent mental and physical effort, a game which implements this approach (requiring either mental or physical effort at a time) could be tested in future research. In addition, deviating from dual-task methodology, since physical and cognitive activities are serving for the same goal of "playing the game", exergames can be evaluated as one combined task of physical and cognitive activities rather than multiple tasks performed simultaneously. Further research is needed to reveal the interplay of these two components in exergame design.

5.2 Limitations

5.2.1 Participants

Since only participants enrolled in mainly undergraduate studies at a major Midwest university participated in the study, this may impact generalizability. Therefore, whether our the findings in this study hold for other populations remains to be determined. Moreover, when we consider the characteristics of participants like gameplay and exercising experiences, a wider range of participants with broader skill levels should be included and these prior experiences can be measured with well-established instruments from the literature. Additionally, there might be a role of "interest" towards what the exergame is promising for participants on their overall experiences. According to Chen and Wang (2017), interest in learning outcomes offered by an exergame may motivate students to engage in physical activity. Similarly, the outcomes offered by the exergame (e.g. energy expenditure, in-game rewards) might have an impact on subjects' motivational experience during gameplay. In the future research, the ratings for such an inclination (e.g. interest in exercising, interest in competition) might be also observed for its effect on overall gameplay experiences.

5.2.2 Exercising Characteristics and Cognitive Task Type

Exercising duration and increased fatigue are some determinants of the effects of physical exertion on cognitive performance (Brisswalter et al., 2002, p.559). For example, with a moderate level of physical exercise, performance can increase and with a more intense exercising activity cognitive performance may start to decrease with an Inverted-U effect (as cited in Brisswalter et al., 2002. In this study, due to the length of procedure in the experimental setting with the calibration sessions, we limited the time of the gameplay to 5 minutes. Since longer gameplay activity may result in higher levels of physical demand, its contributions to the overall gameplay experiences might be different. Moreover, although mental load was manipulated using multiple-object tracking task in the study, other types of cognitive tasks (e.g. requiring complex decision making, reaction time) should be integrated in the exergame design (using different game genres) and tested for future research.

5.2.3 Exercising Device

Lastly, exercising device was another limitation in our study. Since we tested our hypotheses on one exercising medium, gameplay experiences might be different on when using another device. Independently from the game design itself, different exercising settings may require different levels of physical load (and probably cognitive attention) which may affect gameplay experience in turn. A comparison between different exercising devices may also contribute to the research in this direction.

Focusing on the motivational outcomes, this study is promising for the future research to observe the dynamic and continuous interaction between mental and physical load in exergames. A design guideline can be provided for optimized gameplay experiences through the lens of mental-physical load balance in exergames after conducting future research with more significant results on game enjoyment and flow. Considering the limitations listed above, future work can help designers to create better exergames for a permanent interest in exercising and continuous interest in gameplay.

APPENDICES

APPENDIX A

SCREENING SURVEY

A.1 Demographics

1. What is your gender?

   - ☐ Male
   - ☐ Female
   - ☐ Transgender

2. What is your age?

   - ☐ 16
   - ☐ 17
   - ☐ 18
   - ☐ 19
   - ☐ 20
   - ☐ 21
   - ☐ 22
   - ☐ 23
   - ☐ 24
   - ☐ 25
   - ☐ over 25

3. How would you describe yourself?

   - ☐ American Indian or Alaska Native

- ☐ Asian
- ☐ Black or African American
- ☐ Hispanic or Latino
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ White or Caucasian
- ☐ Mixed Race

A.2 Physical Activity Readiness (PAR)

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Canadian journal of sport sciences*.

1. Has your doctor ever said that you have a heart condition and that you should only do physical activity recommended by a doctor?
   - ☐ Yes
   - ☐ No
2. Do you feel pain in your chest when doing physical activity?
   - ☐ Yes
   - ☐ No
3. In the past month, have you had chest pain when you were not doing physical activity?
   - ☐ Yes
   - ☐ No
4. Do you lose your balance because of dizziness or do you ever lose consciousness?
   - ☐ Yes

- ☐ No

5. Do you have a bone or joint problem that could be made worse by a change in your physical activity?
   - ☐ Yes
   - ☐ No
6. Is your doctor currently prescribing medications for your blood pressure or heart condition?
   - ☐ Yes
   - ☐ No
7. Do you know of any other reason why you should not do physical activity?
   - ☐ Yes
   - ☐ No

A.3 Previous Gaming Experience

1. How many years have you been playing video games? _______________ (e.g. 10)
2. How many hours in a week do you play video games? _______________ (e.g. 10)
3. Which genre of video games do you play? (Check all that apply)
   - ☐ Action
   - ☐ Action-Adventure
   - ☐ Simulation
   - ☐ Strategy
   - ☐ Role Playing
   - ☐ Sports

- ☐ Casual
- ☐ Other (Please specify): _______________

4. Are you comfortable playing a casual game including a cognitive task?
   - ☐ Yes
   - ☐ No
5. Are you comfortable tracking task of multiple objects in the screen?
   - ☐ Yes
   - ☐ No
6. Are you comfortable using exercise bike while playing a casual game including a cognitive task?
   - ☐ Yes
   - ☐ No

A.4 Previous Exercising Experience

1. Are you comfortable doing exercise?
   - ☐ Yes
   - ☐ No
2. How many hours do you exercise in a week? _______________ (e.g. 6)
3. Which exercise devices do you use? (Check all that apply)
   - ☐ Treadmill
   - ☐ Exercise Bike
   - ☐ Elliptical

- ☐ Weight Benches
- ☐ Home Gyms
- ☐ Rowers
- ☐ Other (Please specify): _______________

4. What type of exercises do you do? (Check all that apply)

- ☐ Walking
- ☐ Running
- ☐ Swimming
- ☐ Cycling
- ☐ Hiking
- ☐ Fitness
- ☐ Yoga / Plates
- ☐ Strength & Weight Training
- ☐ Other (Please specify): _______________

5. Are you familiar with exercise bikes?

- ☐ Yes
- ☐ No

APPENDIX B

CONSENT FORM

INVESTIGATORS

Irem Gokce Yildirim, yildiri4@msu.edu Kuo-Ting Huang, huangku1@msu.edu Tom Day, daythoma@msu.edu Dr. Taiwoo Park, twp@msu.edu, 517-353-2198. 423 Communication Arts & Sciences, MSU Dr. Wei Peng, pengwei@msu.edu, 517-432-8235, 429 Communication Arts & Sciences, MSU

GENERAL INFORMATION ABOUT RESEARCH STUDIES

You are being asked to take part in a research study. To join the study is completely voluntary. You may refuse to join, or you may withdraw your consent to be in the study at any time, for any reason, without penalty or loss of any benefits to which you are entitled. Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study.

PURPOSE OF RESEARCH

You are being asked to participate in a research study on the perception of physical and cognitive loads in exergames. You have been selected as a possible participant in this study because of your involvement with the MSU SONA system. From this study, the researchers hope to understand experiences of exergames.

LENGTH OF STUDY

Participants will complete the study only once and the procedure will take about 60 minutes.

PROCEDURE OF STUDY

The study is part of an academic research project to examine how people experience exergames. You completed an online screening survey before coming into the lab and we will use this information for research only if you agree to participate in this study.

In the lab, the first session will be the individual based ‘calibration’ session to gauge your baseline physical and cognitive performance levels in two separate phases.

- In the cognitive challenge calibration phase, you will be asked to play the game (without physical load - no exercise device), while gradually increasing the difficulty of game (e.g., by increasing the number of objects to track on the screen) over time. Anytime when you do not want to play the game, you are allowed to stop the process. At the end of cognitive challenge calibration phase, you will be given 1 minutes to rest.

- In the physical challenge calibration phase, you will be asked to cycle at certain speed (e.g., 100 RPM) on the exercise bike, while gradually increasing the tension level over time, to identify the maximum physical challenge for you. Anytime when you feel it becomes either too hard or uncomfortable for you to cycle, you are allowed to stop the process. At the end of physical challenge calibration phase, you will be given 5 minutes to rest.

