TESTING THE REALITY OF EXERCISE PARTNERS AS A MODERATOR OF THE KÖHLER EFFECT

By

Samuel Thomas Forlenza

A DISSERTATION

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

Kinesiology – Doctor of Philosophy

2014
ABSTRACT

TESTING THE REALITY OF EXERCISE PARTNERS AS A MODERATOR OF THE KÖHLER EFFECT

By

Samuel Thomas Forlenza

Few people engage in recommended amounts of physical activity, and as such, it is important methods are sought out that could help people improve their motivation to sustain exercise. One potential way to do this is to leverage the power of group dynamics, specifically, the Köhler effect (Stroebe, Diehl, & Abakoumkin, 1996). When two people of moderately unequal abilities are partnered together on a conjunctive task, the weaker group member tends to perform better, and this performance gain is known as the Köhler effect (Kerr & Hertel, 2011).

A growing body of research supports the idea that the Köhler effect can be combined with exergames to produce improvements in the length of time people hold plank exercises (e.g., Feltz, Kerr, & Irwin, 2011). However, the research thus far has only made use of pre-recorded confederates purported to be live human partners. While this approach has been successful, it may not be practical on a large scale. Therefore, the purpose of this dissertation is to explore whether or not the Köhler effect can be achieved with different types of partners.

One approach is to use a computer-generated partner. Some research shows that people will team up with computers, suggesting that there is the potential for these types of interactions to improve motivation and performance (Nass, Fogg, & Moon, 1996). Another potential partner is an imagined partner. Research has shown that using associative imagery (imagery focused on the task) and dissociative imagery (imagery focused on something unrelated to the task) helps extend performance on muscular endurance tasks, suggesting that combining imagery and the Köhler effect could be an effective match (Razon et al., 2010).
To test this, participants (220 undergraduate students) were randomly assigned to one of six experimental conditions (virtually-real partner, altered virtually-real partner, software-generated partner, imagined partner, associative imagery, dissociative imagery) or to a no-partner, no-imagination control condition. Participants performed two blocks of plank exercises, holding each exercise for as long as possible. The first block was a baseline measure, performed individually and without imagery. For the experimental conditions, the second block was performed either with a partner or with the use of imagery. Measures of exercise duration, perceived exertion, self-efficacy beliefs, enjoyment, intention to exercise the following day, and impressions of one’s partner and group were collected.

Results from the main analysis indicated that participants in the experimental groups generally held plank exercises for longer during the second block compared to the control participants, supporting the idea that computer-generated partners, imagined partners, and imagery can be used to increase the time people hold plank exercises. Ancillary analyses indicated that participants in the partner conditions perceived themselves as working harder during the second block, while participants in the imagery and control conditions did not. There were no differences in self-efficacy beliefs, enjoyment, and intention to exercise the following day.

These findings are consistent with past research (Feltz et al., 2011) and suggest that incorporating computer-generated partners into exergames may boost how long people play these games. The use of imagined partners and imagery may also be effective methods to improve performance during traditional exercise. As a muscular endurance task was used, future researchers could use a similar paradigm to test these results with aerobic exercise, which would have a greater potential to positively influence health.
TABLE OF CONTENTS

LIST OF TABLES........................................................................................................... vii

LIST OF FIGURES........................................................................................................ viii

CHAPTER 1
INTRODUCTION............................................................................................................. 1
  Purpose and Hypotheses........................................................................................... 6
  Primary Research Question and Hypotheses....................................................... 7
  Secondary Research Question and Hypotheses................................................... 9
  Delimitations............................................................................................................ 10
  Definitions............................................................................................................. 11

CHAPTER 2
LITERATURE REVIEW.................................................................................................. 13
  The Problem of Physical Inactivity................................................................. 13
    The Benefits of Exercise..................................................................................... 14
    Meeting Physical Activity Recommendations............................................. 16
  Video Games: Problem and Solution?............................................................... 18
    Effects of Video Game Play................................................................................ 20
    Video Games as Solutions.................................................................................. 22
  Exergames............................................................................................................. 23
    Alternative to Exercise......................................................................................... 24
    Benefits of Playing Exergames.......................................................................... 27
    Improving Motivation......................................................................................... 32
  Group Dynamics and the Köhler Effect............................................................... 35
    Motivation Losses............................................................................................... 38
    Motivation Gains................................................................................................. 39
  Human-computer Interactions............................................................................. 53
    Media Equation................................................................................................... 53
    Uncanny Valley.................................................................................................. 56
  Cognitive Strategies During Exercise............................................................... 59
    Imagery................................................................................................................ 60
    Imagined Partners............................................................................................... 62
  Summary............................................................................................................... 63

CHAPTER 3
METHOD...................................................................................................................... 65
  Design and Participants....................................................................................... 65
  Experimental Task................................................................................................. 66
  Measures.............................................................................................................. 67
    Persistence........................................................................................................... 67
    Rating of Perceived Exertion (RPE)................................................................. 67
APPENDIX G: TEAM PERCEPTION INDEX.................................................. 125
APPENDIX H: GROUP IDENTIFICATION.................................................. 126
APPENDIX I: PARTNER MANIPULATION CHECK..................................... 127
APPENDIX J: IMAGERY MANIPULATION CHECK..................................... 128
APPENDIX K: INDIVIDUAL CONTROL MANIPULATION CHECK............. 130
APPENDIX L: DEMOGRAPHIC QUESTIONS............................................. 131
APPENDIX M: MANIPULATION SCRIPTS............................................... 132
APPENDIX N: SUPPLEMENTAL ANALYSES........................................... 139

REFERENCES.......................................................................................... 144
LIST OF TABLES

Table 1: Means and Standard Deviations of the Imagery Manipulation Check by Condition .......................................................... 83

Table 2: Correlations Between Difference Scores and Imagery Manipulation Check by Condition ............................................................................................................. 84

Table 3: Overall Means and Standard Deviations for All Dependent Variables.......... 89

Table 4: Means and Standard Deviations for All Dependent Variables by Condition ........ 91

Table 5: Correlations Between All Dependent Variables ........................................... 92

Table 6: Means and Standard Deviations for the Partner Conditions for Emotional Responses to their Partners and Attitudes Toward their Partners............... 96

Table 7: Correlations Between Difference Scores and Emotional Responses to their Partners and Attitudes Toward their Partners by Partner Condition ................. 97

Table 8: Correlations Between Subscales of the Godsspeed Indices.......................... 97

Table 9: Correlations Between Subscales of the Alternative Godsspeed Indices............. 98

Table 10: Summary of the Main Analysis, Alternative Follow-up Tests, and Alternative Analyses ................................................................. 142
LIST OF FIGURES

Figure I: A simplified version of Mori’s Uncanny Valley (1970/2005)………………… 57

Figure II: Pictures of the three male partners; from left to right, virtually-real (VR), altered virtually-real (AVR), and software-generated (SG)…………………………………… 74

Figure III: Mean Block 2 – Block 1 difference scores (in seconds) with 95% confidence intervals. Negative difference scores indicate reduced performance compared to baseline, while positive difference scores indicate greater performance compared to baseline…………………………………………………………………………. 90

Figure IV: Adjusted means for each condition of the unstandardized residuals (from the Block 2 on Block 1 regression) with 95% confidence intervals………………………… 141
CHAPTER 1

INTRODUCTION

One of the major contributors towards obesity is physical inactivity, both of which are associated with numerous health problems (Dixon, 2010). There are a multitude of factors that play a role in being physically active, which means that there are no easy ways to increase activity. Just as there are numerous health problems associated with obesity and inactivity, there are also numerous physical and mental health benefits to being physically active and physically fit (Garber et al., 2011). However, not enough people are meeting the basic guidelines for physical activity for a variety of reasons, including exercise programs being too intense (Dishman, 2001). Many who try to initiate an exercise regimen drop out, so it is imperative that researchers and practitioners alike develop strategies that are focused on helping people start and sustain physical activity programs.

One major barrier to participating in exercise is the lure of sedentary activities, such as television, internet use, and playing video games. Because these activities are highly enjoyable, many people choose to engage in them instead of physical activity (Epstein & Roemmich, 2001). When this happens, time that could have been spent exercising is instead spent sitting. There is some mixed evidence that video game play uniquely (i.e., independently of other sedentary activities) displaces physical activity, although there are wider concerns about the effects of video game play (Ballard, Gray, Reilly, & Noggle, 2009; Cummings & Vandewater, 2007). For example, playing violent video games may lead to both short-term and long-term problems with aggression, although those conclusions are disputed (Anderson et al., 2010; Ferguson, Olson, Kutner, & Warner, 2010).
Given all of this, it seems implausible that video games could actually help the problem of inactivity (Guy, Ratzki-Leewing, & Gwadry-Sridhar, 2011). However, exercise video games, or exergames, are a class of video game that has been growing rapidly in popularity in the past decade (Hawn, 2009). These types of active games require movement in order to play and succeed at the game (Staiano & Calvert, 2011). Put another way, these video games require players to engage in physical activity during game play, changing these traditionally sedentary activities into a movement-based activity. Thus, there is potential for exergames to help reduce inactivity.

Before exergames can be recommended as an intervention tool or as an alternative to exercise, it is important to determine if they are a worthy substitute (Daley, 2009). Most research on the matter suggests that exergames are the equivalent of light-to-moderate exercise for young adults and children, similar to walking at a moderate pace (Barnett, Cerin, & Baranowski, 2011; Warburton et al., 2009). Heart rate and energy expenditure from playing exergames is significantly higher compared to traditional video games, though playing exergames typically does not lead to vigorous physical activity (Biddiss & Irwin, 2010; Peng, Lin, & Crouse, 2011).

Similar to exercise, there are many benefits to playing exergames. Cognitive benefits have been observed, though more high-quality research is necessary (Anderson-Hanley et al., 2012a; O’Leary, Pontifex, Scudder, Brown, & Hillman, 2011). The observed physical and health benefits of playing exergames are small but significant, with some studies showing improvements in waist circumference and body mass in children (Maddison et al., 2011). However, a recurring problem with exergames is that any potential improvements are lost once their novelty wears off (Maloney et al., 2008). After a few weeks, the exergame becomes just another game, and many children stop playing them (Bethea, Berry, Maloney, & Sikich, 2012).
This may be less of a problem with young adults (e.g., Warburton et al., 2007), but more research with this population is needed. Thus, a major issue for using exergames is devising strategies to keep people playing them for longer.

Because exergames are a unique combination of exercise and video games, the issue of keeping people engaged in exercise can be approached from a video game perspective or an exercise perspective (Baranowski, Buday, Thompson, & Baranowski, 2008). While many strategies focus on one or the other, a potential way to include both is through the use of group dynamics. Many people play video games with friends, just as many people exercise with friends or in a group setting. The most relevant group dynamics research for this dissertation is the research that is focused on motivation. Though it has been well-documented that people can lose motivation from being in groups, there is also evidence that when certain conditions are present, people can gain motivation from being in a group (Karau & Williams, 1993; Williams & Karau, 1991). One particular motivation gain that shows promise for this problem is the Köhler effect (Witte, 1989).

To state it simply, the Köhler effect is the improvement in motivation or performance by the weaker group member when that person is paired with a superior partner under conditions where the group’s outcome is dependent on that weaker member (Köhler, 1926, 1927, as cited in Stroebe et al., 1996). In this way, an upwards social comparison (i.e., comparing oneself to the superior partner) and indispensability to the group (i.e., when one’s efforts significantly contribute to the group’s outcome) are the two major mechanisms through which the Köhler effect works (Kerr et al., 2007). A number of studies have looked at moderators of the Köhler effect, which include the discrepancy in partner ability, availability of feedback, gender, and physical presence (Weber & Hertel, 2007). The overall body of literature suggests that this effect
is quite strong and effective at improving performance in the weaker member. This is ideal for exercise settings where the inferior group member is essentially the least fit member, although this may be limiting for the superior group member, suggesting a need to explore alternative (i.e., non-real) types of partners.

The Köhler effect has been observed in athletic competitions (e.g., Hüffmeier & Hertel, 2011), but it has also been experimentally induced in exercise settings in combination with exergames (Feltz et al., 2011). The research suggests that the application of the Köhler effect leads people to engage in exercise longer without perceiving themselves to be working any harder and without enjoying the task any less (Feltz, Irwin, & Kerr, 2012; Irwin, Scorniaenchi, Kerr, Eisenmann, & Feltz, 2012). In other words, the Köhler effect helps people to exercise more without negative side-effects.

Although these are promising results, the research to date has only used ostensibly real or live human partners (e.g., Forlenza, Kerr, Irwin, & Feltz, 2012). In reality, these partners were all pre-recorded confederates. While useful for experimental research, it makes their applicability to the wider population and exergame design difficult at best for a number of reasons, such as the need to schedule live partners, the superior partner’s inability to receive a good workout, and the difficulty of controlling the optimal discrepancy between partners. Therefore, it would be beneficial to explore alternative types of partners that people could engage with to help boost their exercise behaviors.

One type of an alternative partner to a real person is a computer-generated one, which would be ideal for exergame design. These partners could be built into the game, removing the need to schedule real people. Also, their fitness level could be adjusted to match the fitness level of the participant. There is evidence that people apply social rules to human-computer
interactions, which includes teaming up to complete some task that requires interdependence between the partners (Reeves & Nass, 1996). But not all studies show people will team up with computers, instead ignoring their computer partner’s contributions to the task and finding those contributions less important (Kiesler, Sproull, & Waters, 1996). One reason for this may be the Uncanny Valley, or the feelings of discomfort people experience when they encounter a robot or computer agent who is simultaneously human-like but not human (Tondu & Bardou, 2011). As such, a design goal should be to avoid the valley to reduce any negative feelings from interacting with the computer (Mori, 1970/2005).

Another type of partner, one that would be useful to anyone in or out of exergame settings, is an imagined partner. There is evidence that many people already engage in imagery while they exercise for many reasons, including motivation (Giacobbi, Hausenblas, Fallon, & Hall, 2003; Hausenblas, Hall, Rodgers, & Monroe, 1999). Other research suggests that associative imagery, or imagery that is task-relevant (i.e., imagery that is focused on one’s body or the task one is performing, such as successfully performing the exercise) and dissociative imagery, or imagery that is task-irrelevant (i.e., imagery that is focused on something unrelated to one’s body or to the task, such as achieving a life goal) can both be used to improve performance on a physical persistence task (Razon et al., 2010). Given that the Köhler effect exergame research has used a physical persistence exercise task, it is plausible that these types of imagery can help improve performance. But more importantly, it is also plausible that people may benefit from exercising alongside a self-created, imagined partner, which has not been explored in the research. Coupling the Köhler effect with a superior imagined partner could be a powerful way for people to receive an extra boost of motivation while they are working out, helping them to persist at the exercise for a longer amount of time.
Purpose and Hypotheses

Improving exercise behaviors, particularly duration and intensity, is important for improving fitness. The Köhler effect is a promising tool for enabling this, but thus far, the research has been conducted with ostensibly real/live partners (actually, pre-recorded confederates). While effective, there are several limitations with this approach, and it could be impractical to implement on a large scale. Therefore, the purpose of this dissertation was to explore how alternative types of partners influence the Köhler effect in exercise settings.

The first two conditions in this dissertation replicated previous Köhler effect experiments. Participants in the individual control (IC) condition exercised individually (without a partner and without the use of imagery), while participants in the virtually-real (VR) condition exercised with an ostensibly live (but actually pre-recorded) person in a conjunctive setting.

The second two conditions utilized two types of computer-generated partners in conjunctive settings. The first partner was entirely software-generated (SG). This partner was modeled to look like a person, but clearly was not human. The second partner was a pixelated version of the virtually-real partner, or altered virtually-real (AVR). For both of these conditions, participants were told that they were working with a computer-generated partner programmed to be moderately superior in exercise ability, but still able to become fatigued like a real person.

The final three conditions were imagery conditions. Participants used either partner imagery (PI), associative imagery (AI), or dissociative imagery (DI) while they performed the exercises. For all three of these conditions, participants were given guidelines on what to imagine during the task. Participants who used PI were given similar instructions as participants in the VR, AVR, and SG conditions, with the exception that these participants were instructed to imagine themselves exercising conjunctively with an imagined partner instead of a real or
computer-generated partner. Participants using AI were instructed to focus on the task and imagine themselves holding the plank exercises successfully, while participants using DI were instructed to imagine themselves doing something motivational but unrelated to the task, such as successfully accomplishing a challenging goal. Also, note that the IC group received no imagery instructions, thus making it a no-partner, no-imagery control group.