After the calibration session, the gameplay session will be conducted. During this session, some gameplay data will be logged. Only the researchers of this study will have access to this data. You will be asked to play the video game using exercise bike for about 5 minutes, and after playing, you will be asked to complete some basic questions about your gameplay experiences, including physical and cognitive aspects. The entire procedure should take no more than 50 to 60 minutes to complete.

POTENTIAL RISKS

While the game used in this study is age-appropriate for teens and adults and there is an individual based adaptation for cognitive challenge and physical challenge on the exercise bike, there is the possibility that you might feel uncomfortable when playing the game and cycling. It is also possible that you may feel fatigue or experience some muscle soreness after performing the physical workout. This will not be more than one you experience in your exercise routine. Therefore, there will be no special problems or risks during or after the study. If you have some physical injury, performing the tasks may worse it. If you have such injury, you may not participate in this study. Should you feel any discomfort as a result of this (or any other) reasons resulting from your participation you may withdraw your consent and data from the study by stopping playing at any time. If you withdraw from the study before its completion, we will not use any of the data you provide us.

If you encounter any injury as a result of your participation in the study, Michigan State University will assist you in obtaining emergency care. If you have insurance for medical care, the bills of treatment will be charged to your insurance carrier and any costs that are not covered by your insurance will be your responsibility.

PRIVACY AND CONFIDENTIALITY

The data for this project are being collected anonymously using your SONA ID. Neither the researchers nor anyone else will be able to link data to you. Your survey answers for the study will be collected and stored on a separate secure website. Information that is printed out will be stored in a locked filing cabinet in the locked office of the principal investigator and destroyed when analysis is complete. Data that is downloaded will be stored on the investigators' laptop computer and will have a double password protected log-on and the files will be password protected as well. Additionally, the IRB will have access to the data. Research records will be kept for at least 3 years after the close of the study. The only people with access to the data will be Dr. Taiwoo Park, Dr. Wei Peng, Irem Gokce Yildirim, Tim Huang, Tom Day and the MSU HRPP. Your confidentiality will be protected to the maximum extent allowable by law.

Participants will not be identified in any report or publication about this study.

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at Olds Hall, 408 West Circle Drive #207, MSU, East Lansing, MI 48824.

COSTS AND COMPENSATION FOR BEING IN THE STUDY

In the SONA system, 1 hour of research participation is worth 1 SONA credit and this credit is pro-rated in 15-minute increments. The duration of this study is approximately 60 minutes. Hence, participants who complete this study will receive 1 SONA credits.

QUESTIONS ABOUT THIS STUDY

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher (Irem Gokce Yildirim, 430 Communication Arts and Sciences Building, yildiri4@msu.edu, 517-348-3866; Kuo-Ting (Tim) Huang, huangku1@msu.edu, 850-345-3805; Tom Day, daythoma@msu.edu, 586-453-1832; Dr. Taiwoo Park, 517-353-2198; Dr. Wei Peng, 517-432-8235).

STATEMENT OF CONSENT

Your signature below means that you voluntarily agree to participate in this research study.

______________ ______________

Signature **Date**

APPENDIX C

NASA-TLX: TASK LOAD INDEX

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage Publications. Retrieved from: `https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf`

Please rate the work load you experienced during the gameplay in the following questions, using the following scale.

1	2	3	4	5	6	7
Very low						Very high

1. How mentally demanding was the task? (MENTAL DEMAND)
2. How physically demanding was the task? (PHYSICAL DEMAND)
3. How hurried or rushed was the pace of the task? (TEMPORAL DEMAND)
4. How successful were you in accomplishing what you were asked to do? (R) (PERFORMANCE)
5. How hard did you have to work (mentally and physically) to accomplish your level of performance? (EFFORT)
6. How insecure, discouraged, irritated, stressed, and annoyed were you? (FRUSTRATION LEVEL)

APPENDIX D

BORG RATING OF PERCEIVED EXERTION SCALE

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med sci sports exerc*, 14(5), 377-381.

Please choose the number from below that best describes your level of exertion.

- ☐ 6 - No exertion at all
- ☐ 7 - Extremely light
- ☐ 8
- ☐ 9 - Very light (easy walking slowly at a comfortable pace)
- ☐ 10
- ☐ 11 - Light
- ☐ 12
- ☐ 13 - Somewhat hard (It is quite an effort; you feel tired but can continue)
- ☐ 14
- ☐ 15 - Hard (heavy)
- ☐ 16
- ☐ 17 - Very hard (very strenuous, and you are very fatigued)
- ☐ 18
- ☐ 19 - Extremely hard (You can not continue for long at this pace)
- ☐ 20 - Maximal exertion

APPENDIX E

GAME ENJOYMENT

Song, H., Peng, W., & Lee, K. M. (2011). Promoting exercise self-efficacy with an exergame. *Journal of health communication, 16*(2), 148-162.

Please rate how well the following adjectives describe your game playing experience, using the following scale.

1	2	3	4	5	6	7
Very poor						Very well

1. boring (R)
2. exciting
3. enjoyable
4. entertaining
5. fun
6. interesting
7. pleasant

APPENDIX F

GAME FLOW

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education, 52*(1), 101-112.

Please indicate your agreement with the following statements, using the following scale.

1	2	3	4	5	6	7
Strongly Disagree						Strongly Agree

1. I forgot about time passing while playing the game.
2. I became unaware of my surroundings while playing the game.
3. I temporarily forgot worries about everyday life while playing the game.
4. I experienced an altered sense of time.
5. I could become involved in the game.
6. I felt emotionally involved in the game.
7. I felt viscerally involved in the game.
8. I enjoyed the game without feeling bored or anxious.
9. The challenge was adequate, neither too difficult nor too easy.

APPENDIX G

PERCEIVED COMPETENCE

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport, 60*(1), 48-58. Retrieved from `http://selfdeterminationtheory.org/intrinsic-motivation-inventory/`

For each of the following statements, please indicate how true it is for you, using the following scale.

1	2	3	4	5	6	7
Not at all true			Somewhat true			Very true

1. I think I am pretty good at this game.
2. I think I did pretty well at this game, compared to other players.
3. After playing this game for awhile, I felt pretty competent.
4. I am satisfied with my performance at this game.
5. I could become involved in the game.
6. I was pretty skilled at this game.
7. This was a game that I couldn't do very well. (R)

APPENDIX H

PERCEIVED EFFORT

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport, 60*(1), 48-58. Retrieved from `http://selfdeterminationtheory.org/intrinsic-motivation-inventory/`

For each of the following statements, please indicate how true it is for you, using the following scale.

1	2	3	4	5	6	7
Not at all true			Somewhat true			Very true

1. I put a lot of effort into this game.
2. I didn't try very hard to do well at this game. (R)
3. I tried very hard on this game.
4. It was important to me to do well at this game.
5. I didn't put much energy into this game. (R)

APPENDIX I

FUTURE PLAY PREFERENCES

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion, 30*(4), 344-360.

For each of the following statements, please indicate how true it is for you, using the following scale.

1	2	3	4	5	6	7
Not at all true			Somewhat true			Very true

1. Given the chance I would play this game in my free time.
2. I would be interested in having my own personal copy of this game.
3. I would like to spend more time playing this game.
4. I am excited to play this game in the future.

APPENDIX J

PERFORMANCE

Performance was calculated as the ratio of the average of the successful feeding number across all of the rounds played during the main gameplay session to the total number of turtles set for the gameplay session. The data was collected during gameplay in a log file.

APPENDIX K

OTHER GAMEPLAY METRICS

Average speed and total gameplay duration were other gameplay log metrics we collected.

- **Average Speed:** Calculated based on the non-zero speed information during gameplay.
- **Total Gameplay Time:** Total time spent on gameplay during 5 min of gameplay duration.
- **Total Resting Time:** Total time spent on resting without cycling (speed=0) during the 5 min of gameplay time.

APPENDIX L

CALIBRATION SESSION STATISTICS

L.1 Mental Load Calibration Phase Data

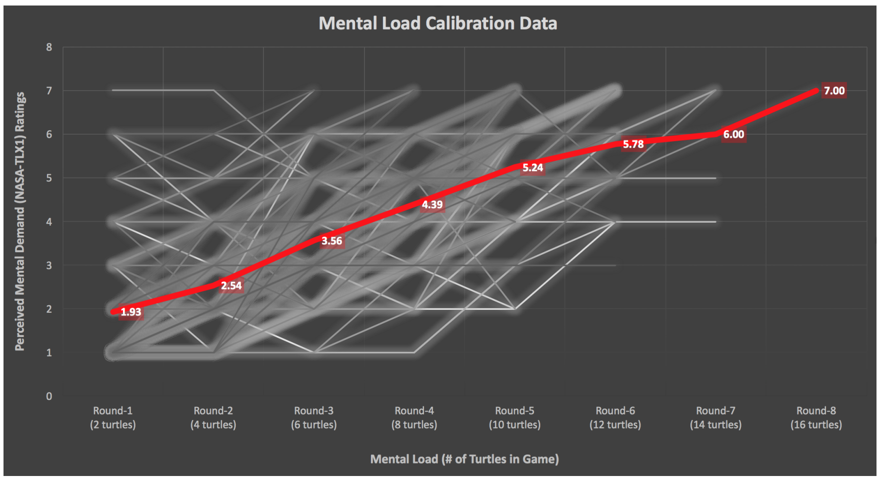

Figure L.1: Perceived mental load ratings across participants as mental load increases during mental load calibration phase trials.

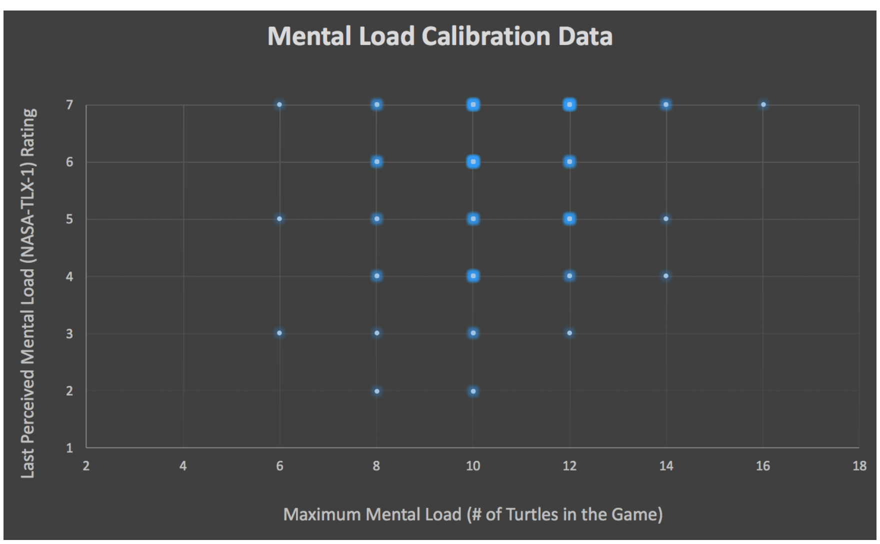

Figure L.2: Subjects' last perceived mental load ratings and maximum mental load levels identified at the end of mental load calibration phase. ($M_{maximum\_mental\_load}$ = 10.44, $M_{last\_mental\_load\_rating}$ = 5.61)

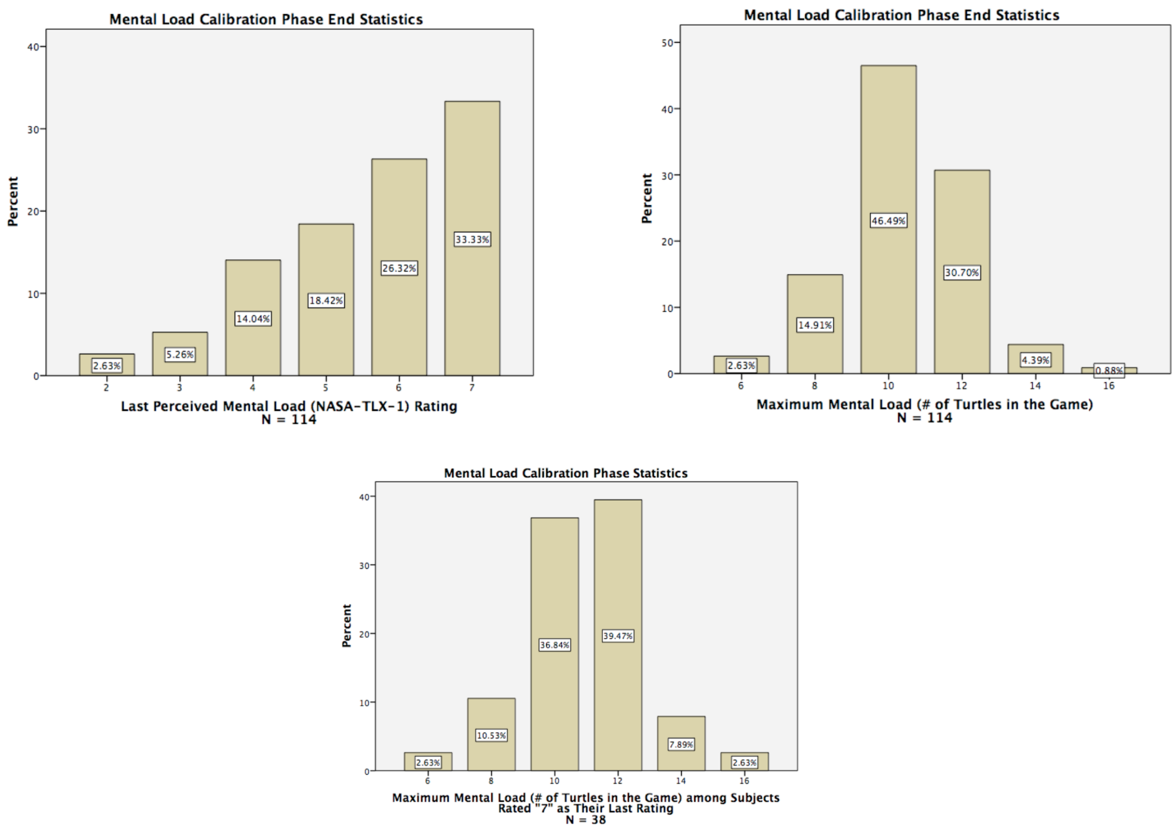

Figure L.3: Subjects' last perceived mental load ratings and maximum mental load level statistics at the end of mental load calibration phase.

L.2 Physical Load Calibration Phase Data

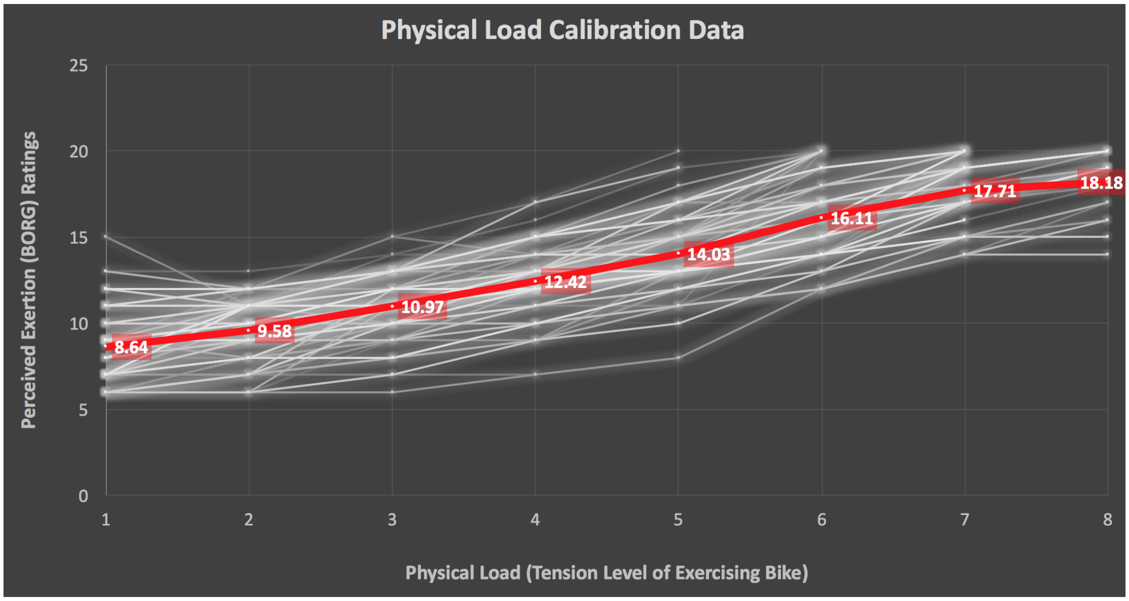

Figure L.4: Subjects' perceived exertion (BORG) ratings as physical load increases during physical load calibration phase.