**Primary Research Question and Hypotheses**

The first research question this dissertation explored was:

**RQ 1: What are the effects of working with computer-generated partners, working with imagined partners, and using imagery on exercise task performance?**

Previous research has shown that the Köhler effect can be elicited by the use of a virtually-real partner (i.e., one who is real but presented virtually) in exergame settings (Feltz et al., 2011). Research has also shown that people will partner with teammates and that the use of imagery can extend the duration of physical performance (Nass et al., 1996; Razon et al., 2010). As such, it seemed plausible that computer-generated partners, imagined partners, and imagery could elicit performance gains. Therefore, the first hypothesis was:

**Hypothesis 1: Participants in each of the six of the experimental conditions (VR, AVR, SG, PI, AI, DI) would significantly improve their exercise performance compared to participants in the control condition (IC).**

Although research has shown participants will team up and work with computers, the research is mixed on how effective computer-generated partners are compared to human partners (e.g., Kiesler et al., 1996). Computer-generated partners are also interesting in an endurance exercise setting. People may not feel it is informative to compare themselves with a computer-generated partner, and thus will not be as motivated by their partner’s superior ability. The indispensability
mechanism also may be attenuated or absent with a computer-generated partner if people do not feel that they are part of a team. Therefore, it was plausible that computer-generated partners would not lead to as large an improvement as human partners, which has been a robust finding in the literature (Weber & Hertel, 2008). Furthermore, while imagery has been shown to improve performance, imagined partners have not been used before and thus there is no evidence that imagined partners will lead to as much of an improvement in performance as compared to a real partner. Therefore, the second hypothesis was:

**Hypothesis 2: Participants exercising with a human partner (VR) would show significantly greater Köhler motivation gains effects compared to participants exercising with computer-generated partners (AVR, SG) or with an imagined partner (PI).**

Finally, although imagery has been shown to improve performance on physical persistence tasks, these results have not always been statistically significant (Razon et al., 2010; Razon, 2012). While improvements are hypothesized to occur, it was plausible that they would not be as great as when working with a partner. Additionally, due to the lack of prior research in this area, it was difficult to hypothesize how PI would compare to the other conditions. Given research has shown that teaming up with partners in these settings do improves performance (see Feltz et al., 2011), adding this group dynamics principle to imagery could improve its effectiveness above either AI or DI. Therefore, the third hypothesis was:

**Hypothesis 3: Participants in each of the four partner conditions (VR, AVR, SG, PI) would show significantly greater improvements on exercise performance compared to participants in the remaining two imagery conditions (AI, DI).**
Overall, then, it was hypothesized that the largest performance gains would occur in the VR condition, followed by the remaining three partner conditions (AVR, SG, PI), followed by the remaining two imagery conditions (AI, DI), all of which would be significantly greater than the no-partner, no-imagery control condition (IC).

Secondary Research Question and Hypotheses

The second research question this dissertation explored was:

*RQ 2: Do computer-generated partners, imagined partners, and imagery affect performance-related constructs, such as perceived exertion, self-efficacy beliefs, enjoyment, and intention to exercise in the future?*

Previous research with the Köhler effect has sometimes found that people do not perceive themselves as working harder despite the fact that they hold plank exercises for significantly longer (Feltz et al., 2011). Participants using associative and dissociative imagery, however, have reported lower perceived exertion compared to a control group, with dissociative imagery use leading to lower exertion levels (Razon et al., 2010). Therefore, the fourth hypothesis was:

*Hypothesis 4: There would be no significant differences between the four partner conditions (VR, AVR, SG, PI) and the control condition (IC) on perceived exertion. However, participants in the remaining two imagery conditions (AI, DI) would report significantly lower perceived exertion compared to the control condition (IC), with DI showing significantly lower perceived exertion than AI.*

Previous research with the Köhler effect and live partners has generally found that improvements in performance come at no expense to self-efficacy beliefs, enjoyment, and intention to exercise in the future, so it was plausible that these findings would be replicated with computer-generated and imagined partners (Feltz et al., 2011; Forlenza et al., 2012). While these constructs had not
yet been measured with associative or dissociative imagery, because it was hypothesized that performance would be greater while perceived exertion would be lower, it followed that self-efficacy beliefs, enjoyment, and future exercise intentions would be unaffected. Therefore, the fifth hypothesis was:

*Hypothesis 5: There would be no significant differences between the six experimental conditions (VR, AVR, SG, PI, AI, DI) and the control condition (IC) on self-efficacy beliefs, enjoyment, and intention to exercise the following day.*

Overall, the secondary hypotheses stated that there would be no differences between the experimental conditions (VR, AVR, SG, PI, AI, DI) and the control condition (IC) on all secondary measures, with the exception of perceived exertion, which would be lower for the imagery groups (AI, DI), with DI showing the greatest reduction of exertion.

**Delimitations**

The participants for this dissertation were undergraduate students at a large Midwestern university who were mostly fit or within a normal weight range (as measured by body mass index). For this reason, the results of this dissertation may not generalize well to other populations, such as people who are sedentary or overweight. Similarly, college students are perceived as being more “wired” and familiar with technology. As a result, their willingness to team up with computer-generated partners may be different from other age groups, such as middle-aged or older adults. Additionally, these results may not be replicated with other types of group task demands (Steiner, 1972), such as with additive or coactive tasks. The computer-generated partners may only be limited to the type of partner that was created for this study. More or less sophisticated partners may elicit different emotional reactions from people, which
may in turn affect performance. Finally, plank exercises were used in this dissertation, and therefore the results may not generalize to other tasks.

**Definitions**

- **Association (cognitive strategy):** while exercising, directing one’s attention towards task-relevant cues, such as one’s own bodily sensations, in order to manage those cues to persist at the exercise (e.g., managing pain level in legs by slowing pace).

- **Conjunctive tasks:** tasks where the group’s outcome is determined by the least capable or weakest member (e.g., mountain climbers tethered together can only go at the pace of the slowest climber).

- **Dissociation (cognitive strategy):** while exercising, directing one’s attention away from the task in order to distract oneself to persist at the exercise (e.g., focusing on what to make for dinner when pain is experienced).

- **Exergames:** exercise video games (or active video games) that require physical movements in order to be played successfully.

- **Indispensability:** how crucial one’s contributions are for the group’s success.

- **Köhler discrepancy effect:** a moderating factor of the Köhler effect; compared to a small or large discrepancy between partners, a medium discrepancy leads to the largest gains in motivation and performance.

- **Köhler motivation gain effect:** the increase in motivation and performance exhibited by the least capable member of a group when performing a task with a superior partner (compared to performing the task individually); alternatively referred to as simply the Köhler effect.
− *Media Equation*: the mindless application of social rules/scripts from human-human interactions to human-computer interactions.

− *Social comparison*: the process whereby people evaluate their own ability or performance by comparing it to the people around them (e.g., fellow group members).

− *Uncanny Valley*: the continuum of affective responses to humanoid robots and computer-generated people; when robots look simultaneously human-like and non-human, they fall into the Uncanny Valley, reducing feelings of familiarity and increasing feelings of strangeness; movement exaggerates these negative feelings.
CHAPTER 2

LITERATURE REVIEW

The purpose of this chapter is to review the literature relevant to the topics of focus of this dissertation. The chapter begins with an overview of the problem physical inactivity before discussing the role video games play in this problem. From there, it is argued that video games can be a way to help the situation via exergames. The literature on exergames is reviewed to explore if they are a possible solution before moving to how group dynamics principles can be incorporated into exergames. Specifically, it is argued that the Köhler effect is a promising approach to take to improve exercise behaviors. Finally, the chapter concludes with a review of human-computer interactions and cognitive strategies associated with exercise to explore the potential viability of using computer-generated partners, and imagery (including imagined partners) to improve exercise duration.

The Problem of Physical Inactivity

Over the past few decades, obesity in the United States of America (and globally) has become a major health crisis, with the number of obese Americans rising dramatically between 1980 and 2000, though recent research suggests the obesity level has plateaued over the past 10 years (Flegal, Carroll, Kit, & Ogden, 2013). Still, the obesity rate remains high (35.5% for men, 35.8% for women) and is even higher when overweight individuals are included (73.9% for men, 63.7% for women) (Flegal et al., 2013). More people than ever are carrying an excess accumulation of fat, which is associated with a variety of physical and psychological health problems (e.g., type 2 diabetes, coronary heart disease, certain cancers, depression, low self-esteem) (Dixon, 2010). Although the rising rate of obesity seems to have leveled off, there is no
evidence that the rates have declined, and given the negative consequences, there is still a major need for interventions.

Increases in inactivity and sedentary behaviors have played a large role in the high levels of overweight and obesity. During the same time period that obesity levels were rising, the total amount of physical activity in which people engaged was dropping. Rates of leisure time physical activity (exercise, sports) have remained fairly consistent, but the amount of physical activity in other areas of life has dropped significantly (Brownson, Boehmer, & Luke, 2005). For example, over the past 50 years, the economy shifted away from agricultural and goods-producing jobs (which typically require moderate amounts of physical activity) towards service jobs (which typically require light or negligible amounts of physical activity). Indeed, it is estimated that in 2008, less than 20% of jobs required moderate activity, a drop of nearly 30% (Church et al., 2011). This translated to over 120 fewer calories expended throughout the working day, which accounted for a large portion of the increase in weight (Church et al., 2011). Other research has suggested that in addition to a reduction of high- and moderate-activity jobs, the opportunities to engage in sedentary behaviors have risen (e.g., driving versus walking, spending money on sedentary activities versus sports and recreation), and people have generally taken these opportunities (Brownson et al., 2005). The cumulative result of the declining rates of everyday activity, increases in caloric intake, and unchanging levels of leisure time physical activity was a steady increase in the number of Americans who were overweight and obese, to the levels seen today.

The Benefits of Exercise

Physical benefits. Given that occupational physical activity has declined and sedentary activities have become more common, it stands to reason that one point of intervention is
boosting the amount of leisure time physical activity and exercise in which people engage. Doing so will provide more opportunities for energy expenditure and weight management, which could partially offset the deleterious effects of our sedentary lifestyle. However, just as obesity is associated with a number of health problems, exercise is associated with a number of health benefits well beyond weight control.

Over 60 years of research has shown that people who exercise on a consistent basis are at a lowered risk for a variety of ailments, including type 2 diabetes, strokes, heart attacks, some cancers, osteoporosis, and obesity (Blair & Morris, 2009). These beneficial effects of exercise apply to all people, regardless of age, ethnicity, gender, or current fitness level (Garber et al., 2011; Janssen & Leblanc, 2010; Vogel et al., 2009). It seems clear that increasing the number of people who exercise and/or increasing the intensity of exercise should help to lower the rates of obesity and the associated health costs.

However, there is evidence that the primary emphasis need not be on reducing weight, but on improving physical fitness. For example, in a study that focused on all-cause mortality and mortality from cardiovascular disease in men, it was found that men who were fit and remained fit (over an average span of 5 years) had the lowest death rates (Blair et al., 1995). For men who improved their fitness levels, there were significant reductions in all-cause mortality (44%) and cardiovascular disease mortality (52%), including in men who were already fit to begin with (15%). More recent research corroborates these findings. Despite significant differences in the amount of weight lost, both “responders” (those who lost weight from participating in an intensive aerobic exercise program) and “non-responders” (those who lost less weight than predicted by the amount of energy expended during program participation) groups showed significant improvements in aerobic capacity, blood pressure, resting heart rate, and
waist circumference, all of which benefitted their health (King, Hopkins, Caudwell, Stubbs, & Blundell, 2009). Therefore, this evidence suggests that the emphasis of exercise programs should be on fitness, not just fatness, as regular exercise improves health even when weight is not lost.

**Cognitive benefits.** In addition to physical and health benefits of exercise and physical fitness, there are numerous benefits for cognition. The short-term or acute effects of moderate aerobic exercise include improved decision making, increases in the speed and efficiency at which cognitive tasks are performed, feelings of positive mood, and reduced symptomatology of psychological disorders (Garber et al., 2011; Tomporowski, 2003). Other research suggests that physical fitness is associated with small, but positive improvements in cognitive performance (compared to a state of non-fitness), which suggests that chronic exercise that improves fitness will have psychological benefits (Etnier et al., 1997). Similar to the physical benefits of exercise, the cognitive benefits of exercise apply to people of all ages, genders, and ethnicities (Garber et al., 2011; Hillman, Erickson, & Kramer, 2008; Vogel et al., 2009). Overall, then, it is clear that there are numerous health benefits – both physiological and psychological – to engaging in regular amounts of exercise.

**Meeting Physical Activity Recommendations**

In order to achieve the physical and mental benefits of exercise, the U.S. Department of Health and Human Services (DHHS) (2008) (along with several other organizations) recommends that every adult should engage in 150 min per week of moderate aerobic physical activity, or 75 min per week of vigorous aerobic activity, or some combination (e.g., 40 min of vigorous activity and 70 min of moderate activity) (Garber et al., 2011). Additionally, the DHHS recommends adults engage in exercises that build muscular endurance, muscular strength, and
flexibility at least 2 days each week. Consistently meeting these guidelines will lead to short-
term and long-term benefits.

However, as evidenced by the unchanging rate of leisure time physical activity and
increases in obesity, not enough people are meeting these guidelines. Indeed, adopting and
adhering to exercise programs is difficult, as nearly half of all people who start an exercise
program drop out within the first few months, particularly when the exercise intensity increases
(Dishman, 2001). This is problematic as the health benefits of exercise are greater for higher-
intensity exercise (Swain & Franklin, 2006). There are several reasons or barriers for why
children and adults do not engage in enough exercise, including a perceived lack of time,
tiredness, poor weather, a lack of facilities, poor body image, pain, and a lack of social support
(Korkiakangas, Alahuhta, & Laitinen, 2009; Trost, Owen, Bauman, Sallis, & Brown, 2002;

Another major reason people do not engage in enough physical activity each week is the
lure of sedentary behaviors. All people have limited time and must therefore choose activities in
which to participate. The more enjoyable and desirable an activity is, the more likely people are
to participate in that activity over a less enjoyable alternative. As such, engaging in physical
activity (a hard, strenuous activity) often means choosing it over other appealing sedentary
activities (that are easy and relaxing), such as watching television, reading a book, or playing
video games (Epstein & Roemmich, 2001).

Activity choice depends on many factors, including how enjoyable an activity is and how
easy it is to engage in. Therefore, to make the choice for physical activity more likely, one can
make it easier to participate in, or make sedentary activities more challenging to participate in (or
alternatively, attempt to have the behaviors complement each other, as in watching television
while riding a stationary bike) (Epstein & Roemmich, 2001). From this perspective, people do not engage in physical activity as often as recommended because it is easier for them to engage in sedentary activities, plus these sedentary activities often are seen as more enjoyable.

Therefore, overcoming the barriers to exercise and the lure of sedentary activities is crucial if people are to exercise more than they are currently. One way to help people overcome these barriers is to build their exercise self-efficacy beliefs. Self-efficacy is a person’s specific belief about their confidence to perform a given task (in this case, exercise) (Bandura, 1997). A positive relationship between exercise self-efficacy and physical activity has been observed in the literature consistently (Trost et al., 2002). While there are several different ways to build exercise self-efficacy, one particularly effective approach has been to leverage the social influence of groups and significant others. Research has shown that as social support increases (from family and important others), so too do self-efficacy beliefs and the effectiveness of the exercise intervention, and similar results are obtained when people exercise in groups (Burke, Carron, Eys, Ntoumanis, & Estabrooks, 2006; Carron, Hausenblas, & Mack, 1996). Overall, then, using groups may be a motivating and effective way to help engage in physical activity instead of sedentary activity.

**Video Games: Problem and Solution?**

One of these easily enjoyable sedentary activities that people of all ages (children and adolescents in particular) choose to engage in on a regular basis is playing video games. Indeed, the Pew Internet and American Life Project estimates that 53% of American adults play video games, with 21% playing games every day or almost every day (Lenhart, Jones, & Macgill, 2008). Similarly, the Entertainment Software Association (ESA) estimates that 51% of households own a dedicated video game console (ESA, n.d.). Other research suggests those
figures are even higher for youth, estimating that 97% of teenagers play video games (Lenhart et al., 2008). These numbers show that many children, young adults, and even adults play video games on a regular basis, which may contribute to a lack of physical activity.

Playing video games is thought to displace time that people can spend engaging in other activities, such as exercise and sports. This would make intuitive sense: if a child chooses to stay indoors and play video games instead of going outside to play, playing video games is displacing physical activity. Some evidence does support this: research has found a positive relationship between playing video games (particularly Massively Multiplayer Online Role Playing Games, or MMORPGs) and having a higher body mass index (BMI), and a negative relationship between playing these games and frequency and duration of exercise (Ballard et al., 2009).

However, other findings are less convincing, making the evidence for this displacement effect much less straightforward. Ballard and colleagues (2009) also found that the frequency of playing sports video games was significantly and positively related to the frequency of exercise and number of days engaged in vigorous physical activity, suggesting the displacement effect may only be true for certain types of video games (i.e., for MMORPGs, but not for sports games). A different study analyzed time diaries of over 1,400 adolescent boys and girls and found very small effects: male gamers spent 8 min less on sports and active leisure on weekend days compared to male non-gamers, but there were no significant differences on weekdays, and no effects at all for female gamers (Cummings & Vandewater, 2007). An earlier study with children under the age of 12 years found that moderate video game use was associated with higher weight in children, but not light or heavy video game use (Vandewater, Shim, & Caplovitz, 2004). This study did find a significant linear relationship between overall time spent
in sedentary activities and weight status, but this measure included all sedentary activities, not just video game play.

Overall, it seems safer to say that sedentary activities displace physical activity without the need to single out video game play. While video game play is a sedentary activity and contributes to the lack of physical activity, it is not the only sedentary activity that does. Still, the point remains that video games are highly enjoyable and easily accessible for most of the population, which means that time to play video games will compete with time to engage in moderate and vigorous physical activity.

Effects of Video Game Play

In addition to the worry that video games displace physical activity and contribute to the overly-sedentary lifestyle many Americans live, there is a widespread concern about the behavioral and psychological effects of playing video games. Are video games leading to shorter attention spans? Do playing violent video games make people aggressive? Can there be any good to playing video games, or are they at best just a diversion?

Positive effects. There is mounting evidence that video games have beneficial effects for cognition, although more longitudinal studies are needed to determine evidence of causality in real-world settings. Reviews of research have shown that playing video games leads to improvements in a variety of cognitive constructs, including working memory, visual attention, selective attention, spatial ability, and auditory perception (Barlett, Anderson, & Swing, 2008; Barlett, Vowels, Shanteau, Crow, & Miller, 2009). Other research suggests that expert gamers outperform novice gamers on a variety of cognitive tasks tapping several cognitive constructs (e.g., short-term memory, object tracking, task switching, and mental rotation) (Boot, Kramer, Simons, Fabiani, & Gratton, 2008). It is also speculated that video game play is associated with
improvements in hand-eye coordination, faster reaction time, communication ability, subjective well-being, and positive mood (Barlett et al., 2008). In short, there appears to be a number of cognitive benefits to playing video games, although more experimental research is necessary before firm conclusions can be drawn (e.g., Irons, Remington, & McClean, 2011).

**Negative effects.** There is also a need for more research on the negative effects of playing video games, as much of it is still up for debate. Some research suggests that playing video games has deleterious effects on schoolwork. Playing video games was found to be related to lower grades in children (but was not related to BMI) (Jackson, von Eye, Fitzgerald, Witt, & Zhao, 2011). Time diary analysis has shown that adolescent gamers spend slightly less time reading and doing homework compared to non-gamers, suggesting some displacement effects exist (although the authors note the data are correlational – gamers could simply be faster readers and do their homework more efficiently) (Cummings & Vandewater, 2007). Additionally, review and meta-analytic studies suggest that there is considerable evidence that playing violent video games leads to short-term and long-term increases in aggressive behavior, aggressive cognitions, and aggressive affect, and is associated with a decline in empathy and prosocial behavior (Anderson & Bushman, 2001; Anderson et al., 2010; Barlett et al., 2008).

However, research has also shown the opposite findings. Undergraduates who played video games for a moderate amount of time had higher grades than those who played lightly or heavily, which suggests there is not a straightforward relationship between video games and school achievement (Ventura, Shute, & Kim, 2012). Even the findings on the relationship between violent video games and aggression are disputed. One study of over 1,200 middle school children found that even though 81% had played violent video games, delinquent behaviors and bullying were predicted by trait aggression and stress, not by exposure to violent
games (Ferguson et al., 2010). Additionally, the overall review and meta-analytic findings of relationships between violent video games and aggression have been criticized for a number of reasons, such as very small effect sizes and a publication bias (Ferguson & Kilburn, 2010). Clearly, there is a need for additional research to determine what effects – both positive and negative – playing video games has on children, adolescents, and adults, as much of it is contradictory.

**Video Games as a Solution**

In short, there is evidence both for and against video games being a unique contributor to the obesity epidemic (Ballard et al., 2009; Vandewater et al., 2004). There is little doubt that most Americans are engaging in too much sedentary behavior, which includes playing video games. Exactly how much video game play is displacing physical activity is unclear, especially compared to other sedentary behaviors. More research is needed to determine the positive and negative behavioral and cognitive effects of playing video games.

Given all of this, it would not be surprising to come to the conclusion that video games are part of the problem with inactivity and other behavioral issues. What may be surprising is the thought that video games could be part of the solution, but many people believe video games can be used to promote healthy behaviors. However, given how ubiquitous video games are in today’s society – Nielsen estimates that 97 million Americans use a video game console each month (Nielsen, 2013) – they represent an opportunity for intervention. Recent reviews argue that because most children and adolescents have a video game console in their bedroom, there is a great potential to turn sedentary screen time into educational and physically active screen time (Guy et al., 2011). And organizations such as the National Institutes of Health and the Centers for Disease Control have funded the development of health games (Hawn, 2009). Indeed, these
health games that are focused on improving education and changing behavior are already showing modest benefits, such as increasing physical activity and improving nutritional knowledge, and these games are still relatively early in their development (Hawn, 2009). By harnessing the widespread enjoyment of video games and combining it with research, theory, and knowledge, there is the potential for video games to be part of the solution, not just part of the problem.

**Exergames**

Exercise video games, or exergames, are video games that require movement in order to be played. The first exergames were developed in the 1980s and were designed to hook up to stationary bikes (Staiano & Calvert, 2011). However, these games were too expensive and too complicated for the average person, so they were largely unused. But significant advancements have occurred since then that have made exergames much less costly and therefore much more accessible. Games can now be built into stationary bikes directly, removing the need to buy and build new equipment. Current generation video game consoles take advantage of motion sensor technology and small mounted cameras for full body tracking (Staiano & Calvert, 2011). Modern exergames are able to capture a person’s image and project them directly into the screen, allowing for movement-based interaction with and feedback from the game environment. Considering this technology is relatively new to video game consoles, it can only be refined and improved going forward.

Because of the drop in price and ease of use, more people of all ages are able to play exergames now than ever before. However, there is a general lack of data as to how many people are actually playing exergames. One study found that 24% of Montreal-area adolescents reported playing exergames, and that exergamers were more likely to be female (O’Loughlin, Dugas,
Encouragingly, those who reported playing exergames reported playing on average twice per week, for 50 min each bout, at a moderate or vigorous intensity. This suggests that, at least for the Montreal area youth, exergaming is a popular activity that can make contributions towards reducing sedentary time. Another study found that 63% of high school students in the United States played active video games in the past year (although intensity and frequency were not measured) (Brener et al., 2013). Clearly more research is needed to understand who is purchasing and playing these games (particularly in the adult population), how often they are being played, and for what reasons.

**Alternative to Exercise?**

Although exergames are much more common nowadays, there are still questions about their use. Perhaps the most important question surrounding exergames is, “Are exergames viable alternatives to traditional exercise such that they can help reduce sedentary time?” One way to begin to answer this question is to examine how comparable a bout of exergame play is to a bout of sedentary game play and to a bout of traditional exercise. Several studies, reviews, and meta-analyses have explored this question and generally have found promising results in both young adults and children.

Overall, research with young adults has shown that playing exergames is comparable to light-to-moderate physical activity and may even be effective at eliciting higher intensity exercise. Unsurprisingly, compared to playing sedentary video games or resting, playing exergames elicits higher heart rate and energy expenditure (Bonetti, Drury, Danoff, & Miller, 2010; Haddock et al., 2012; Kretschmann, Dittus, Lutz, & Meier, 2012; Leatherdale, Woodruff, & Manske, 2010). The amount of energy expended by young adults while playing exergames is generally comparable to light-to-moderate or moderate physical activity (Haddock et al., 2012;
Siegel, Haddock, Dubois, & Wilkin, 2009). Exergame play may even meet the requirements for vigorous physical activity when gameplay is structured to mimic interval training (Noah, Spierer, Tachibana, & Bronner, 2011).

Similar findings are reported with young adults when comparing exergame play to exercise directly. Compared to a 30 min session of treadmill walking, participants who played Dance Dance Revolution (DDR) (Konami of America) for 30 min had similar levels of oxygen consumption and perceived exertion (Leininger, Coles, & Gilbert, 2010), while heart rate and perceived exertion were higher when playing Wii Fit (Douris, McDonald, Vespi, Kelley, & Herman, 2012). Two studies have compared riding a bike ergometer to riding a bike ergometer with an exergame. One of these experimental studies reported that at matched sub-maximal low-to-moderate workloads, more energy was expended and heart rate was higher when playing the exergame compared to biking without the game (Warburton et al., 2009). The other study focused on self-selected intensities to better mirror a home exercise environment and found that biking with the exergame led to higher heart rates and levels of perceived exertion (Kraft, Russell, Bowman, & Selsor, 2011). In short, for young adults, playing exergames seems to be comparable to light-to-moderate physical activity, but may also be a good vehicle for increasing the intensity of the performed exercise.

Experimental studies with children corroborate these findings. When children aged 10 to 14 played five different exergames for the Sony PlayStation®2, results showed that compared to playing a sedentary game, playing exergames led to significant increases in energy expenditure and heart rate (Maddison et al., 2007). Similarly, when children were allowed to play six different exergames (three commercial games, three consumer games), playing them led to significant increases in energy expenditure compared to rest for all of the games (Bailey &
McInnis, 2011). Another study had 20 overweight children ride a stationary bike on two separate occasions for 20 min each time (Haddock, Siegel, & Wikin, 2009). On one occasion, children were able to play a game that was hooked up to the bike, while on the other occasion, children simply pedaled without a game. The results showed that oxygen consumption, heart rate, and perceived exertion were higher in both the exergame and exercise condition compared to rest, while energy expenditure was slightly higher in the exergame condition. Taken together, these studies suggest that playing exergames is equivalent to light-to-moderate activity, with some children playing them at even higher intensities. Other experimental research suggests comparable findings to the four reviewed here (e.g., Mellecker & McManus, 2008; Straker & Abbott, 2007).

Review studies and meta-analyses also support this conclusion. The average amount of energy expenditure during exergame play in children across the experimental literature is around 3.3 METs, which is equivalent to light-to-moderate activity (Barnett et al., 2011; Biddiss & Irwin, 2010). None of the reviewed articles used exergames that reached 6.0 METs, suggesting exergames are not a source of vigorous activity for youth (Barnett et al., 2011; Biddiss & Irwin, 2010). Games that utilize both upper and lower body movements (i.e., whole body movements) lead to higher energy expenditure in children and adults (Biddiss & Irwin, 2010; Peng et al., 2011). These increases in energy expenditure, heart rate, and oxygen consumption are significantly higher than playing sedentary video games, are comparable to physical activity, and are high enough to elicit health benefits (Peng et al., 2011; Warburton, 2009). Therefore, it seems that exergames encourage players to work hard enough to burn similar amounts of energy as light or moderate physical activity, such as brisk walking. This would suggest that exergames are potentially a promising alternative to exercise for both adults and children.
Benefits of Playing Exergames

**Cognitive benefits.** Because of these promising results, many researchers have attempted to use exergames as an intervention to improve cognitive functioning and physical health. Studies that have explored the cognitive benefits of exergame play have used participants of all ages and found generally favorable results, although this research is still very new. Research with children aged 6 to 10 years showed significant improvements in executive functioning following a bout of exergame play (compared to watching a movie or playing a sedentary game) by way of greater visuospatial efficiency (Best, 2012). Similar improvements in executive functioning were also seen when autistic children played exergames (Anderson-Hanley, Tureck, & Schneiderman, 2011). One study of overweight adolescents showed that playing exergames once per week for 10 weeks led to improvements on a test of executive functioning (compared to a control group), although this was only true for the adolescents who played the exergames competitively (versus cooperatively) (Staiano, Abraham, & Calvert, 2012).

The results are less clear when participants are college students. One study showed that a mere 10 min of exergame play led to significant improvements on two different cognitive tasks compared to playing a sedentary version of the same game (Gao & Mandryk, 2012). The other study including college students, however, showed that despite similar increases in heart rate between the exergame and exercise conditions, only exercising led to significant improvements on a measure of cognitive control (O’Leary et al., 2011).

Research with older adults also tends to show improvements in cognitive functioning, which is particularly important for that population due to natural declines in cognition. The most rigorous study with older adults used a randomized control trial to show that riding a stationary bike with an exergame led to significantly greater improvements on three measures of executive
functioning compared to traditional stationary biking (Anderson-Hanley et al., 2012a). These greater improvements occurred for exergames despite both groups exerting similar amounts of physical effort, suggesting that the additional improvements of exergaming occurred due to the interactivity of the exergame, which required more mental exertion than simply pedaling on a stationary bike. This pattern also held up (though not as strongly) with a sub-set of older adults who had type 2 diabetes (Anderson-Hanley, Arciero, Westen, Nimon, & Zimmerman, 2012b).

It appears that there are cognitive benefits to be reaped from playing exergames. Given that exergames are generally the equivalent of light-to-moderate exercise, this is somewhat unsurprising. Previous research has established the cognitive benefits of exercising (e.g., Etnier et al., 1997; Garber et al., 2011; Tomporowski, 2003) and the potential cognitive benefits of playing video games (Barlett et al., 2008; Barlett et al., 2009), so it would make sense for exergaming (a combination of exercise and video games) to lead to cognitive improvements. More research is necessary to replicate and extend these findings, however, as this line of study is still very new.

**Physical benefits.** The next area to address is whether or not exergaming can lead to significant improvements in physical health, particularly in the areas of weight, fitness, and physical activity levels. The evidence supporting this is mixed, as some studies show exergame interventions lead to significant improvements, while others do not. And unlike the research that has explored the cognitive benefits of exergaming, the literature on the physical benefits is primarily focused on children.

One particularly promising study with youth is a large, randomized control trial that took place in New Zealand over a 6 month span. Over 300 children between the ages of 10 and 14 years were randomly assigned to either a control group or to an intervention group, which
received an active video game package consisting of three different exergames (Maddison et al., 2011). The children were encouraged to meet daily physical activity recommendations by substituting exergames for inactive games and activities, such as watching television. By the end of the 6 month intervention, the exergame group showed reductions in BMI (-0.24), percent body fat (-0.83%), and waist circumference (-1.21 cm) (Maddison et al., 2011). Although these reductions were small, they suggest a slowing of weight gain as children grew. Secondary analyses indicated that these improvements were likely the result of improvements in aerobic fitness and a decline in snack food consumption (Maddison et al., 2012; Maddison, Jull, Marsh, Direito, & Ni Mhurchu, 2013). Similarly, the 12 week pilot study for this randomized control trial showed a series of trends: children in the exergame group played video games for less time, engaged in more physical activity, and reduced their body weight and waist circumference (Ni Mhurchu et al., 2008). Although the changes were small, they showed that exergames can be one tool (among many) to help slow or reverse weight gain in children.

Other research has shown favorable effects resulting from exergame play in children. Overweight children were given DDR and instructed to use it in their homes 5 days per week for 12 weeks (Murphy et al., 2009). Most of the children (75%) complied with these instructions (as determined by the assessment of daily logs) and this resulted in a host of physiological and behavioral changes, including improvements in mean arterial pressure, aerobic capacity, time spent exercising, and less weight gain compared to a delayed-treatment control group. A similar intervention (though without a control group) found that playing DDR led to improvements in physical fitness after 12 weeks, and these improvements were maintained through the end of the study (at 30 weeks), with many children still playing the game (Bethea et al., 2012).
In addition to home-based interventions, exergames have been used in community- and school-based interventions, with some success. A pediatric weight management program incorporated exergaming and traditional exercise 1 hr per week for 10 weeks (Christison & Khan, 2012). Exergaming was only one component of the program, with the others being focused on nutrition education and psychosocial behavior change strategies. Ultimately, the overweight children had a lower BMI, were exercising for more time, had improved self-worth, and were drinking less soda and watching less television, which are promising results after only 10 weeks. A 2 year school-based intervention showed that DDR and traditional exercise led to improvements in physical fitness and improvements in math scores, although it was not possible to determine the effect of one mode of exercise versus the other (Gao, Hannan, Xiang, Stodden, & Valdez, 2013). This study showed that turning unstructured recess time into structured physical activity time can lead to physical and academic benefits.