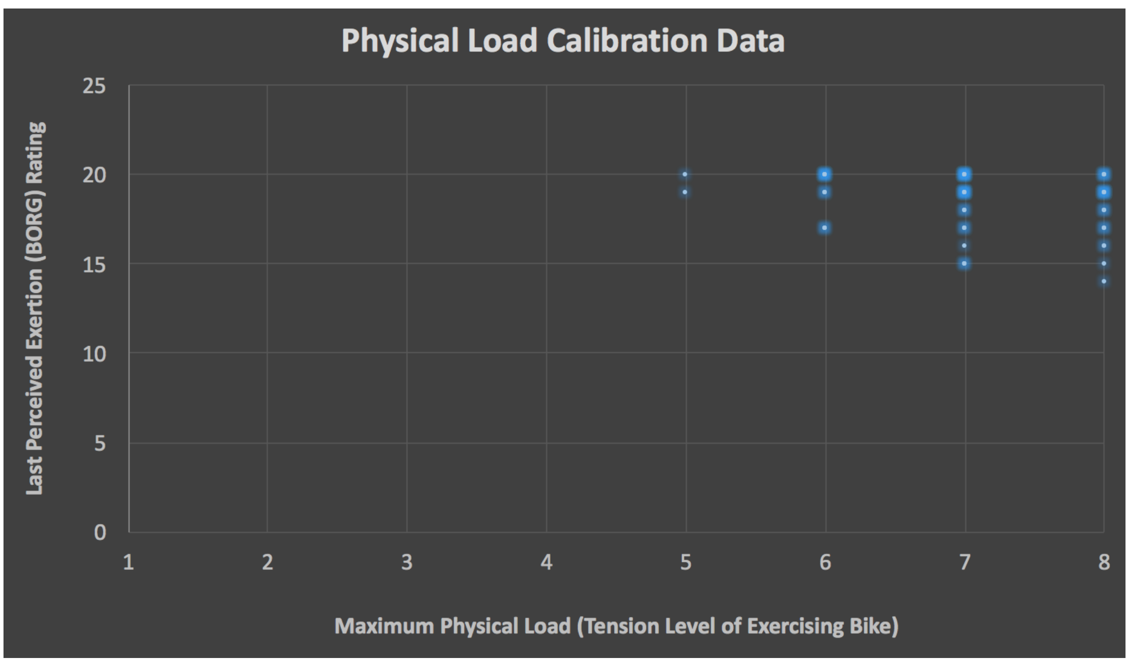

Figure L.5: Subjects' last perceived exertion (BORG) ratings and maximum physical load level statistics at the end of physical load calibration phase.
($M_{maximum\_physical\_load}$ = 7.08, $M_{last\_physical\_load\_rating}$ = 18.50)

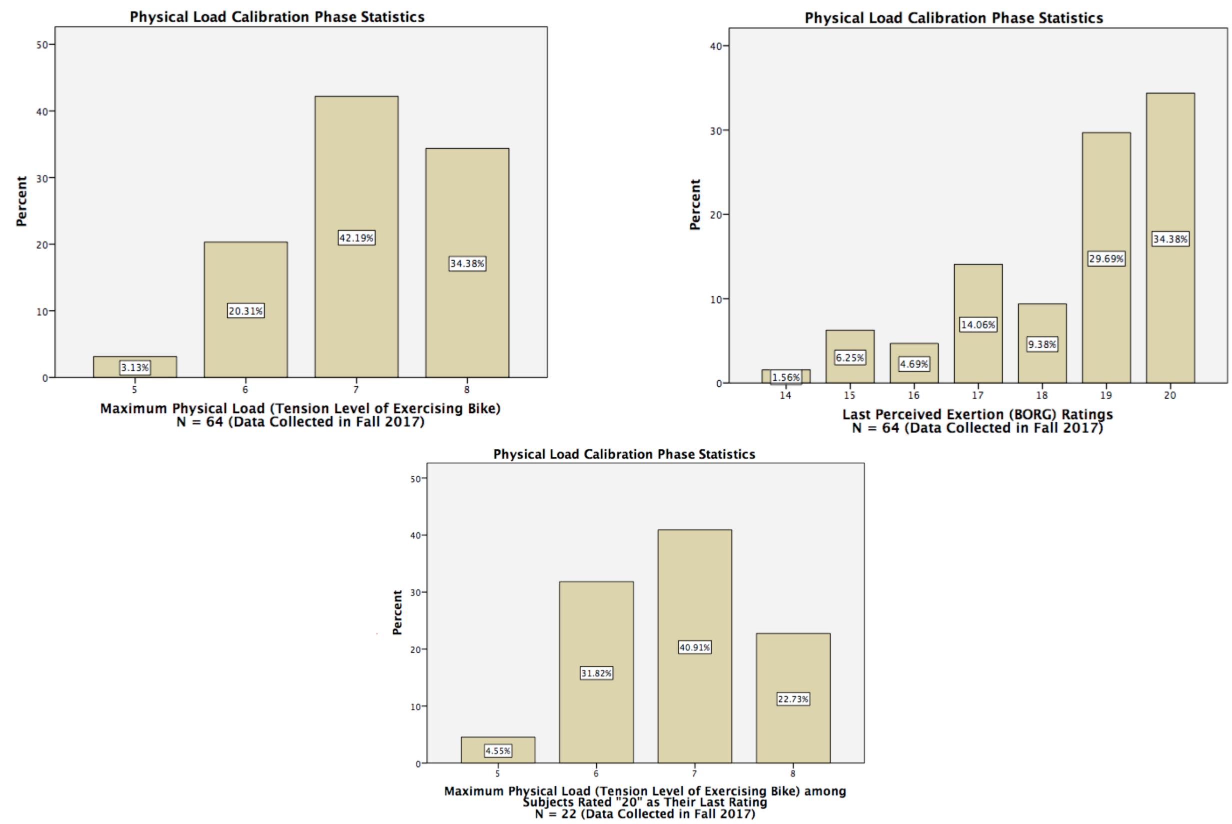

Figure L.6: Subjects' last perceived exertion ratings and maximum physical load level statistics at the end of physical load calibration phase.