Other studies, however, have had much less success or no success at all. One study utilized a peripheral active game that encouraged children to take steps while playing traditional video games (Graves, Ridgers, Atkinson, & Stratton, 2010). After 12 weeks, there were no significant changes in body fat or BMI, plus both active game time and sedentary game time were increased, suggesting the intervention inadvertently led to an increase in sedentary video gaming.

Although there are some exceptions (e.g., Bethea et al., 2012), a recurring problem with many exergaming studies is that any initial success is difficult to sustain over time, particularly once the novelty of the games wear off. Another home-based DDR intervention found that DDR use peaked during the first week of the study, with participants playing an average of 145 min per week (Maloney et al., 2008). After 10 weeks, the average play time per week had dropped
down to 89 min despite over 90% of parents and youth stating they enjoyed playing the game. This intervention also used weekly coaching sessions during the first 5 weeks that focused on principles of motor learning (Errickson, Maloney, Thorpe, Giuliani, & Rosenberg, 2012). Children who had weekly coaching sessions played DDR more, but these increases disappeared once the coaching sessions ended. Similar studies have shown that giving children exergames with little or no instruction on how much to play the games or how to play them do not engage in any more physical activity than children without the games (Baranowki et al., 2012; Madsen, Yen, Wlasiuk, Newman, & Lustig, 2007). Indeed, even in the one study that showed children still played the exergame after 30 weeks, nearly 42% of children had lost interest by 12 weeks (Bethea et al., 2012).

A notable exception to the above problems with exergame research in children took place in a study with 14 sedentary young adult males (Warburton et al., 2007). In this study, college-aged males were randomly assigned to either bike with or without an exergame for 30 min per day, 3 days per week, for 6 weeks. Adherence to the training program was higher for the exergame group (78%) compared to the traditional cycling group (48%). Additionally, participants in the exergame condition significantly improved their VO₂max by 11% and reduced their systolic blood pressure by 9 mmHg. In the traditional cycling group, there was no improvement in aerobic fitness and the reduction of systolic blood pressure (by 3 mmHg) was significantly smaller compared to the exergame group. Attendance mediated the relationship between condition and the health outcomes of VO₂max and systolic blood pressure. This study showed that exergames improved adherence to the training protocol, which in turn resulted in greater physiological improvements and suggests that for young adults, exergames may be used to keep people exercising enough to receive health benefits.
Overall, then, the physical benefits of exergaming appear to be small, but positive. Exergaming is light-to-moderate physical activity, so it would follow that regularly playing exergames leads to fitness and health improvements, which the research generally shows. However, once the uniqueness of exergames wears off, children play these games less, or not at all, which means any improvements in fitness and health will be lost. And some exergames may not work at all. For these reasons, it seems unwise to recommend exergames as an alternative to exercise in children (Daley, 2009). However, many of these problems with exergames are similar to problems that the general population has with starting and sustaining exercise routines. Therefore, one could view the lack of long-term success with exergames as an issue of reduced motivation.

**Improving Motivation**

Compared to traditional exercise, exergames have some unique characteristics that allow for different motivational strategies to be used to increase engagement. One strategy is relatively simple: purchase more games. The most successful exergame interventions tended to use multiple games, not just one game (Lu, Kharrazi, Gharghabi, & Thompson, 2013). The obvious drawback to this strategy is cost, although the price of an exergame is likely to be less than the cost of a monthly gym membership. Still, an advantage to exergames is that they are a one-time cost, so it would be more beneficial to determine the best ways to increase attraction to exergames and to improve motivation to maintain exergame play so as to reduce the number of game purchases (Baranowski, Baranowski, Thompson, & Buday, 2011).

Motivational strategies to increase exergame play can come from two sources: the exergame itself (via improved game design) or the environment of the player (e.g., playing with someone else). It seems that the sole focus of virtually all exergames is on exercising
(Baranowski et al., 2008). Part of the reason video games (and other sedentary activities, like reading a book and watching a movie) are engaging and enjoyable is because they tell a story. So why not include stories in exergames to make them more appealing?

Including stories in exergames is something that some researchers are attempting to do. *Escape from Diab* is a health video game designed to lower the risk of obesity and diabetes in at-risk children and adolescents by improving diet and physical activity levels (Lu, Thompson, Baranowski, Buday, & Baranowski, 2012). After playing the game for 2 months, children showed a number of favorable changes, including feeling more confident in their own ability to eat fruits and vegetables and to exercise. Additionally, there were moderate correlations between immersion (how engaged or absorbed one is in an activity) and these outcomes, suggesting that stories enhance immersion in games, which also enhances the message they are sending (Lu et al., 2012). While this is not an exergame, it is a great example of how stories can be included in health games.

Another way to improve game design to increase adherence is to build in-game elements in accordance with existing psychological/behavioral theories. Researchers did exactly this when designing the exergame *Olympus* (Peng, Lin, Pfeiffer, & Winn, 2012). In addition to having a story (players were citizens of ancient Greece, training for athletic competitions), the exergame incorporated elements from self-determination theory (Deci & Ryan, 1987). Specifically, the game systematically altered various elements of the game to improve autonomy (e.g., character customization, dialogue choice, skill improvement choice) and competency (e.g., achievements, adjusting difficulty levels) (Peng et al., 2012). Results showed that both autonomy- and competency-supportive game features led to need satisfaction, which in turn led to improved motivation, enjoyment, recommendation of the game to others, and the overall rating of the
game. Comparing exergames with stories to exergames without stories would be an interesting avenue of future research, particularly over the course of several months.

Research has also shown that a third component from self-determination theory – relatedness – is important. Surveys with adults and undergraduates suggest that feelings of autonomy and relatedness are the two most important reasons they exercise, play video games, and play exergames (Osorio, Moffat, & Sykes, 2012). Similarly, in one of the unsuccessful exergame interventions, interviews revealed that many of the children wanted someone else to play with (Baranowski et al., 2012). And there is some evidence to support this notion. In one intervention, children were given an exergame to play at home, but half of the participants also attended a weekly 1 hr session at the local fitness center where they played the exergame with other children (Chin A Paw, Jacobs, Vaessen, Titze, & Van Mechelen, 2008). Results showed that dropout was much lower in the multiplayer group (15%) versus the at-home group (64%), with participants in the multiplayer group playing the exergames for 525 min longer over the course of the 12 week intervention. Additionally, by the end of a 10 week study, participants who played an exergame competitively lost more weight than those playing the game cooperatively, suggesting more energy was expended by competitive play (Staiano et al., 2012).

Taken together, this research suggests that exergame developers and researchers should consider including multiplayer options.

Including multiple players in exergames opens up additional interesting avenues of research because not only are there two (or more) people playing the game simultaneously, there are also opportunities for interactions between the players. Players can influence each other to play harder and/or for longer, which translates to exergaming at higher intensities and for greater durations. While many exergames have incorporated social-psychological principles for behavior
change (see Lyons & Hatkevich, 2013, for a review of 18 exergames available to consumers), few (if any) have taken the principles of group dynamics into consideration. Doing so may be another way to increase the appeal and effectiveness of exergames.

**Group Dynamics and the Köhler Effect**

Groups are “a collection of two or more individuals who possess a common identity, have common goals and objectives, share a common fate, exhibit structured patterns of interaction and modes of communication, hold common perceptions about group structure, are personally instrumentally interdependent, reciprocate interpersonal attraction, and consider themselves to be a group” (Carron & Hausenblas, 1998, pp. 13-14) and are a large part of every person’s life (Steiner, 1972). Sometimes people are forced into groups (e.g., collaborative projects), but one’s membership in many groups is voluntary (e.g., clubs, exercise classes). The focus of a group depends on its goals. Some groups are mostly focused on a particular task (e.g., completing a sales pitch by the end of the quarter), while others are primarily focused on being social (e.g., weekly community dinners). Much of the group dynamics literature has focused on work groups, or groups that have some task-based objective. How effective a group is depends on what that task is and how it is structured (i.e., what needs to be done and how to do it), so it is important to know how to define tasks when studying groups. Because there are many types of tasks, Steiner (1972) created a typology to organize how work groups are studied and classified.

One way Steiner (1972) organizes tasks is by how much a task allows group members to work together. *Divisible tasks* are those that allow for the labor to be divided (e.g., on an assembly line, each person performs one of three separate jobs that contribute to the overall product). More importantly for this dissertation, however, are *unitary tasks*, or tasks that cannot be divided into smaller tasks where different group members can work on each subtask. In
unitary tasks, the person performing the job performs all parts of it (e.g., each person performs all three jobs on the assembly line, or in a soccer shootout, each athlete performs the same penalty kick).

Unitary tasks can be divided into additive, disjunctive, and conjunctive tasks (Steiner, 1972). In *additive tasks*, the group’s product is determined by the contributions of all members, as the group product is usually the sum or average of all individual contributions. An example of an additive task is a game like horseshoes, where the team’s score is the sum of individual scores after each set of throws. In a *disjunctive task*, the group’s overall performance is determined only by one member (and this tends to be the most capable or best member). For example, in a group trivia contest where the group can give only one answer, the person who knows the most about the given subject will most likely answer, while the other people in the group do not need to make any contributions to answering the question correctly. Finally, *conjunctive tasks* are tasks where the group’s product is determined by the worst member. Everyone is doing the same thing, but whether or not the group achieves its goal depends on the least capable member having enough ability or resources to achieve that goal (Steiner, 1972). For example, in partnered rock climbing, each climber is going at his/her own pace, but the group itself can go no quicker than the slowest member. (Although not included in Steiner’s (1972) classification scheme because it is not a group task, *coactive tasks*, where two people work at the same task, but independently of each other (i.e., they are not a team/group as there is no interdependence) is also an important task type to note here).

In addition to determining how the goal is to be achieved, task structures can also influence how motivated individual group members are to exert effort toward reaching that goal. How motivated a person is to contribute is commonly viewed as a function of two different
factors: valence and expectancy (Vroom, 1964). Valence is one’s affect towards a particular outcome. When the group’s goal is highly meaningful to a person (e.g., receiving an “A” on an assignment), that person will put forth a lot of effort to make that outcome more likely. Included but distinct from valence is instrumentality, which is how crucial one’s actions are for acquiring certain outcomes. If one believes that by performing well, the group will perform well, effort will be increased. For example, in a string quartet, the cellist must play well in order for the whole group to make beautiful music, so the cellist should put forth energy toward ensuring that happens. Expectancy is a belief about how likely the outcome and any associated rewards will result from one’s actions. Individuals who believe that they have a high amount of control over an outcome will put forth more effort because they expect their actions to make significant contributions toward the group’s product, while the converse is also true. For example, if a play is being run to the left side of the football field, the team’s right guard might not put forth much effort because the ball is heading toward the opposite direction and, therefore, the right guard has little say in whether the play will be successful. But if the play is being run to the right side, the right guard will exert more effort trying to block for the running back. Additionally, if the right guard expects his performance to lead to some reward (e.g., the game ball, a touchdown), effort should be higher, and if the guard expects his performance will not increase the likelihood of receiving a reward (e.g., because a wide receiver is having an excellent game), then effort may be unaffected, or even lowered. How motivated a person is to act in a given group situation is a product of these two functions, Expectancy x Valence (Vroom, 1964).

The combination of Steiner’s (1972) task typology and Vroom’s (1964) expectancy-value theory allows researchers to explore the relationship between task structure and motivation. From this foundation, numerous motivational processes have been studied. Depending on the
type of task (and other related variables, such as group size), an individual’s motivation to contribute to the group can either decrease or increase. These are termed motivation losses and motivation gains, respectively, and have been a large focus of group dynamics research.

**Motivation Losses**

Perhaps the most studied type of motivation loss in group settings is *social loafing*, which is the tendency for individual group members to expend less effort on a task when performing it in a group setting compared to when that same task was performed individually (Latane, Williams, & Harkins, 1979). Furthermore, the magnitude of certain particular motivation losses increases as group size increases. One such motivation loss is known as the Ringelmann effect and is more likely to occur in additive tasks (Latane et al., 1979). Social loafing seems to occur for several reasons, including when individual contributions to the group’s outcome are not identifiable or when the task is not perceived as meaningful (Kerr & Bruun, 1981; Williams & Karau, 1991). Indeed, a number of variables moderate social loafing, most of which have to do with expectancy, instrumentality, or valence (Karau & Williams, 1993; Shepperd, 1993).

Although social loafing is a robust and generalizable finding, applicable to many situations, it is not an inevitable or unavoidable effect. In fact, at times it can be a hard effect to produce depending on the environment surrounding the group’s performance. For example, a series of experiments could not elicit a social loafing effect until task motivation was manipulated to be lowered beforehand (Price, 1993). The author speculated that the participants were highly motivated to perform because high performance norms were standard for their major. This indicates that if a group’s productivity norm is high, individual effort will remain high because group members view doing well on the task as important.
Other research suggests that identifiability and one’s perception of their indispensability to the group play important roles in mitigating social loafing. Increasing the size of a group can make members feel anonymous, which means it is easier for them to “hide in the crowd” and not put forth as much effort as they would if they had to perform the task individually (Kerr & Bruun, 1981). Increasing identifiability, however, significantly attenuates the social loafing effect, suggesting that when one’s contributions to the group are visible, effort will improve. The same pattern of findings emerges for dispensability. In disjunctive tasks, the most capable group members performed the best and reported low feelings of dispensability (meaning they believed their contributions were very important for the group’s outcome), while they performed the worst and reported high feelings of dispensability in a conjunctive setting (where performance depends on the least able member) (Kerr & Bruun, 1983). The opposite pattern was observed for low ability group members, where they felt their performance was indispensable for the group’s success in a conjunctive task but dispensable in a disjunctive task.

In short, although social loafing is a common occurrence in groups, there are ways to prevent these motivation losses from happening. In accordance with expectancy-value theories (Vroom, 1964), increasing the value of individual and group outcomes should keep effort high, as should ensuring individual contributions are noticed and crucial for group success (Shepperd, 1993). But more than simply preventing a decline in motivation, certain mechanisms can also increase overall motivation, leading to enhanced performances and greater contributions of effort.

**Motivation Gains**

When working with other people, one can be either equal in ability to one’s teammates, superior in ability to one’s teammates, or inferior in ability to one’s teammates. Motivation and
performance gains can be elicited in each of these conditions. When ability is equal, competition seems to be one way to enhance performance. When group members are of unequal ability, however, one of two major motivation gains effects can occur (Witte, 1989).

The first effect is *social compensation*, which occurs when a group member works harder than usual to compensate for the expected poor performances of the other members (Williams & Karau, 1991). This is most likely to occur when the task is viewed as meaningful and there are questions about one’s teammates’ willingness, ability, or reliability to make significant contributions. For example, a very talented volleyball player on a team without other talented players may attempt to compensate for the rest of the team’s expected poor performances by trying to cover more of the court than typical. In this situation, the talented player is exerting more effort and energy to compensate for other players’ poor performance, which would hopefully help the team perform well. However, the sucker effect (a motivation loss) can occur when high-ability group members believe the rest of the group is deliberately slacking, particularly when the task is not seen as important or meaningful (Kerr & Bruun, 1983). In this case, they will reduce their efforts to avoid being taken advantage of – that is, they avoid being a ‘sucker’ by not covering for other capable group members.

Social compensation occurs in high-ability group members. However, when one is the weakest link in a group, certain conditions can trigger improvements in performance. This is known as the *Köhler effect* and it occurs when inferior group members exhibit motivation and performance gains when working with a superior other compared to working individually. The Köhler effect seems to be more useful in exercise settings versus the social compensation effect, which is why the rest of this section will focus on the Köhler effect. Active exercisers presumably do not need as much motivation to exercise compared to someone who is sedentary
or less active. Thus, taking advantage of the Köhler effect may help individuals who are unfit or unmotivated engage with exercise for longer.

**The Köhler effect.** In the 1920s, Otto Köhler, a German industrial psychologist, conducted two studies that showed the performance of weaker group members can increase in certain situations (Köhler, 1926, 1927, as cited in Stroebe et al., 1996). In his first study, men on a rowing team completed a weight-lifting task until exhaustion, either individually or in pairs, over the span of several weeks. The results showed that if one partner was moderately (~25%) better than the other, it led to a 35% performance gain. (This effect also occurred for triads, but the performance gain was smaller, only 10%). A similar pattern of results emerged in his second study, where participants (either alone or in pairs) were instructed to turn a hand wheel in rhythm with a metronome. However, if one partner was much better than the other partner, or if there were no major differences in ability, performance was reduced (Köhler, 1926, 1927, as cited in Stroebe et al., 1996).