APPENDIX M

MEAN SCORES FOR THE DEPENDENT VARIABLES

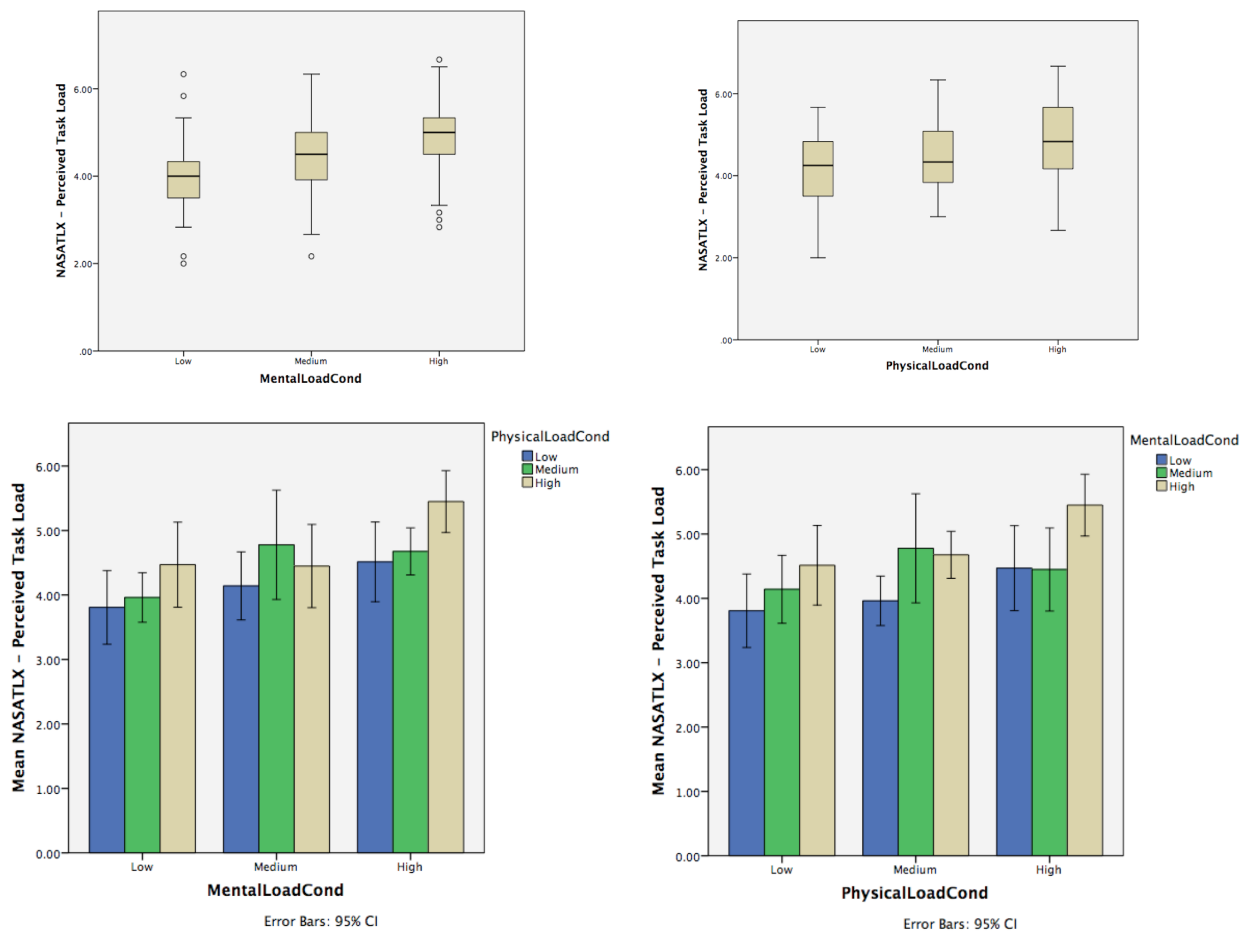

Figure M.1: Mean scores for NASA-TLX across mental and physical load conditions respectively.

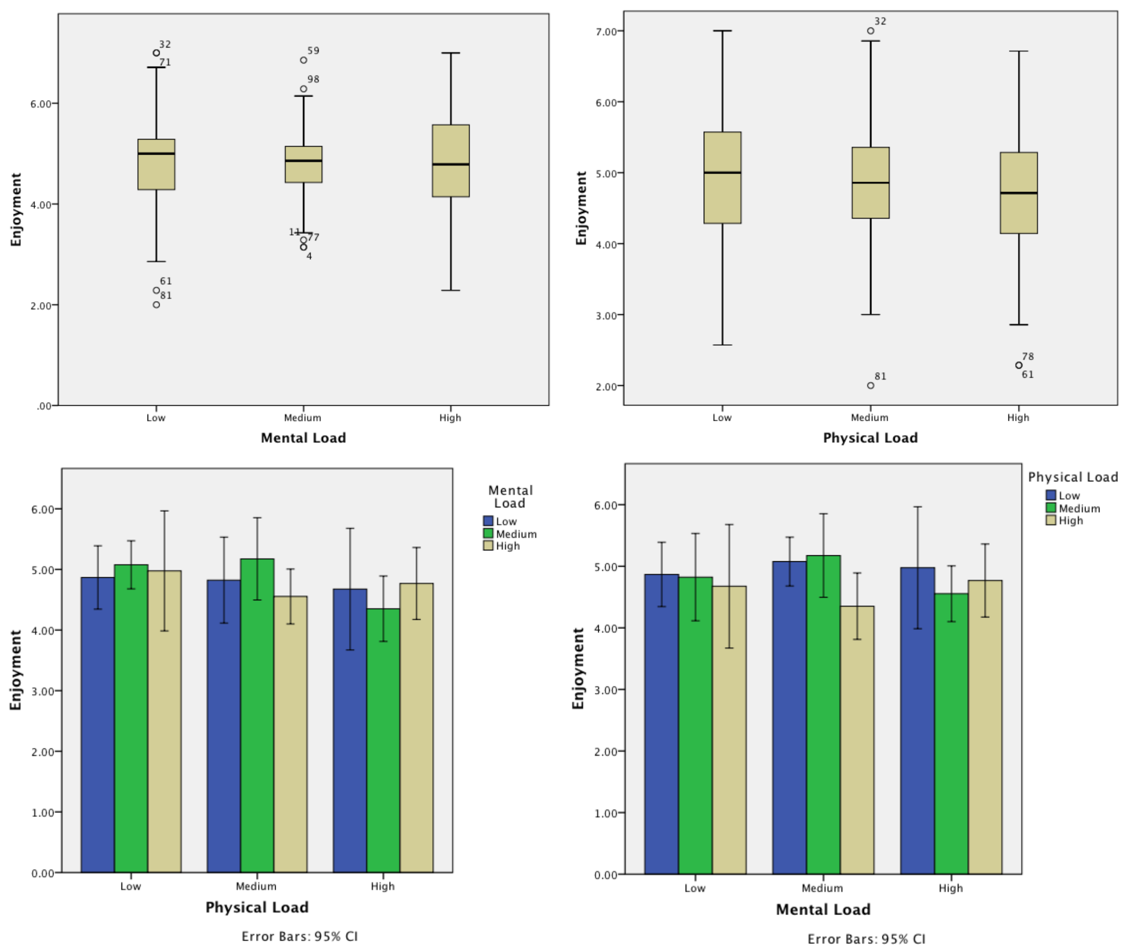

Figure M.2: Mean scores for enjoyment across mental and physical load conditions respectively.