There are three important features of Köhler’s research to consider. First, the tasks used in Köhler’s research are conjunctive tasks. Pulling weights or turning a hand wheel in rhythm to exhaustion depends on the least capable member. Once the least capable member is unable to continue, the group must stop, making the group’s performance dependent on the weaker partner. Second, these tasks are both persistence tasks, meaning the goal is to perform the action for as long as possible. (Steiner, 1972, refers to these type of tasks as *maximizing tasks*, where the goal is to maximize performance in some way, such as by persisting for as long as possible or by performing the most repetitions). Third, these studies showed that differences in partner ability matter. When the discrepancy between partners is large or if there is no discrepancy,
performance gains are reduced. But when the discrepancy between partners is moderate, the performance of the group seems to improve due to improvements by the weaker member.

Researchers (e.g., Hertel, Kerr, Scheffler, Geister, & Messé, 2000) have split these findings into two effects. First, the Köhler motivation gain effect refers to the increase in motivation and performance of inferior group members in conjunctive tasks. Second, the Köhler discrepancy effect is a moderation effect, showing that motivation gains are highest when the differences in ability between partners are only moderate. For the sake of clarity, in the rest of this dissertation, the Köhler effect will refer to the motivation gain effect that weaker group members display, while the Köhler discrepancy effect will refer to the moderating effect of relative partner ability.

Although Köhler conducted his research in the 1920s, it was seemingly forgotten until it was “rediscovered” over 60 years later (Witte, 1989). This rediscovery eventually led to a series of research lines intent on exploring the mechanisms that underlie the Köhler effect and the factors that moderate its strength. Stroebe and colleagues (1996) conducted a series of five experiments with the goal of replicating and extending Köhler’s original research. Their first experiment replicated the weight lifting task and found evidence for a discrepancy effect: when moderate ability differences existed between partners, group performance was higher when compared to partners of similar ability. These gains in group performance were mostly due to gains by the weaker member. The next three experiments attempted to replicate the hand wheel turning task, however, the results were more mixed. While Köhler’s hand wheel turning task was paced, and therefore a persistence task, Stroebe and colleagues’ (1996) hand wheel turning task was not paced. Participants were instructed to turn the wheel as many times as possible in a
given time frame. The authors speculate that these task demands either fostered competition, left little room for improvement after the individual trial, or both, which eliminated the Köhler effect.

The fifth experiment used a new type of persistence task, one that would go on to be used in several additional Köhler effect studies. Participants were instructed to hold their arm straight out in front of them for as long as possible while holding a weight in their hand (Ruess, 1992, as cited by Stroebe et al., 1996). Results showed a Köhler effect, albeit with smaller performance gains than usual. This series of studies showed that the Köhler effect is not a by-product of the 1920s, but is a replicable and interesting avenue for improving motivation in weaker group members.

A number of explanations for why the Köhler effect occurs have been put forth. Köhler (1926, 1927) himself suggested that competition for leadership could explain the findings, while Witte (1989) offered a similar explanation, status incongruency. Both of these explanations state that because leadership (or status) is not clear in conditions when partners are at equal ability, performance will be highest. However, given the highest performances were seen when there were moderate discrepancies, these explanations do not hold up to scrutiny. Köhler (1926, 1927) also suggested interpersonal resonance, where these conditions created a way for the stronger member to influence the weaker member. But the process by which this occurred was not specified. Stroebe and colleagues (1996) explained the Köhler effect in terms of goal comparison. The weaker member sets the stronger member’s performance as the standard to match, thus enhancing the weaker member’s effort. But if the stronger member’s performance is too great, the weaker member will think reaching the standard is impossible, and thus will not put forth any extra effort. There is some evidence for this view, though goal setting alone is not a sufficient explanation (Hertel et al., 2000).
More recent research, however, has zeroed in on two mechanisms through which the Köhler effect occurs: social comparison and indispensability. People have a tendency to evaluate their own ability, but when objective evaluative methods are not available, people will compare their ability to the ability of people around them, thus engaging in social comparison (Festinger, 1954). When the performance of one’s peers are moderately better compared to one’s own performance, the bar is raised and performance is increased to match those more capable others (Seta, 1982).

However, social comparisons alone are not enough to explain the Köhler effect (Kerr et al., 2007). Social comparison is enough to improve the performance of a weaker member, but that cannot explain why motivation gains are greater under conjunctive task conditions compared to additive or coactive tasks. What explains this added boost is indispensability, or how crucial one’s contributions to a group are for group success. Indeed, research has shown that because conjunctive tasks are dependent on the weaker member (thus improving indispensability to the group), the weaker member performs better when working at a conjunctive task than when social comparison is the only mechanism available (as in a coactive task setting) (Hertel, Niemeyer, & Clauss, 2008; Gockel, Kerr, Seok, & Harris, 2008).

Therefore, it is this combination of upwards social comparison and indispensability to the group that underlie the Köhler effect. A recent review article and meta-analysis shows this to be the case, as the largest effect sizes of motivation gains in inferior group members occurred in conjunctive settings (which include social comparison and indispensability) compared with additive and coactive settings (which only include social comparison) (Kerr & Hertel, 2011; Weber & Hertel, 2007). Indeed, the Köhler effect is a rather strong effect, as the overall effect size for motivation gains in conjunctive settings is large ($g = 0.72$). Furthermore, the Köhler
effect can persist over time, which Köhler’s original research (1926, 1927) and more recent research both show (Lount, Kerr, Messé, Seok, & Park, 2008a). In addition to exploring how and why the Köhler effect occurs, other research has focused on factors that may strengthen or weaken the size of the Köhler effect.

**Moderators of the Köhler effect.** There are a number of factors that can moderate the Köhler effect, including partner ability, gender, competition, social ties with the group, how feedback is given, and presence. Perhaps the most crucial moderator is the amount of discrepancy between the two partners (the Köhler discrepancy effect). As Köhler (1926, 1927) showed in his research, moderate differences in ability (~25%) led to the largest motivation gains (and this finding was generally replicated by Stroebe et al., 1996). Subsequent research (using the physical persistence “hold a weight above a tripwire” paradigm) initially failed to produce the Köhler discrepancy effect (Hertel, Kerr, & Messé, 2000). However, this study did not allow participants to receive feedback about how well they were doing compared to their partner. When participants had knowledge of their partner’s ability, the Köhler discrepancy effect was observed (Messé, Hertel, Kerr, Lount, & Park, 2002). This research concluded that while working with a better partner generally means the weaker member will improve performance, the performance gain is largest when the better partner is only moderately (~40%) better.

Receiving feedback about one’s partner’s ability is important for the Köhler effect. One study showed that receiving continuous feedback about how one’s partner is doing or only at the end of each trial about their relative performances are both sufficient for the Köhler effect to occur (Kerr, Park, & Sambolec, 2005). Other research suggests that effort is highest when the partner’s performance is continuously available to view (Hertel et al., 2008; Weber & Hertel,
Thus, allowing people to see how they are doing in relation to their partner while performing strengthens the Köhler effect.

Another moderator of the Köhler effect is gender. Women tend to have a large motivation gain effect in conjunctive conditions for both physical persistence and cognitive tasks, while men show similar motivation gains under conjunctive and coactive conditions (though motivation gains are still highest in conjunctive conditions) (Kerr et al., 2007; Weber & Hertel, 2007; Weber, Wittchen, & Hertel, 2008). This suggests that for men, an upwards social comparison is enough to trigger motivation gains, while for women, the indispensability component is very important.

Why only being paired with a superior partner is enough to improve motivation in men can be explained in terms of stereotypical sex roles. Specifically, men are stereotypically more competitive than women, who themselves are stereotyped to be more cooperative. As such, men are more sensitive to being told they are the weaker member of a pair and will try harder to raise their performance so as to not be the weak link. When men are primed for competition, their performance under coactive task conditions improves to match their performance under conjunctive task conditions (which is not improved by competition) (Sambolec, Kerr, & Messé, 2007). Similarly, when women are primed for competition, they also exhibit significant motivation gains under coactive conditions (Kerr et al., 2007). Together, these results fit with traditional stereotypes about how men and women behave. Having a favorable social comparison with others is more of an individual, competitive goal, so it follows that the presence of a discrepancy in ability is enough to trigger performance gains in men. On the other hand, indispensability, which is more of a group goal (contributing to the group’s success), is more important for women, who are stereotyped as being more cooperative.
However, not all studies have found that gender moderates the factors that contribute to the Köhler effect. Research using abdominal planks (discussed in more depth later) has not always shown differences between coaction and conjunctive tasks (Feltz et al., 2011). The authors suggest that interpersonal competitiveness may have been primed during the experiment, and in such contexts, women who more typically exhibit an indispensability effect may not do so.

Much of the Köhler effect research uses artificial, temporary groups. Participants come into the lab, are grouped with another person they do not know, asked to perform together, and then the group breaks up once the experiment is over. However, in real working situations, oftentimes group members know each other already, if only a little. Some research has explored how social ties with group members moderate the Köhler effect. Research suggests that simply being acquainted with one’s partner beforehand does not change much: the Köhler effect still occurs (Weber et al., 2008). However, when paired with a friend, the Köhler effect is strengthened, leading to larger gains in motivation and performance (Kerr & Seok, 2011). And the opposite is also true: when one is ostracized from a group, the Köhler effect is weakened (though not completely eliminated) (Kerr, Seok, Poulsen, Harris, & Messé, 2008). Both friendship and ostracism moderate the Köhler effect by strengthening or weakening the importance placed on group success when one is indispensable, respectively. When working with a friend, one does not want to let their friend down. But when working with someone who already excluded their future partner, the excluded member values the group outcome less, so the importance placed on their contributions also declines.

With the rise of computers and particularly the internet, groups can (and do) work on the same project together, even when they are hundreds of miles apart. This brings up an interesting
question from a group dynamics perspective: how is motivation affected when people are remotely working together? When conjunctive tasks are used, the usual Köhler effect by the weaker member is observed. This was seen when participants worked with anonymous group members on a computer-based cognitive task (Hertel, Deter, & Konradt, 2003). Similarly, when working with a virtually-present partner, the Köhler effect still occurred (Lount, Park, Kerr, Messé, & Seok, 2008b). However, presence does matter: motivation gains are strongest when working with physically-present others due to a greater concern of being evaluated negatively (Weber & Hertel, 2007).

Overall, the Köhler effect is a rather robust effect in that it occurs in a number of situations and is difficult to eliminate (though note Irwin, Feltz, & Kerr, 2013, and Seok, 2007). Working with a moderately more capable partner seems to be crucial for maximizing motivation and performance gains, as is receiving constant feedback about how one’s partner is doing. Competition and friendship seem to strengthen the Köhler effect, while ostracism weakens it. The effect persists over several trials, and is present even when working with people who are only virtually present. As such, the Köhler effect is a promising group dynamics principle that may be a key part of the solution to increasing the intensity and duration at which people play exergames. Research has already shown that indispensability and outcome valence are crucial components behind the improvement of times in relays (compared to individual times) for both swimmers and track and field athletes (Hüffmeier & Hertel, 2011; Hüffmeier, Krumm, Kanthak, & Hertel, 2012; Osborn, Irwin, Skogsberg, & Feltz, 2012). If these types of motivation gains are observable in athletic competitions, it is plausible that this effect will also extend to exercise and exergame play.
**Exergames and the Köhler effect.** Initial research combining the Köhler effect with exergames has yielded promising results (Feltz et al., 2011). The basic task paradigm is similar to previous Köhler effect research, except the task is exercise-based. Participants arrive at the lab and are instructed to play along with an exergame, the EyeToy: Kinetic™ (Sony Computer Entertainment America LLC) for the Sony PlayStation®2. The exergame demonstrates a series of five plank exercises, but only for a pre-determined amount of time (around 10 s). To make this task a persistence task, the exergame is paused before it advances to the next plank, and participants are instructed to hold the plank for as long as they can.

Following this first block of exercises, which are performed individually, participants are told they will be performing the exercises again, but (in all except the control condition) with a partner. An introduction with the partner takes place via a simulated webcam-like connection where basic information is exchanged. In actuality, this virtual partner is a pre-recorded video of a confederate, though participants are not told this; they believe they were speaking with another person live.

After the introduction, the Köhler effect manipulation occurs. Participants are told they will be performing as a team with their partner, each starting to hold the plank exercises at the same time and holding them for as long as possible. The group’s score is the score of the person who stops holding the exercise first, so the group’s score depends on the weaker member (making this a conjunctive task). Participants are then truthfully told the average length of time they held the planks and falsely told the average length of time their partner held the planks. The partner’s time is always manipulated to be roughly 40% better than the participant’s time. So for example, if the participant held the plank exercises for an average of 100 s, the partner is purported to have held the planks for an average of 140 s. In this way, both mechanisms driving
the Köhler effect are present (i.e., the partner having greater ability creates an upwards social comparison, while the conjunctive task with a more capable partner enhances indispensability.

Following the partner introduction and manipulation, participants are given a short break. The second block of exercises is the same as the first block of exercises with one exception: the partner’s performance is projected onto a screen. Like with the introduction, these are a series of looped pre-recorded videos, which means the partners never tire and quit, hence, participants can never match their partner’s performance. In this way, participants are able to see how their partner is doing while performing the exercises (continuous feedback). Once the second block is over, participants are thanked and debriefed.

Using these basic task procedures, an ongoing program of research has demonstrated that motivation gains can occur in weaker members on exercise tasks. The first study to come out of this research showed that motivation gains occurred in conjunctive, coactive, and additive task conditions, with no significant differences among the three (Feltz et al., 2011). These results seem puzzling because they do not align with predictions based on the Köhler effect. The authors speculated that the task may have been perceived as competitive, thus feelings of competition in the additive and coactive conditions matched feelings of indispensability in the conjunctive conditions. Regardless, participants with a partner held the plank exercises an average of 54 s longer (a 24% increase) compared to individual controls. In addition, when performance incentives were removed in a subsequent study, a stronger conjunctive condition effect was found (Kerr, Feltz, & Irwin, 2012).

Additional research has resulted in findings that are more in-line with the Köhler effect. One study compared cycling on a stationary bike that had a built-in exergame under conjunctive conditions to cycling on the same bike and exergame under coactive conditions (Irwin et al.,
Participants completed a total of six sessions, with the first session serving as a baseline performance measure. Results showed that starting with the third session, participants in the conjunctive condition performed significantly better than participants in the coactive condition, and this difference grew in magnitude through the final session. On the final trial, participants in the conjunctive condition pedaled a remarkable 19 min (~177%) longer than individual controls, and 11 min longer than participants in the coactive condition (Irwin et al., 2012). These results clearly fit with most Köhler effect research and show that these performance gains may even grow over time (though note Kerr, Forlenza, Irwin, & Feltz, 2013, which did not show a growth in gains over time).

Other studies have focused on variables that may moderate the observed motivation gains. The Köhler discrepancy effect was replicated in these settings, as a partner discrepancy of 40% led to significantly stronger motivation and performance gains compared to partner discrepancies of 1% and 100% (Feltz et al., 2012). Working with a partner who was older and/or heavier did not moderate this effect among college students (though note that there was a trend of men working harder with heavier partners), suggesting a wide range of partners can be used to produce motivation gains in exercise settings for college students (Forlenza et al., 2012). Certain forms of partner verbal encouragement moderated the Köhler effect, but by weakening it, leading to smaller gains, which suggests that making motivational comments may backfire (Irwin et al., 2013, though note Max, 2014). And, as noted previously, incentives may undermine the Köhler effect in exercise settings, as they were also shown to weaken motivation gains (Kerr et al., 2012).

One other study explored how patterns of inferiority and superiority on two different tasks affected performance (Kerr et al., 2013). In this study, participants performed one plank
exercise and one wall-sit exercise a total of four times each, with the last three following the partner introduction. However, the participant was either inferior to the partner on both tasks or superior to the partner on one of the tasks. Results showed the usual Köhler effect, suggesting participants did not receive any additional boost in effort as a result of sometimes being the superior partner. Additionally, the Köhler effect weakened across trials, but was still significant on the final trial.