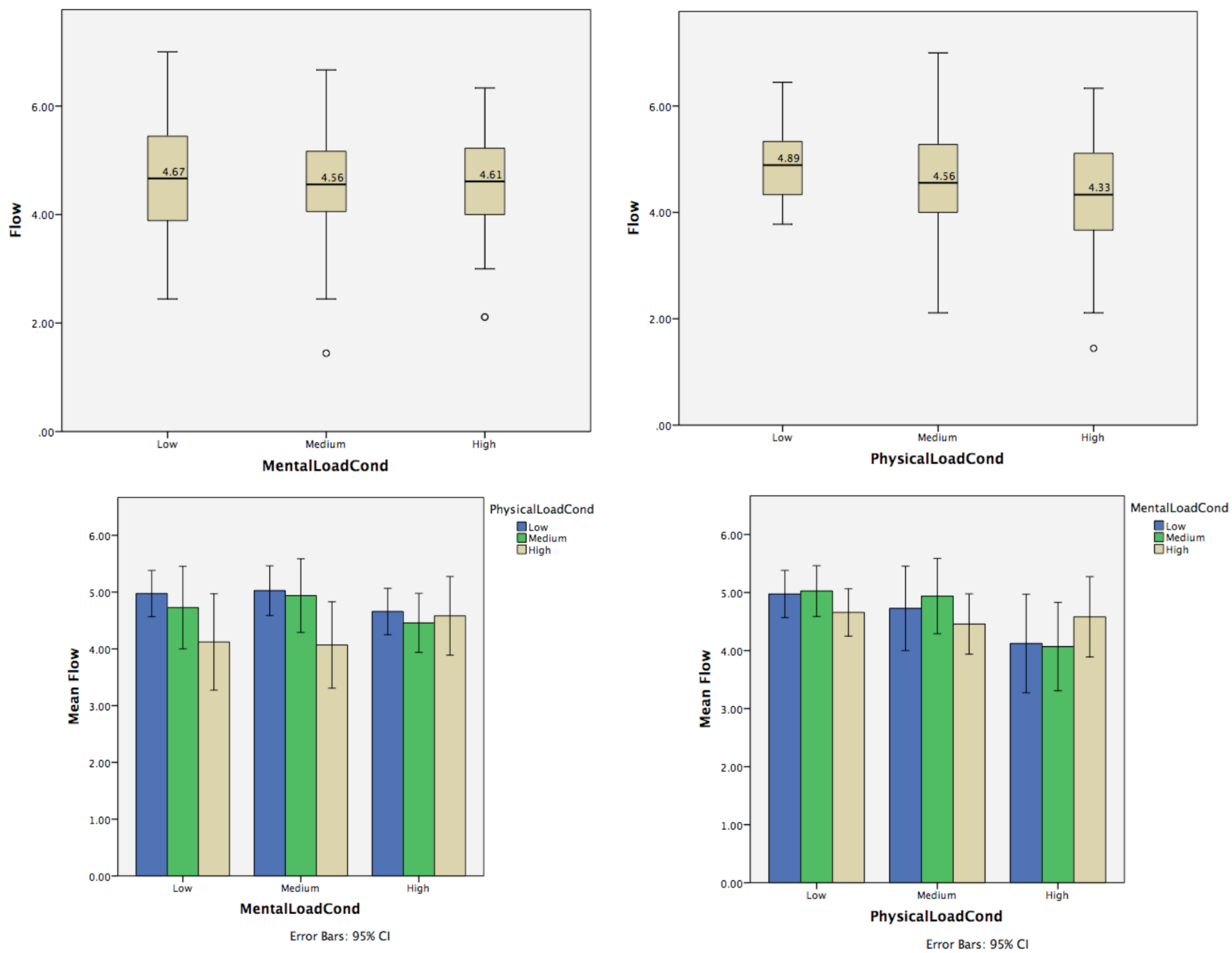

Figure M.3: Mean scores for flow across mental and physical load conditions respectively.

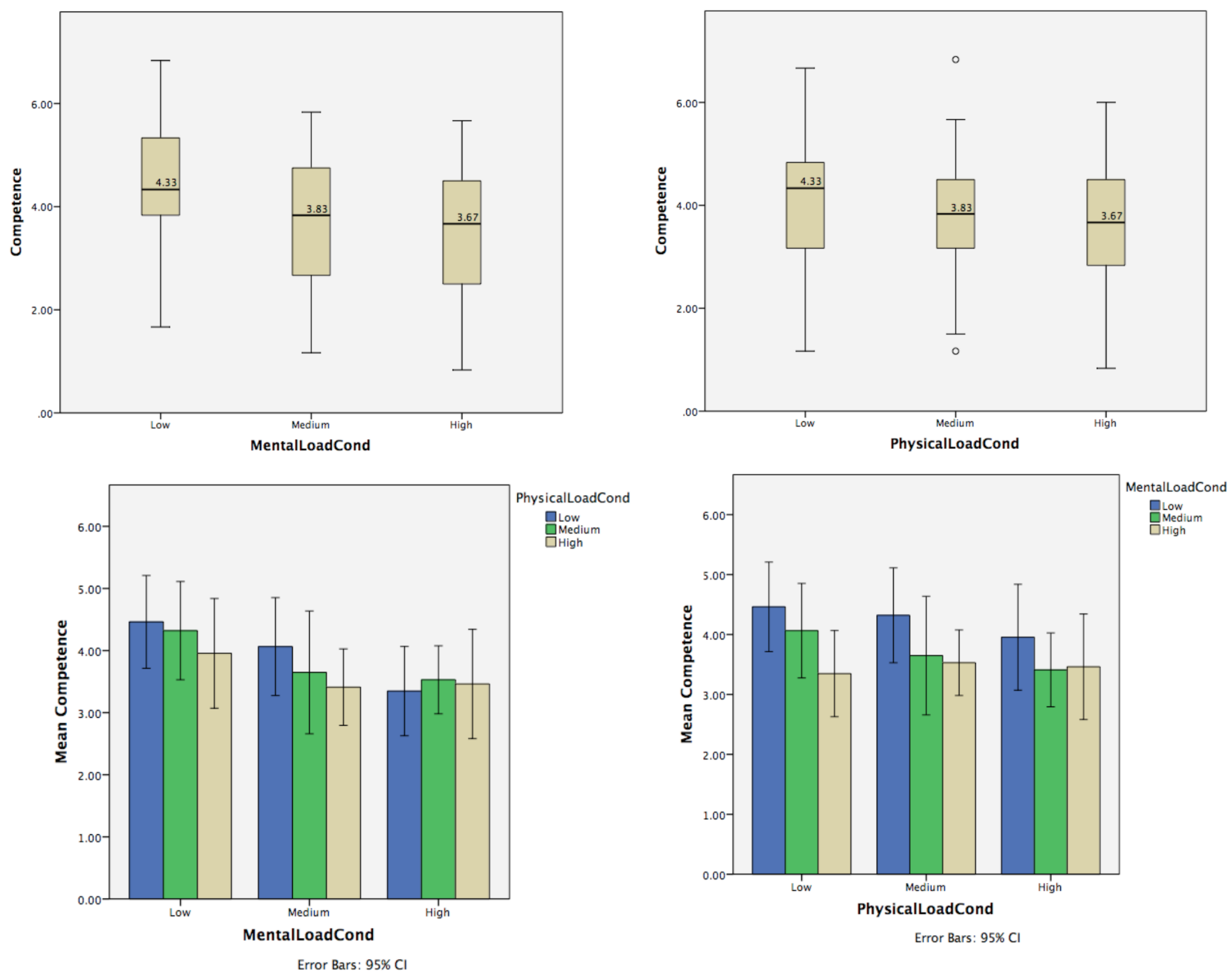

Figure M.4: Mean scores for competence across mental and physical load conditions respectively.

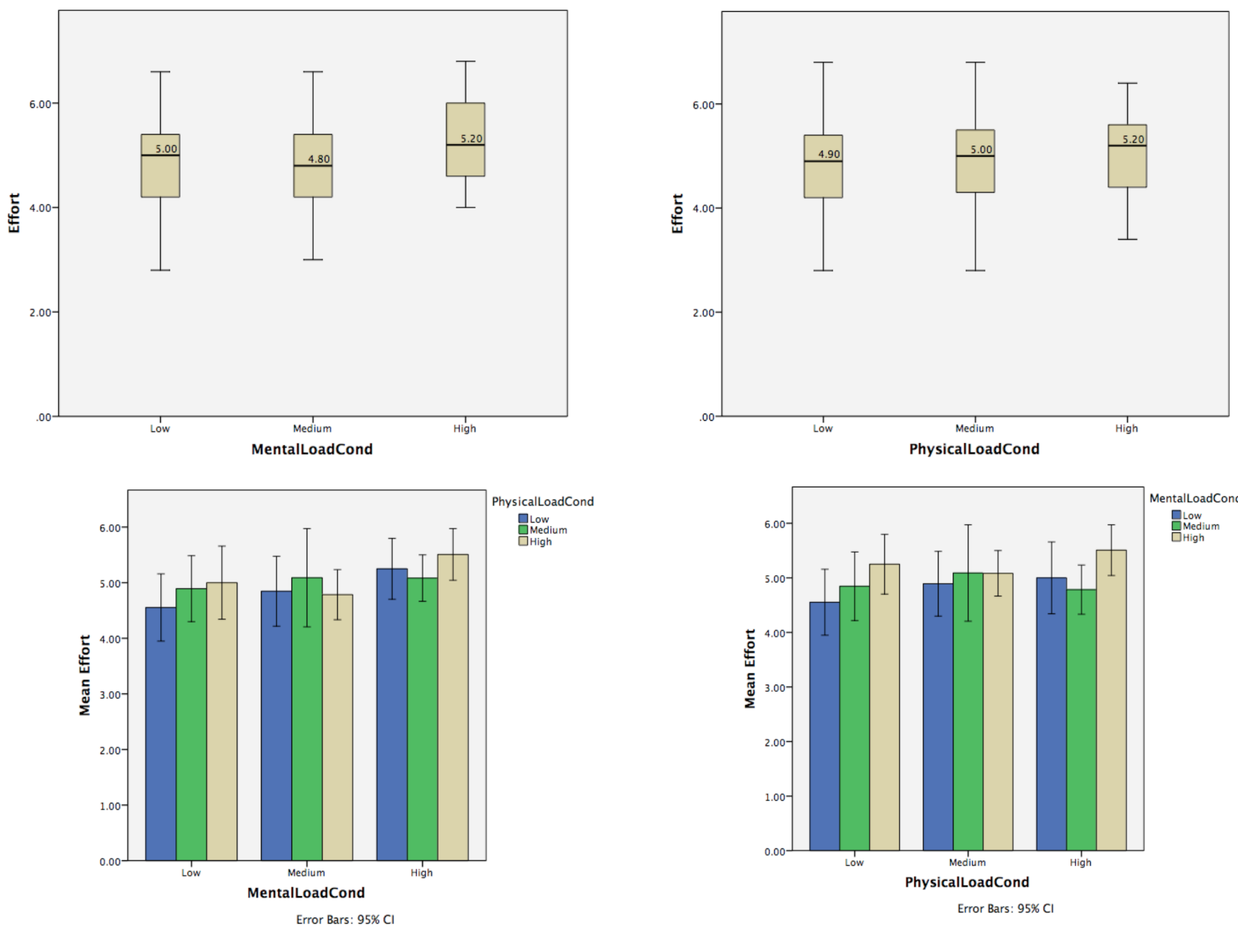

Figure M.5: Mean scores for effort across mental and physical load conditions respectively.

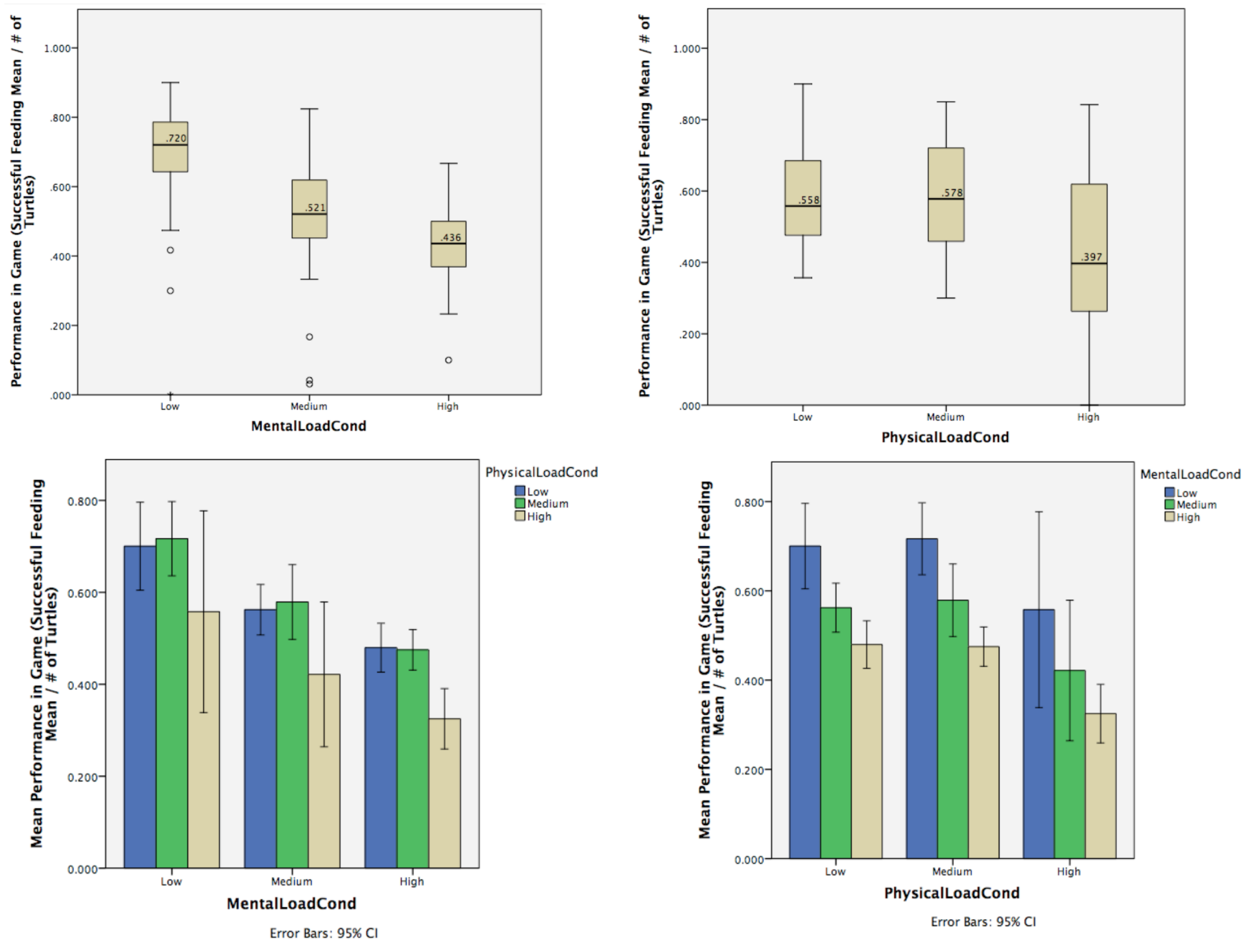

Figure M.6: Mean scores for performance across mental and physical load conditions respectively.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, R., Mcgeorge, P., Pearson, D. G., and Milne, A. (2006). Multiple-target tracking: A role for working memory? *The Quarterly journal of experimental psychology*, 59(6):1101–1116.

Ang, C. S., Zaphiris, P., and Mahmood, S. (2007). A model of cognitive loads in massively multiplayer online role playing games. *Interacting with computers*, 19(2):167–179.

Annetta, L. A., Minogue, J., Holmes, S. Y., and Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, 53(1):74–85.

Arroyo-Gómez, N., Laparra-Hernández, J., Soler-Valero, A., Medina, E., and de Rosario, H. (2017). Physiological model to classify physical and cognitive workload during gaming activities. In *International Conference on Applied Human Factors and Ergonomics*, pages 246–254. Springer.

Borg, G. (1990). Psychophysical scaling with applications in physical work and the perception of exertion. *Scandinavian journal of work, environment & health*, pages 55–58.

Borg, G. (1998). *Borg's perceived exertion and pain scales.* Human kinetics.

Botterill, K., Allen, R., and McGeorge, P. (2011). Multiple-object tracking. *Experimental psychology*.

Brisswalter, J., Collardeau, M., and René, A. (2002). Effects of acute physical exercise characteristics on cognitive performance. *Sports medicine*, 32(9):555–566.

Brunken, R., Plass, J. L., and Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational psychologist*, 38(1):53–61.

Capodaglio, P. (2001). The use of subjective rating of exertion in ergonomics. *Giornale italiano di medicina del lavoro ed ergonomia*, 24(1):84–89.

Chen, A. and Wang, Y. (2017). The role of interest in physical education: A review of research evidence. *Journal of Teaching in Physical Education*, 36(3):313–322.

Constant, T., Levieux, G., Buendia, A., and Natkin, S. (2017). From objective to subjective difficulty evaluation in video games. In *IFIP Conference on Human-Computer Interaction*, pages 107–127. Springer.

Coppin, A. K., Shumway-Cook, A., Saczynski, J. S., Patel, K. V., Ble, A., Ferrucci, L., and Guralnik, J. M. (2006). Association of executive function and performance of dual-task physical tests among older adults: analyses from the inchianti study. *Age and ageing*, 35(6):619–624.

Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. *Harper Perennial modern classics*.

Damos, D. (1991). *Multiple task performance*. CRC Press.

Davey, C. (1973). Physical exertion and mental performance. *Ergonomics*, 16(5):595–599.

DiDomenico, A. and Nussbaum, M. A. (2008). Interactive effects of physical and mental workload on subjective workload assessment. *International journal of industrial ergonomics*, 38(11):977–983.

DiDomenico, A. and Nussbaum, M. A. (2011). Effects of different physical workload parameters on mental workload and performance. *International Journal of Industrial Ergonomics*, 41(3):255–260.

Franconeri, S., Jonathan, S., and Scimeca, J. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological science*, 21(7):920–925.