Ancillary analyses of this body of research also yield promising results (Feltz et al., 2011; Feltz et al., 2012; Forlenza et al., 2012; Irwin et al., 2012; Irwin et al., 2013; Kerr et al., 2012; Kerr et al., 2013). Despite the fact that the participants working with a partner persisted at the exercises for anywhere between 24% and 177% longer than participants working individually, they only sometimes perceived themselves as working harder (Forlenza et al., 2012; Irwin et al., 2013; Kerr et al., 2012). Furthermore, these better performances did not lead to a drop in self-efficacy beliefs, enjoyment, or intention to exercise the next day (with the exception of Kerr et al., 2013, where participants enjoyed the task more if they were sometimes superior). This suggests that when conjunctive task demands and social comparison were present, participants exercised for a longer duration without experiencing any adverse consequences, even when they perceived themselves as exerting more effort.

Overall, then, it seems that incorporating the Köhler effect into exergames is a worthwhile approach as it has led to favorable results thus far. However, there are other avenues for exploring how the Köhler effect can be used to increase exercise behavior. In particular, two areas for exploration that can expand the situations in which the Köhler effect is present, both of which center on the source of the partner, include the use of software-generated partners and the use of imagined partners.
Human-computer Interactions

The research thus far has only used pre-recorded confederates, which means participants believed they were working with another person. It would be beneficial if the motivation gains seen in this research could be obtained without the need for another human being, i.e., with a fake virtual partner instead. While exercising alongside another human would most likely be optimal, this is not always possible, particularly for people who have conflicting schedules or who cannot find someone of the optimal ability discrepancy. But if the exergame provided the exercise, the task conditions that can improve exercise intensity, and the partner, all of that should lead to greater motivation to play the exergame. However, no research to date has tested whether people will exercise harder with a fake, software-generated partner. Fortunately, a body of research that has explored how humans interact with computers does exist, and this research may be able to shed light on how well such interactions would work.

Media Equation

As computers have become used for more and more reasons, humans are forced to interact with computers on a regular basis. In addition to using personal computers and laptops, people are talking to their phones, not only to people on the other end. These types of interactions have led researchers to examine how people are engaging with their technology, not from a physical (e.g., touching, talking) standpoint, but from a social one. Overall, this research – known as the Media Equation theory – has shown that people treat computers as social actors, applying social rules from human-human interactions to human-computer interactions (Nass, Steuer, & Tauber, 1994; Reeves & Nass, 1996).

Of course, people know what computers are, and they know computers are not human. Yet, people unconsciously engage with computers socially, inappropriately invoking human-
human interaction norms with computers (Nass & Moon, 2000). For example, people treat the same voice to be the same actor regardless of which machine the voice is coming from, even though it is ridiculous to think of a voice coming from a computer as an actor (Nass et al., 1994). Three explanations have been put forth for why this happens. People may be anthropomorphizing computers by viewing them as more human-like than they really are. People may also believe they are interacting not with the computer itself, but with the person who programmed the computer interfaces. Neither of these arguments seemed to hold up, however, as participants consistently denied treating computers as humans and treated the computer as the source of the interaction, not the programmer (Nass & Moon, 2000; Sundar & Nass, 2000).

The third explanation deals with mindlessness. Mindless behavior can be triggered by cues from one’s environment that activate simplistic scripts for behavior. Because humans are wired for social interaction, they are generally ready to activate these social scripts, and computers increasingly have functions that can trigger these scripts (e.g., words for output, interactivity, performing human jobs). When these social scripts are activated, people mindlessly respond to computers as if they were responding to another person, thus inappropriately applying a human-human interaction script to a computer (Johnson & Gardner, 2007a; Nass & Moon, 2000). Thus, mindlessness can explain why people are polite to computers (Nass, Moon, & Carney, 1999), experience stronger feelings of presence when a computer’s personality matches their own (Lee & Nass, 2003), and engage in reciprocal behaviors with computers (Fogg & Nass, 1997).

If people will apply various social scripts to interactions, it is plausible that they could also view computers as teammates, and some research supports this idea. Participants who were interdependent with a computer on a task reported higher feelings of team affiliation compared to
non-interdependent participants (Nass et al., 1996). These participants reported themselves as being more similar to their teammate (the computer), perceiving the computer to be friendlier and providing higher quality information, and more open to influence from their teammate. Similarly, participants cooperated with a human agent on a computer as much as they did with an actual human, keeping more of their promises in a social dilemma game (Parise, Kiesler, Sproull, & Waters, 1999). Thus, people seem willing in at least some senses to “team up” with computers.

However, some research suggests caution should be made when people and computers are partnered together. One study showed that while experienced computer users had feelings of team affiliation when working with a computer on a task, they ultimately responded negatively to that computer, rating its information as lower quality and more different from their own answers, and being less influenced by the computer’s choices (Johnson & Gardner, 2007b). However, this is consistent with the Media Equation theory and mindlessness explanations, as these experienced users did perceive the computer to be a part of the team; they just did not perceive the computer to contribute to the group productively. Indeed, less experienced users did not show any Media Equation effect, reporting no feelings of team affiliation (Johnson & Gardner, 2007b). Other research shows that people are flat-out less cooperative with computers, choosing to engage in more non-cooperative behaviors when playing a social dilemma game with a computer compared to playing the game with another person (Kiesler et al., 1996; Miwa & Terai, 2012). In these experiments, participants rated the computer partner as less likeable and desirable, which in turn affected how much they wanted to cooperate with the computer.

Research on the Media Equation theory generally supports the idea that people tend to interact with computers in a manner similar to how they interact with humans from a social rules
perspective. However, the research is less clear when it comes to team formation. Some studies show that people will partner and cooperate with computers and computer agents, while others will not. One interesting implication of this human-computer team research seems to be that the more human-like the computer partner is (i.e., the more human elements the computer partner contains), the more likely people are to like and cooperate with the partner (Kiesler et al., 1996; Parise et al., 1999). Another possible explanation for why people do not always form cooperative teams with computer partners has to do with these partners being too human-like.

Uncanny Valley

The Uncanny Valley was first applied to robotics and states that as robots become more human-like, our sense of familiarity (or affinity or likeability, depending on the translation, see Tondu & Bardou, 2011) with them increases (Mori, 1970/2005). However, at some point, robots become too human-like, and this has the effect of reducing our feelings of familiarity, making us feel strange. This drop in familiarity is the Uncanny Valley (see Figure I). The other side of the valley is a healthy human being, with which we experience the greatest feelings of familiarity. Movement exaggerates the peaks and valleys; a shambling zombie should provoke more feelings of discomfort compared to a still corpse. Similarly, a toy robot that can move on its own should generate a greater feeling of familiarity compared to one that is unable to move by itself (Mori, 1970/2005). While the Uncanny Valley was initially conceptualized for robots, recent research has extended it to include computer-generated characters (Flach, de Moura, Musse, Dill, Pinho, & Lykawka, 2012).
As an example, imagine a robot that has square, metal hands and fingers. We would not experience much familiarity with that robot’s hands because they are clearly robot-like. Now imagine that the robot’s hands were shaped just like a human’s, with a wrist, palm, and five fingers, but all the wires were visible and there was no protective layering. Our sense of familiarity would be heightened because it looks more human-like, but it is still clearly robotic. Now imagine the engineers covered that robot’s hands with a thin protective coating and painted them to look like human skin. These prosthetic hands might appear to be human at a glance, but when it is noticed that the hands do not look or move quite like human hands, our sense of familiarity decreases and our sense of uneasiness increases. These hands have fallen into the Uncanny Valley: they are human-like, but clearly not human. Finally, imagine the hands are now fully-functioning human hands. Our sense of familiarity with them would rise dramatically (and our sense of discomfort would fall dramatically) because we are very familiar with human hands.
The previous examples illustrate the range of how we react to human-like robots. One explanation put forward for why we experience discomfort is cognitive dissonance, which is a tension that people experience when they hold two conflicting beliefs in mind (Tondu & Bardou, 2011). Viewing a robot or computer agent that looks and acts as a human would, but is artificial, leads us to feel tension because we are viewing something that we want to interpret as human, but we know it is not. When people are engaging in social interactions, they typically do not want to feel tension, but human-like robots can cause tension to develop because they are simultaneously human-like and not human (Tondu & Bardou, 2011). An important note is that the amount of cognitive dissonance experienced depends on the person; reactions are individualized, and people do not have to react with discomfort (Hanson et al., 2005). Therefore, when engaging socially with machines, experiencing these computer agents as both human-like and not human can lead to feelings of tension, and these feelings may ultimately explain why some people do not cooperate with computers.

Because of the possibility to experience discomfort and tension with robots or computer agents that are too human-like, Mori (1970/2005) recommended that designers should avoid aiming for the second peak, or avoid trying to make machines look just like humans. Instead, designers should aim for the first peak, right before falling into the Uncanny Valley. However, those recommendations were made over 40 years ago and technology has advanced considerably since that time. Indeed, other roboticists believe it is possible to reach the second peak and should be the goal (Hanson et al., 2005). Still, the Uncanny Valley is something that needs to be considered when designing software-generated partners. If the goal of creating these partners is to motivate people to exercise harder, the partners cannot be designed in such a way that they cause tension or discomfort simply by virtue of their appearance and behavior. That could hurt
the game play experience and make people less likely to continue playing. As such, the software-generated partners should either be so realistic that they are virtually indistinguishable from people, or clearly non-human. Either way, exploring whether or not people will find software-generated partners as motivating as human partners has interesting and important implications for exergame design.

**Cognitive Strategies During Exercise**

Another type of non-human partner that could be used to improve exergame play is an imagined partner. Before the feasibility of an imagined partner can be assessed, however, it is important to have a sense of what people are thinking about while exercising. Overall, people purposely engage in different types of thought processes during exercise and these thought processes can ultimately affect how they feel and perform.

When people exercise, they engage in two broad strategies to help them avoid pain and pace how hard they are working. When people use *associative strategies*, they are focusing on their own body and using that information to adjust their effort accordingly. For example, a runner may pay attention to how her legs and feet are feeling, and if there is some pain, she can run a little slower so that the pain goes away, thus preventing her from becoming injured and allowing her to continue running. By contrast, when people use *dissociative strategies*, they are directing their attention away from their own body to avoid feeling pain. For example, a runner might think about what he is going to make for dinner after his run is over, thus keeping any pain or discomfort out of his mind (Masters & Ogles, 1998; Morgan & Pollock, 1977; Stevinson & Biddle, 1998). It should be noted that these strategies are not necessarily mutually exclusive; exercisers can switch between them as needed, though they cannot be performed at the same time.
Generally, the use of each type of strategy is associated with different outcomes. Associative strategies result in faster performances (as in swimming or running) and are preferred in competition, while dissociative strategies are more useful for improving performance on endurance activities and lowering perceived exertion (LaCaille, Masters, & Heath, 2004; Masters & Ogles, 1998; Salmon, Hanneman, & Harwood, 2010; Stanley, Pargman, & Tenenbaum, 2007). However, much of this research has not taken exercise intensity into account, which limits the definitiveness of these findings (Lind, Welch, & Ekkekakis, 2009). Exercise intensity is important to consider, as people overwhelmingly report using associative strategies as intensities increase (Hutchinson & Tenenbaum, 2007). It seems that attention becomes dominated by physiological perceptions at higher intensities, while at lower intensities people are free to dissociate or associate as they please.

**Imagery**

Although there has been a good amount of research on cognitive strategies in general, there is a notable lack of research on the use of imagery as a specific strategy (Razon et al., 2010). This is somewhat puzzling as some of the earliest literature on cognitive strategies mentioned that people imagine things while exercising (e.g., a runner who builds an entire house during a marathon, see Morgan & Pollock, 1977). Imagery has received a considerable amount of attention in the sport and exercise psychology literature, however, and is one of the most studied mental skills.

Imagery is a mental process whereby a person creates or recreates situations in one’s mind using one or more senses (Vealey & Greenleaf, 2009). For example, a goaltender could use imagery in an attempt to predict how the opposing team’s star skater will approach the penalty shot, or to recall a past penalty shot situation to think about how to avoid making the same
mistake again. Research with elite athletes has shown that an overwhelming majority used imagery systematically to help them perform at their highest level (Orlick & Partington, 1988). The overall body of literature supports this notion, with a variety of evidence existing that shows imagery positively influences performance and psychological states related to performance (such as confidence and motivation) in many different sports (Vealey & Greenleaf, 2009; Weinberg, 2008).

In sport settings, imagery has generally been shown to serve motivational and cognitive functions, such as planning a strategy, learning a new skill, managing arousal, picturing goals, and maintaining confidence (Hall, Mack, & Paivio, 1998; Paivio, 1985). Although imagery use by exercisers has been studied less than imagery use by athletes, research has shown that exercisers use imagery for a variety of reasons as well. These reasons include improving technique, developing routines, picturing one’s ideal appearance, improving competitive outcomes (such as a 5k time), improving fitness and health, building self-efficacy, and managing arousal (Giacobbi et al., 2003; Hausenblas et al., 1999).

Exercisers use imagery in a variety of settings and times, including while they are exercising (Giacobbi et al., 2003). Existing evidence supports the idea that using imagery can improve exercise behaviors, thus making it a potential intervention tool. Exercise imagery is positively correlated with exercise behavior, such that people who use more exercise imagery exercise more than those who use less exercise imagery (Gammage, Hall, & Rodgers, 2000; Hausenblas et al., 1999). Imagery use has also been linked with various types of efficacy to exercise and autonomous motivation, which in turn are linked with exercise behaviors (Cumming, 2008; Stanley, Cumming, Standage, & Duda, 2012). And imagery has been used successfully as an intervention tool, its use leading to greater motivation, action planning, and
ultimately more physical activity (Chan & Cameron, 2012). Thus, it seems promising that imagery can be used to improve the intensity and duration at which people exercise.

To date, it appears that only a few studies have directly assessed the effectiveness of imagery as a cognitive strategy during physical activity. The first study had participants compete with a confederate on a wall-sit exercise, with instructions to hold it for as long as possible (Feltz & Riessinger, 1990). While holding the exercise, participants used mastery-oriented imagery (imagery focused on eliciting feelings of competence) facilitated by listening to an audio recording that guided this type of imagery, which led to stronger self-efficacy and better performance. The second study asked participants to hold a handgrip dynamometer at 30% of their maximum grip for as long as possible. Overall, both associative imagery (imagery focused on the current task, such as feeling stronger or more aware of one’s own bodily sensations) and dissociative imagery (imagery focused on scenes unrelated to the current task, such as achieving a major life goal) led to better performances and lower perceptions of exertion compared to a control group, although these results were not statistically significant (Razon et al., 2010).

The third study asked participants to perform the handgrip dynamometer task and separately to cycle at 10% above their anaerobic threshold, which is vigorous exercise. Dissociative imagery led to reduced feelings of exertion and better performance on the handgrip task, while associative imagery led to improved performance on the cycling task (Razon, 2012). It was suggested that the imagery helped improve task-related feelings of motivation, in turn improving performance. Thus, the research on using these particular types of imagery while exercising to improve performance is somewhat conflicting, though it appears that imagery can be used to extend performance time on strength and exercise persistence tasks.

**Imagined Partners**
Overall, the research on cognitive strategies during exercise and exercise imagery suggests there is the potential for imagery to be used while exercising to improve exercise performance. In the context of the Köhler effect, this would entail that participants imagine they are exercising with a partner. Given that the Köhler effect is a robust effect and can be elicited regardless of whether one’s partner is virtually or physically present, or even anonymous, (Hertel et al., 2008; Lount et al., 2008b), and is able to extend exercise duration (Feltz et al., 2011), it seems plausible that participants can use imagery to imagine working with a partner.

In this way, imagery during exercise would be focused on creating and maintaining an image of a superior partner throughout the duration of the exercise. For example, a jogger could picture herself running with a fitter friend under conjunctive task conditions for motivation to keep her pace up during the last few minutes of a run. However, the literature on imagined interactions seems to be focused on improving attitudes towards members of outgroups to reduce prejudice (e.g., Crisp & Turner, 2009) and conversations with significant others (e.g., Honeycutt, Zagacki, & Edwards, 1990). This literature does suggest that imagining a social interaction can lead to similar effects as if one were to actually have that interaction in real life, which can lead to better attitudes towards others (Crisp, Stathi, Turner, & Husnu, 2009). Drawing from the literature on the Köhler effect, cognitive strategies during exercise, exercise imagery, and imagined interactions with others, using imagined partners to improve motivation and enhance performance is an interesting proposition that could ultimately provide a relatively simple way to boost exercise behaviors.