Fu, F.-L., Su, R.-C., and Yu, S.-C. (2009). Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, 52(1):101–112.

Hancock, P. (1989). The effect of performance failure and task demand on the perception of mental workload. *Applied Ergonomics*, 20(3):197–205.

Hardy, S., Dutz, T., Wiemeyer, J., Göbel, S., and Steinmetz, R. (2015). Framework for personalized and adaptive game-based training programs in health sport. *Multimedia Tools and Applications*, 74(14):5289–5311.

Hart, S. G. and Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183.

Huang, W. D. and Johnson, T. (2009). Instructional game design using cognitive load theory. In *Handbook of research on effective electronic gaming in education*, pages 1143–1165. IGI Global.

Hutchinson, J. C. and Tenenbaum, G. (2006). Perceived effort—can it be considered gestalt? *Psychology of Sport and Exercise*, 7(5):463–476.

Jakobsen, M. D., Sundstrup, E., Persson, R., Andersen, C. H., and Andersen, L. L. (2014). Is borg's perceived exertion scale a useful indicator of muscular and cardiovascular load in blue-collar workers with lifting tasks? a cross-sectional workplace study. *European journal of applied physiology*, 114(2):425–434.

Kalyuga, S., Chandler, P., and Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied cognitive psychology*, 13(4):351–371.

Kalyuga, S. and Plass, J. L. (2009). Evaluating and managing cognitive load in games. In *Handbook of research on effective electronic gaming in education*, pages 719–737. IGI Global.

Kiili, K. and Perttula, A. (2013). A design framework for educational exergames. *New pedagogical approaches in game enhanced learning: Curriculum integration*, pages 136–158.

Kiili, K., Perttula, A., Arnab, S., and Suominen, M. (2013). Flow experience as a quality measure in evaluating physically activating serious games. In *International Conference on Games and Learning Alliance*, pages 200–212. Springer.

Kooiman, B. J. and Sheehan, D. P. (2014). The efficacy of exergames played proximally and over the internet on cognitive functioning for online physical education. *American Journal of Distance Education*, 28(4):280–291.

Krausman, A. S., Crowell III, H. P., and Wilson, R. M. (2002). The effects of physical exertion on cognitive performance. Technical report, ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD.

Kyndt, E., Dochy, F., Struyven, K., and Cascallar, E. (2011a). The direct and indirect effect of motivation for learning on students' approaches to learning through the perceptions of workload and task complexity. *Higher Education Research & Development*, 30(2):135–150.

Kyndt, E., Dochy, F., Struyven, K., and Cascallar, E. (2011b). The perception of workload and task complexity and its influence on students' approaches to learning: A study in higher education. *European journal of psychology of education*, 26(3):393–415.

Lai, Y.-C., Wang, S.-T., and Yang, J.-C. (2012). An investigation of the exergames experience with flow state, enjoyment, and physical fitness. In *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on*, pages 58–60. IEEE.

Larsen, L. H., Schou, L., Lund, H. H., and Langberg, H. (2013). The physical effect of exergames in healthy elderly—a systematic review. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 2(4):205–212.

Lyons, E. J. (2015). Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. *Games for health journal*, 4(1):12–18.

Maillot, P., Perrot, A., and Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychology and aging*, 27(3):589.

Martin-Niedecken, A. L. and Götz, U. (2017). Go with the dual flow: Evaluating the psychophysiological adaptive fitness game environment "plunder planet". In *Joint International Conference on Serious Games*, pages 32–43. Springer.

McAuley, E., Duncan, T., and Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport*, 60(1):48–58.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81.

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

Nankar, M. (2016). The influence of different mental processes (cognitive loads) on gait: A study of dual task function. Master's thesis, University of Manitoba.

O'Neil, H. F., Wainess, R., and Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Cirriculum Journal*, 16(4):455–474.

Owen, N., Sparling, P. B., Healy, G. N., Dunstan, D. W., and Matthews, C. E. (2010). Sedentary behavior: emerging evidence for a new health risk. In *Mayo Clinic Proceedings*, volume 85, page 1138. Mayo Foundation.

Paas, F., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, 38(1):63–71.

Paas, F. G. and Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human factors*, 35(4):737–743.

Paas, F. G. and Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational psychology review*, 6(4):351–371.

Park, T., Lee, U., Lee, B., Lee, H., Son, S., Song, S., and Song, J. (2013). Exersync: facilitating interpersonal synchrony in social exergames. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 409–422. ACM.

Park, T., Lee, U., MacKenzie, S., Moon, M., Hwang, I., and Song, J. (2014). Human factors of speed-based exergame controllers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1865–1874. ACM.

Pimenta, A., Gonçalves, S., Carneiro, D., Fde-Riverola, F., Neves, J., and Novais, P. (2015). Mental workload management as a tool in e-learning scenarios. In *Pervasive and Embedded Computing and Communication Systems (PECCS), 2015 International Conference on*, pages 25–32. IEEE.

Pisan, Y., Marin, J. G., and Navarro, K. F. (2013). Improving lives: using microsoft kinect to predict the loss of balance for elderly users under cognitive load. In *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*, page 29. ACM.

Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. tracking without keeping track of object identities. *Visual cognition*, 11(7):801–822.

Qin, H., Rau, P.-L. P., and Salvendy, G. (2009). Effects of different scenarios of game difficulty on player immersion. *Interacting with Computers*, 22(3):230–239.

Robertson, R. J. and Noble, B. J. (1997). 15 perception of physical exertion: Methods, mediators, and applications. *Exercise and sport sciences reviews*, 25(1):407–452.

Rubio-Valdehita, S., López-Higes, R., and Díaz-Ramiro, E. (2014). Academic context and perceived mental workload of psychology students. *The Spanish journal of psychology*, 17.

Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30(4):344–360.

Ryu, K. and Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, 35(11):991–1009.

Sinclair, J., Hingston, P., and Masek, M. (2009). Exergame development using the dual flow model. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment*, page 11. ACM.

Song, H., Peng, W., and Lee, K. M. (2011). Promoting exercise self-efficacy with an exergame. *Journal of health communication*, 16(2):148–162.

Staiano, A. E. and Calvert, S. L. (2011). The promise of exergames as tools to measure physical health. *Entertainment computing*, 2(1):17–21.

Strobach, T., Frensch, P. A., and Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta psychologica*, 140(1):13–24.

Sun, H. (2013). Impact of exergames on physical activity and motivation in elementary school students: A follow-up study. *Journal of Sport and Health Science*, 2(3):138–145.

Sween, J., Wallington, S. F., Sheppard, V., Taylor, T., Llanos, A. A., and Adams-Campbell, L. L. (2014). The role of exergaming in improving physical activity: a review. *Journal of Physical Activity and Health*, 11(4):864–870.

Sweetser, P. and Wyeth, P. (2005). Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3(3):3–3.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4):295–312.

Tozman, T., Magdas, E. S., MacDougall, H. G., and Vollmeyer, R. (2015). Understanding the psychophysiology of flow: A driving simulator experiment to investigate the relationship between flow and heart rate variability. *Computers in Human Behavior*, 52:408–418.

Van Muijden, J., Band, G. P., and Hommel, B. (2012). Online games training aging brains: limited transfer to cognitive control functions. *Frontiers in human neuroscience*, 6.

Velazquez, A., Martinez-Garcia, A. I., Favela, J., Hernandez, A., and Ochoa, S. F. (2013). Design of exergames with the collaborative participation of older adults. In *Computer Supported Cooperative Work in Design (CSCWD), 2013 IEEE 17th International Conference on*, pages 521–526. IEEE.

Wahlström, J. (2005). Ergonomics, musculoskeletal disorders and computer work. *Occupational Medicine*, 55(3):168–176.

Wickens, C. D. (2017). Mental workload: Assessment, prediction and consequences. In *International Symposium on Human Mental Workload: Models and Applications*, pages 18–29. Springer.

Xie, J., Xu, G., Wang, J., Li, M., Han, C., and Jia, Y. (2016). Effects of mental load and fatigue on steady-state evoked potential based brain computer interface tasks: a comparison of periodic flickering and motion-reversal based visual attention. *PloS one*, 11(9):e0163426.