**Summary**

Physical inactivity is associated with major health problems, so it is important to explore ways to reverse these trends. While video games, as a sedentary activity, seem to contribute to
this problem, a new generation of active video games has the potential for contributing to the solution. Building group dynamics principles into exergames, specifically the Köhler effect, has been shown to be a useful approach for extending the duration of exercise in exergame play. However, the research to date primarily used ostensibly real partners. Extending these findings to different types of partners would be beneficial for future researchers, exergame designers, and practitioners who are looking for new motivational strategies to keep people engaged with exercise. This dissertation will explore the viability of two new types of partners, computer-generated and imagined, in the hopes that they could be a useful extension of the current Köhler effect research.
CHAPTER 3

METHOD

Data collection was conducted in two waves. The first wave focused on computer-generated partners, using an exergame program that had four conditions built into it: virtually-real (VR) partner, altered virtually-real (AVR) partner, software-generated (SG) partner, and the individual-control (IC) condition. This portion ran from December 2012 to April 2013. The second wave ran from September 2013 to November 2013 and focused on the imagery conditions: partner imagery (PI), associative imagery (AI), and dissociative imagery (DI). All study procedures were approved by the university’s Institutional Review Board (see Appendix A).

Design and Participants

The overall design for the experiment was a 7 (Condition: VR, AVR, SG, PI, AI, DI, IC) x 2 (Block: 1, 2) between-subjects design, with Block being a repeated measures variable. Participants were undergraduate psychology students who completed the study for course credit. Recruitment took place through the Psychology department’s experimental sign-up system. This system allows researchers to post brief descriptions of their studies, along with any special notes or criteria. Students view all the available experiments and are able to select which ones they would like to participate in by signing up for a timeslot. Students were disqualified from participation if they were under the age of 18 or if they had any injuries or illnesses that would affect their performance on the exercises (e.g., injuries to their arms, legs, or shoulders). An a priori power analysis was conducted with G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). Results of the power analysis showed that 23 participants in each of the seven conditions (for a
A total of 161 participants would be sufficient to obtain a medium effect size of 0.30 with a power of 0.8 and an alpha level of 0.05.

During the first wave of data collection, 112 participants (56 male, 56 female) were randomly assigned to one of four partner conditions: a “live” human partner (VR), a filtered human partner (AVR), a fully software-generated partner (SG), or a no-partner control condition (IC). During the second wave of data collection, 8 participants (4 male, 4 female) were randomly assigned to the VR, AVR, SG, or IC conditions to bring the per condition total to 30 participants. An additional 100 participants also completed the experiment during the second wave. Within that group of 100, 90 participants (45 male, 45 female) were randomly assigned to one of three imagery conditions: working with an imagined partner (PI), using associative imagery (AI), or using dissociative imagery (DI). To account for history effects and to determine if control group participants used imagery, an additional 10 participants (5 male, 5 female) were added to the IC group and were randomly assigned within the group of 100. Thus, overall in the second wave, 108 participants completed the experiment.

In total, 220 participants (110 male, 110 female) completed the experiment. Each condition contained 30 participants, with the exception of the IC condition, which contained 40 participants. The mean age of the participants was 19.47 (SD = 1.45) and the mean Body Mass Index (BMI) score was 23.24 (SD = 3.10). A majority of the participants were Caucasian (61.4%, n = 135) and were first- or second-year students (67.7%, n = 149).

**Experimental Task**

Keeping aligned with previous studies on the Köhler effect and exercise, the experimental task was holding a series of five abdominal plank exercises for as long as possible, consistent with one’s own health and wellbeing (Feltz et al., 2011; Forlenza et al., 2012). These
exercises do not require much coordination and are mostly effort-based, which makes them ideal for physical persistence tasks to gauge motivation gains and losses. There were 30 s breaks between each exercise. Each series of plank exercises was performed twice, with a 10 min break between each Block of exercises.

Participants completed a front plank, two side planks (right and left), and two one-legged planks (right and left) on an exercise mat. While on the mat, front planks required participants to face down, extend their legs straight out so they were on their toes, and prop their bodies in the air using their forearms, with only their toes and forearms touching the ground. The legs, back, and neck should have been aligned in a straight line. Side planks similarly required the body to be elevated off the ground. Participants shifted their bodies so that they were only on their right (or left) foot and right (or left) forearm, facing outwards. (Consider this a 90 degree rotation from the front plank position). Finally, the one-legged planks were like the front planks, except that the right (or left) leg was off the ground and held straight out in the air, with the left (or right) foot firmly on the ground.

**Measures**

**Persistence**

The primary outcome variable was exercise persistence, which was defined as the length of time the plank exercises were held (measured in seconds). At the end of a 10 s countdown, participants moved into the proper position. Once in position, the experimenter started a stopwatch. When participants stopped holding the plank and dropped back down to the mat, the stopwatch ended and the time was recorded. Block scores were calculated by summing the length of time each individual exercise was held.

**Rating of Perceived Exertion (RPE)**
After each individual plank exercise, participants rated what their overall feeling of exertion was at the moment just before they stopped holding the exercise. Participants used the Borg Scale for their exertion ratings, which ranges from 6 (no exertion at all) to 20 (maximal exertion), to rate their exertion levels (Borg, 1998; see Appendix B). Rating exertion with the Borg Scale was explained to participants before the start of the experiment (including the verbal anchors) and a large poster of the scale with anchors was placed in the room for easy reference. As reported in Borg (1998), the psychometric properties of the scale are acceptable, showing adequate retest reliability (> .70) and intratest reliability (.93). Strong correlations between self-ratings and observer ratings of over .90 were reported, suggesting content validity. Ratings of perceived exertion using the Borg Scale were also strongly correlated with heart rate (> .90) and were used to predict performance (with coefficients ranging from .56 to .82), suggesting adequate concurrent and predictive validity, respectively (Borg, 1998).

**Self-efficacy Beliefs (SE)**

Participants recorded (in seconds) how long they believed they could hold each exercise at three different time points: prior to Block 1, prior to Block 2, and after completing Block 2. Each SE measure consisted of five items, one for each plank. The second measurement (before Block 2) occurred after the partner or imagery manipulations and after participants were told the average length of time they held the exercises during Block 1 (see Appendix C).

**Enjoyment and Intention to Exercise**

Enjoyment was measured with an 8-item version of the Physical Activity Enjoyment Scale (PACES) following completion of the experiment (Raedeke, 2007; see Appendix D). The PACES asked participants to rate their enjoyment of the exercise on a 1 to 7 scale (e.g., 1 = loved it, 7 = hated it). Preliminary psychometric support for the longer, 18-item version of the PACES
came from two developmental studies, which found that the scale’s internal consistency was strong (Cronbach’s $\alpha = .93$) (Kendzierski & DeCarlo, 1991). Subsequent research found that participants reported higher levels of enjoyment when listening to music while exercising (versus no music), a negative correlation between likelihood to experience boredom and enjoyment existed, and that scores on the PACES predicted activity choice (Kendzierski & DeCarlo, 1991). The 8-item version was created due to concerns that items on the full scale assessed the antecedents and consequences of enjoyment in addition to enjoyment itself (Raedeke, 2007). The shortened version had a strong correlation ($r = .94$) with the complete scale (Raedeke, 2007). Additionally, participants were asked to rate the likelihood they intended to exercise the following day on a -3 (not at all true for me) to +3 (completely true for me) scale (Mohiyeddini, Pauli, & Bauer, 2009).

**Uncanny Valley**

Participants in four of the conditions (VR, AVR, SG, PI) completed two questionnaires designed to assess emotional responses to their partner. Both questionnaires asked participants to rate a series of bipolar items from 1 to 5 (e.g., 1 = artificial, 5 = natural; 1 = unfriendly, 5 = friendly). The first questionnaire was the Godspeed Indices, which consisted of 24 items and was composed of five sub-scales: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety (Bartneck, Kulić, Croft, & Zoghbi, 2008; see Appendix E). Though widely used in the literature, few studies have explored their psychometric properties. With the exception of the perceived safety sub-scale, both studies found that each sub-scale had adequate internal consistencies (all Cronbach’s $\alpha$ values > .80) (Ho & MacDorman, 2010; Schillaci, Bodiroža, & Hafner, 2013).
However, Ho and MacDorman (2010) concluded that the Godspeed Indices have poor discriminant validity and thus created a new questionnaire, the Alternative Godspeed Indices, which consisted of 19 items across three sub-scales: humanness, eeriness, and attractiveness (Ho & MacDorman, 2010; see Appendix F). These new sub-scales also demonstrated adequate internal consistency (all Cronbach’s α values > .70), but additionally, the sub-scales showed adequate discriminant validity and the overall questionnaire had its theoretical structure verified by a confirmatory factor analysis. Both questionnaires were used in this dissertation in order to obtain as much feedback as possible about the partners and to explore some of the psychometric properties of the questionnaires.

**Attitudes Towards Partner**

Participants in the four partner conditions (VR, AVR, SG, PI) completed two questionnaires assessing what they thought of their working relationship with their partner. Both questionnaires asked participants to rate a series of statements from *Strongly Disagree* to *Strongly Agree* (e.g., I felt like I was part of a team). The first questionnaire was a five-item Team Perception index previously used in human-computer interaction research (Nass et al., 1996; see Appendix G). This index showed adequate internal consistency (Cronbach’s α = .89). The second questionnaire was a six-item measure based on a group identification scale, modified for the exercise task (Brown, Condor, Matthews, Wade, & Williams, 1986; see Appendix H). This measure displayed adequate internal consistency (Cronbach’s α = .71). Participants who reported positive and negative comments about the group had significantly different scores on this measure, suggesting the measure examines positive and negative reactions to group affinity (Brown et al., 1986).

**Manipulation Checks**
In all conditions, participants were given an open-ended question asking, “Was anything confusing or odd about the experiment?” Participants in four of the conditions (VR, AVR, SG, PI) answered two questions asking about the experimental protocol to ensure they understood the conjunctive nature of the task (see Appendix I). Participants in the three imagery conditions (PI, AI, DI) completed an 11-item manipulation check to ensure they followed the imagery protocol (based on Razon et al., 2010, and Razon, 2012; see Appendix J). Participants in the second wave control group (n = 10) also answered three short questions to ensure they did not use imagery during the experiment (see Appendix K).

Demographics

At the end of the experiment, all participants were asked to self-report basic demographic information, including gender, age, class year, height, weight, and race/ethnicity (see Appendix L).

Procedures

The basic procedures of this experiment followed previous research (Feltz et al., 2011; Forlenza et al., 2012; Irwin et al., 2013). The lab in which the study took place was a small room in an academic building. The lab had a spot for participants to sit (small desk and chair) with a laptop, on which participants watched the instructional video and completed the questionnaires. The lab also had a desk, chair, and two computers for the experimenter. A projector screen was hung on a wall and a projector was placed on a small cart. Finally, an exercise mat was placed on the floor, which is where participants performed the exercises. A poster of the Borg Scale was hung on the wall, along with the order of the five plank exercises.

After arriving at the lab, participants gave their informed consent after reading the consent form, and watched a 9 min video that explained the experimental procedures by
demonstrating each of the five plank exercises and instructing participants on how to rate their exertion using the Borg Scale. Following the video, participants had an opportunity to ask questions. After any questions were answered, participants completed the first SE measure, then moved to the exercise mat and performed the first set of exercises, Block 1. Participants were instructed to hold each of the five plank exercises for as long as possible, consistent with their physical wellbeing, and to give their perceived exertion score following the completion of each exercise. After Block 1 finished, there was a 10 min break before Block 2, during which time the manipulations took place.

**Manipulations**

This section outlines the procedures for each of the seven conditions. For specific details of exactly what was said, please refer to Appendix M.

**Virtually-real (VR) partner.** In this condition, during the 10 min break, participants were told that they would be completing the same set of exercises a second time working with a same-sex partner towards a team score, which was defined by the score of the person who stopped holding the exercise first (thus creating conjunctive task demands). The experimenter explained to participants that their partner was another person in a different lab. Participants were led to believe that this other person was real by way of a simulated partner introduction. That is, the experimenter pretended to coordinate a webcam-like meeting between the two labs by pretending to speak with an experimenter in the other lab. After a brief exchange, the experimenter facilitated an introduction by having the partner and participants share basic information (e.g., name, hometown, hobbies). This was done for two reasons: to enhance the believability of the manipulation and to remove anonymity and acquaint the participants with their partner. Participants believed this was “live,” but in fact, the video was pre-recorded and
manipulated with a hidden slide advancer. The pre-recorded video presented a frontal view of the partner (i.e., head and shoulders) wearing exercise clothes. After the introduction, the conjunctive nature of the task was explained again to ensure understanding. Next, participants were truthfully given the average of how long they held each of the five exercises, and falsely given the average of how long their partner held each exercise. The partner’s average was always 40% larger than participants’ averages to establish an upwards social comparison. Following this, participants completed the second SE measure and were told to sit quietly on the exercise mat until the experimenter indicated that the second block would begin.

**Altered virtually-real (AVR) and software-generated (SG) partners.** In these conditions, the manipulations were the same as in the VR condition, with a few exceptions. First, participants were informed that their partner was not a real person, but was instead computer-generated. Second, participants were still introduced to their partner, but because the partner was not real, the experimenter did not need to pretend to coordinate a meeting with people in a separate lab. Third, the partner’s image was different in these conditions. In the VR condition, participants were presented with a pre-recorded person. In the AVR condition, this same video was used, however, a computerized filter effect was applied to the image to make it appear almost animated. In the SG condition, partners were 3D graphical characters rigged to perform plank exercises. Thus, these conditions formed a continuum from most realistic (VR) to least realistic (SG) (see Figure II). Finally, when participants were told their average and their partner’s average, they were also told that although their computer-generated partner was programmed to be somewhat better at holding the exercises, the partner could not hold the exercises forever. So like a real person, this partner would eventually become fatigued and have to stop holding the exercises.
Figure II. Pictures of the three male partners; from left to right, virtually-real (VR), altered virtually-real (AVR), and software-generated (SG).

**Associative imagery (AI).** In this condition, during the 10 min break, participants were told that they would be completing the same set of exercises a second time, only with the use of imagery. Participants were given a definition of imagery, a short explanation of how it is used, and a description of three critical components to effective imagery (vividness, controllability, use of multiple senses). Following this explanation, participants were told that they will use associative imagery, or imagery that is focused on the task, while they are holding the exercises during the second block. Participants were instructed to draw motivation from imagining themselves holding the exercises successfully while they were performing them, such as by seeing themselves holding the exercises firmly, feeling stronger the longer they held it, or by feeling more confident and focused. Following this explanation, participants were told the average length of time they held the exercises during the first block and completed the second SE measure. The experimenter then directed participants to practice this imagery while sitting on the exercise mat, but to also verbalize the imagery to ensure participants understood the types of images they should be creating. Experimenters were instructed to gently correct participants if
the imagery was not task-relevant. The practice imagery session lasted until the final minute of the 10 min break or until participants stopped practicing.

**Dissociative imagery (DI).** The manipulation in the DI condition was the same as in the AI condition, but with one major exception. Here, participants were told that while they are holding the exercises during the second block, they will use dissociative imagery, or imagery that is focused on something unrelated to the task. Participants were instructed to draw motivation from imagining themselves doing something that they found to be motivating but unrelated to the plank exercises, such as achieving a lifelong goal or feeling confident and focused during a challenging situation. Experimenters were instructed to gently correct participants if the imagery was task-relevant.

**Partner imagery (PI).** Manipulations in this condition were a combination of the manipulations in the previous conditions. The language used was similar to the language used in the VR, AVR, and SG conditions, with the exception that participants were told they would be working with an imagined partner. As such, participants were told that they will be working with a partner on the second block of exercises towards a team score (defined by the score of the person who quit first), except that their partner would be created by their own imagination. This was followed by the same explanation of imagery and its critical components as in the AI and DI conditions. The introduction with the partner was imagined; participants were instructed to imagine themselves meeting a same-sex person they did not know who was wearing exercise clothes. Participants were then told the average length of time they held the exercise and the average of how long their imagined partner held the exercises, which was 40% longer. As with the AVR and SG conditions, participants were instructed that their imagined partner should be unable to hold the exercises forever. Next, participants completed the second SE measure, then
practiced their imagery aloud. Participants were instructed to create a clear image of their partner performing the exercises successfully.

**Individual-control (IC).** During the 10 min break, participants in this condition were instructed to wait quietly for further instructions. After a couple minutes, participants were told that they would perform the same set of exercises again after a break. They were told truthfully the average of how long they held the exercises during Block 1 and then completed the second SE measure. After this, participants were instructed to continue to wait quietly until the second block began.

**During Block 2.** Following the manipulations and the break, participants were given a brief recap of the instructions immediately before performing Block 2, which was the same as Block 1, but with minor changes. In the VR, AVR, and SG conditions, the exergame was projected onto a screen. Participants were informed that they could see how their partner was doing, and their partner could see how they were doing. In the VR condition, similar to the introduction, the partner’s images were a series of pre-recorded videos manipulated by the experimenter. Thus, the partner always out-performed the participants, which meant that the participants always stopped exercising first. The AVR condition was exactly the same as the VR condition only with the computerized filter effect present. The SG partner was in a similar setting, only it was clearly artificial, much like the partner, as the same program was used for both.

In the PI, AI, and DI conditions, participants were told that they should start using imagery as soon as they began holding each plank and continue to use imagery until they stopped holding the plank. In the IC condition, participants performed the Block 2 exercises without a partner and without using imagery. Thus, it was exactly the same as Block 1. Following Block 2,
participants completed the third SE measure, post-experimental questionnaires, were thanked, and were given a debriefing sheet.

**Analyses**

**Persistence**

To analyze persistence scores, the times for each exercise were summed within each Block. The primary dependent variable was the difference score between both blocks (Block 2 – Block 1). This approach can show any changes in persistence while controlling for individual differences in fitness. The difference scores were analyzed in a 7 (Condition: VR, AVR, SG, PI, AI, DI, IC) x 2 (Participant Gender: Male, Female) between-subjects analysis of variance (ANOVA), supplemented with planned contrasts and post-hoc tests. (This approach generally produces the same pattern of results as more rigorous analyses, e.g., Forlenza et al., 2012, and Kerr et al., 2013. See Appendix N for a summary of these additional analyses).

**Secondary Analyses**

Participants in all seven conditions completed measures of RPE, SE, enjoyment, and intention to exercise the following day. For RPE, a 7 (Condition) x 2 (Gender) x 2 (Time: Block 1, Block 2) ANOVA with repeated measures on Time was performed. For SE, scores at the first measurement point (prior to Block 1) were used as a covariate in a 7 (Condition) x 2 (Gender) x 2 (Time: SE prior to Block 2, SE following Block 2) ANCOVA with repeated measures on Time. For both enjoyment and intention to exercise, 7 (Condition) x 2 (Gender) ANOVAs were performed, and post-hoc tests were used to determine any significant differences between the conditions when applicable.

Participants in the four partner conditions (VR, AVR, SG, PI) also completed two measures of the uncanny valley and two measures of attitudes towards their partner. A 4
(Condition) x 2 (Gender) MANOVA on the uncanny valley measures’ subscales was performed. For both team perception and group identification, a 4 (Condition) x 2 (Gender) ANOVAs was performed, and post-hoc tests were used to determine any significant differences between the conditions when applicable. For all analyses, the alpha level was set at .05.
CHAPTER 4

RESULTS

The purpose of this study was to explore the Köhler effect in exercise settings using different types of partners and to extend previous research on imagery to plank exercises. This chapter is organized into three major sections. The first section presents preliminary analyses, including data screening and the various manipulation checks that were used to ensure that the partner and imagery protocols were followed. The second section presents the major analyses, including descriptive statistics and hypothesis testing. The third section presents secondary analyses, including descriptive statistics, hypothesis testing, and correlations.

Preliminary Analyses

Missing Data

For the main dependent variable (difference scores), there was one count of missing data. For the first set of exertion scores, there was one count of missing data, and there were eight counts of missing data for the second set. For the second measurement point of self-efficacy beliefs, there was one count of missing data, and seven counts of missing data for the third measurement point. For enjoyment, there were six counts of missing data, and for intention to exercise the following day, there were five counts of missing data. These cases were not included in their respective analyses.

Outliers

The data were assessed for univariate and multivariate outliers. Univariate outliers fell outside of \( \pm 3 \, SD \) from the overall mean. Based on \( z \)-scores, several outliers were identified. Four outliers were identified for the difference scores, one each for the first and second set of exertion scores, four each for the first and second self-efficacy belief measurements, and two for the final
self-efficacy belief measurement. All analyses were conducted twice, once with and without the outliers. The results of the analyses did not differ; therefore, results from the more inclusive analyses (with the univariate outliers) were reported.

Multivariate outliers were assessed for the analyses of exertion scores and self-efficacy beliefs by calculating Mahalanobis’ distance. Results indicated no outliers for exertion scores and six outliers for self-efficacy beliefs. These cases were removed and the analysis was conducted a second time. The results did not differ; therefore, the full analysis (with the multivariate outliers) was reported.

**Normality**

Univariate normality was assessed using skewness statistics for each of the main variables. Only two instances of skewness were observed for the first (2.52) and third (5.89) self-efficacy belief measurement points. A square root transformation was applied to each variable, which resulted in normal distributions (Raykov & Marcoulides, 2008). The self-efficacy analysis was conducted twice, once with normal data and once with skewed data. The pattern of results was the same; therefore, the non-transformed analysis is reported.

**Manipulation Check – Partner Conditions**

Results from the partner manipulation check indicated that most participants in the VR (87%, n = 26), AVR (90%, n = 27), SG (83%, n = 25), and PI (87%, n = 26) conditions understood that they performed these exercises as part of a two-person team. Additionally, participants in the VR (89.7%, n = 26), AVR (79.3%, n = 23), SG (75.6%, n = 22), and PI (66.7%, n = 20) conditions generally understood the conjunctive nature of the task by indicating that the team score was the time of whomever quit holding the exercise first.
A total of 13 participants reported suspicion on an open-ended question asking if there was anything confusing or odd about the experiment. These participants either suspected that their partner was not real \((n = 9)\) or that their partner would not tire and could hold the exercises indefinitely \((n = 4)\). The main analysis was performed with and without the suspicious participants, but the pattern of results did not change. The analysis with the suspicious participants included is reported.

**Manipulation Check – Imagery Conditions**

Before analyzing the specific imagery manipulation check questions, a 3 (Condition: AI, DI, PI) x 2 (Gender) ANOVA was performed to explore if there were any significant differences on the amount of imagery practice time (measured in seconds). Results indicated a significant main effect for gender, \(F(1,84) = 4.08, p = .047, \eta^2 = .05\), such that female participants \((M = 160.36, SD = 34.52)\) practiced their imagery more than male participants \((M = 143.49, SD = 43.76)\). No main effect emerged for Condition, indicating that participants in the AI \((M = 149.67, SD = 47.49)\), DI \((M = 157.03, SD = 41.88)\), and PI \((M = 149.07, SD = 29.74)\) groups practiced their imagery for approximately the same amount of time.

Overall, participants in the three imagery conditions reported that their imagery was vivid \((M = 8.14, SD = 1.99)\), under their control \((M = 8.03, SD = 2.33)\), and that they were engaged with the images \((M = 8.16, SD = 1.94)\). The scenes were moderately challenging to generate \((M = 5.49, SD = 2.67)\), but participants reported they were able to generate them early during each exercise \((M = 3.08, SD = 1.80)\) and that they held the images in their mind for most of the length of each exercise \((M = 7.72, SD = 2.25)\). Participants also reported that they were motivated to deal with the effort required by the exercises \((M = 7.84, SD = 1.98)\) and to reduce their exertion during the exercises \((M = 6.92, SD = 2.67)\), that the imagery helped them to tolerate their
exertion \((M = 6.89, SD = 2.41)\), but that they still perceived a heavy amount of exertion \((M = 7.34, SD = 2.06)\). Participants also indicated a general likeliness to use imagery while exercising in the future \((M = 7.58, SD = 2.88)\). Refer to Table 1 for means and standard deviations for each condition for the imagery manipulation check.

Only one significant difference emerged by condition: based on a Student-Newman-Keuls (SNK) post-hoc test, participants in the PI condition \((M = 8.10)\) reported perceiving more exertion while using imagery compared to participants in the DI condition \((M = 7.10)\) and AI condition \((M = 6.83)\), \(F(2,85) = 3.26, p = .043\). Only one significant difference emerged by gender: female participants \((M = 2.58)\) reported that they started to imagine the scene quicker than male participants \((M = 3.59)\), \(t = 2.75, p = .007\).

There were also several significant correlations between items on the imagery manipulation check and difference scores (see Table 2). Overall, small-to-moderate positive correlations \((r\) values ranged from .30 to .40) emerged for feeling engaged with the images, being motivated to deal with the effort required by the exercises, feeling that imagery helped them tolerate the exertion, and indicating a likeliness to use imagery in the future. These correlations make sense – for example, it stands to reason that people who were engaged with their images and felt that it helped them would hold the exercises for longer than people who were not engaged with their imagery or who found their images unhelpful.
Table 1.

Means and Standard Deviations of the Imagery Manipulation Check by Condition.

<table>
<thead>
<tr>
<th>Imagery Scene Check</th>
<th>AI</th>
<th>DI</th>
<th>PI</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vividness and clarity of imagery</td>
<td>7.97 (2.19)</td>
<td>8.17 (2.25)</td>
<td>8.30 (1.49)</td>
<td>8.14 (1.99)</td>
</tr>
<tr>
<td>Control over imagery</td>
<td>7.30 (2.52)</td>
<td>8.40 (2.34)</td>
<td>8.40 (1.98)</td>
<td>8.03 (2.33)</td>
</tr>
<tr>
<td>Engaged with the images</td>
<td>7.67 (2.32)</td>
<td>8.40 (1.81)</td>
<td>8.40 (1.59)</td>
<td>8.16 (1.94)</td>
</tr>
<tr>
<td>Ease of generating images</td>
<td>5.73 (2.72)</td>
<td>5.07 (2.82)</td>
<td>5.69 (2.49)</td>
<td>5.49 (2.67)</td>
</tr>
<tr>
<td>Time to generate images</td>
<td>3.34 (2.02)</td>
<td>2.60 (1.43)</td>
<td>3.30 (1.88)</td>
<td>3.08 (1.80)</td>
</tr>
<tr>
<td>Length of holding imagined scenes</td>
<td>7.73 (2.46)</td>
<td>7.80 (2.19)</td>
<td>7.63 (2.17)</td>
<td>7.72 (2.25)</td>
</tr>
<tr>
<td>Motivated to deal with effort</td>
<td>7.67 (2.12)</td>
<td>7.60 (1.92)</td>
<td>8.27 (1.89)</td>
<td>7.84 (1.98)</td>
</tr>
<tr>
<td>Motivated to reduce exertion</td>
<td>6.83 (2.79)</td>
<td>7.27 (2.64)</td>
<td>6.67 (2.63)</td>
<td>6.92 (2.67)</td>
</tr>
<tr>
<td>Imagery helped tolerate exertion</td>
<td>6.27 (2.45)</td>
<td>7.10 (2.38)</td>
<td>7.31 (2.35)</td>
<td>6.89 (2.41)</td>
</tr>
<tr>
<td>Perception of exertion during imagery</td>
<td>6.83 (2.35)</td>
<td>7.10 (1.54)</td>
<td>8.10 (2.02)</td>
<td>7.34 (2.06)</td>
</tr>
<tr>
<td>Likelihood to use in the future</td>
<td>7.87 (2.64)</td>
<td>7.07 (3.15)</td>
<td>7.80 (2.86)</td>
<td>7.58 (2.88)</td>
</tr>
</tbody>
</table>

Note. All items rated on a 0-10 scale, with scores of 10 representing a higher amount of that item (e.g., a score of 10 on “engaged with the images” indicates a high degree of involvement with one’s created images), with two exceptions noted.

*High scores indicate greater difficulty imagining images. **High scores indicate that it took longer to generate images.
Table 2.

*Correlations Between Difference Scores and Imagery Manipulation Check by Condition.*

<table>
<thead>
<tr>
<th>Imagery Scene Check</th>
<th>AI</th>
<th>DI</th>
<th>PI</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vividness and clarity of imagery</td>
<td>0.19</td>
<td>0.08</td>
<td>-0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Control over imagery</td>
<td>0.19</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>Engaged with the images</td>
<td>0.40*</td>
<td>0.24</td>
<td>0.29</td>
<td>0.33**</td>
</tr>
<tr>
<td>Ease of generating images</td>
<td>0.09</td>
<td>-0.06</td>
<td>-0.23</td>
<td>-0.03</td>
</tr>
<tr>
<td>Time to generate images</td>
<td>0.17</td>
<td>-0.30</td>
<td>-0.24</td>
<td>-0.07</td>
</tr>
<tr>
<td>Length of holding imagined scenes</td>
<td>-0.13</td>
<td>0.28</td>
<td>0.29</td>
<td>0.11</td>
</tr>
<tr>
<td>Motivated to deal with effort</td>
<td>0.42*</td>
<td>0.43*</td>
<td>0.30</td>
<td>0.40**</td>
</tr>
<tr>
<td>Motivated to reduce exertion</td>
<td>0.25</td>
<td>0.10</td>
<td>0.10</td>
<td>0.14</td>
</tr>
<tr>
<td>Imagery helped tolerate exertion</td>
<td>0.42*</td>
<td>0.33</td>
<td>0.14</td>
<td>0.32**</td>
</tr>
<tr>
<td>Perception of exertion during imagery</td>
<td>0.42*</td>
<td>-0.22</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>Likelihood to use in the future</td>
<td>0.31</td>
<td>0.54**</td>
<td>-0.05</td>
<td>0.30**</td>
</tr>
</tbody>
</table>

* = p < .05; ** = p < .01

Additionally, participants in the DI condition reported that their scenes were mostly unrelated to the task (M = 2.14), while participants in the AI (M = 8.27) and PI (M = 9.40) conditions reported their scenes were mostly related to the task. The differences between the DI
and AI conditions and between the DI and PI conditions were statistically significant based on SNK post-hoc tests, $F(2,86) = 87.04, p < .001$. Specific to the PI condition, a majority of the participants (80%, $n = 24$) reported that they stopped holding each exercise before their imagined partner. These results indicate that most of the participants in the three imagery conditions followed the imagery instructions.

Results from open-ended questions generally confirmed that participants imaged the proper types of scenes. Nearly all participants in the AI condition kept the focus of their imagery on themselves ($n = 29$; e.g., “I was imaging that i was doing each of the position correctly without body shaking and my muscles were working perfectly without being weak and sour; and i was confidently smiling without feeling any exertion” [sic]). No participants reported exercising with a partner; however, five reported that people were cheering them. All participants in the DI condition reported imagining scenes unrelated to the task ($n = 30$; e.g., “I imagine myself the day of graduation. Walking down as my name was being called. Receiving my diploma, family and close friends were there to support me. Nothing but positive atmosphere was flowing through the air. I was happy, and so was my family and that was all that mattered”). All participants in the PI condition reported imagining an exercise partner ($n = 30$; e.g., “I imagined a girl who was in way better shape then me so I could try to think in my mind that I would hold it longer then her” [sic]). However, five participants reported seeing their partner as competition instead of as a teammate. (This may have sometimes been the case for participants in the other partner conditions, too. However, this was not explicitly measured).

A majority of participants in the imagery conditions reported using the visual (97.8%, $n = 88$), kinesthetic (85.6%, $n = 77$), and auditory (76.7%, $n = 69$) senses. Fewer participants reported using smell (15.6%, $n = 14$) or taste (6.7%, $n = 6$) during their imagined scenes. When
asked to report which sense they used the most while imaging, sight was chosen most frequently (61.1%, n = 55), with hearing (18.9%, n = 17) and kinesthetic feel (17.8%, n = 16) chosen the next most.

Finally, participants in the second wave IC condition (n = 10) did not report using imagery during the plank exercises. Instead, participants reported thinking about how long they were holding the exercises (n = 4), counting or backwards counting (n = 3), thinking about trying to go longer (n = 2), and the discomfort of holding the exercises for as long as possible (n = 1). Participants reported thinking those thoughts for various reasons, such as setting a goal and beating the previous time (n = 4), to know how long they were holding the exercises (n = 2), and because the feelings of fatigue were too strong not to think about (n = 2). When participants first experienced fatigue, their thoughts centered on persisting and continuing to hold the exercises for a few more seconds (n = 5), that they would need to stop holding the exercise soon (n = 2), and that they should exercise more often (n = 2).

**History Effects**

To test for history effects, independent samples t-tests were conducted between participants in the first wave (n = 112) and second wave (n = 108). For demographic characteristics, results indicated no significant differences for age (Wave 1 = 19.45, Wave 2 = 19.49, p = .852), while a trend emerged for BMI differences (Wave 1 = 23.64, Wave 2 = 22.82, p = .056). For the dependent variables (Block 1 sum, Block 2 sum, difference scores, RPE averages for Blocks 1 and 2, SE scores at Times 1, 2, and 3, enjoyment, and intention to exercise), no significant differences emerged between the waves (p values > .08).

This analysis was also conducted by Condition for participants in the VR, AVR, SG, and IC groups. (Data for the PI, AI, and DI conditions were all collected during the second wave and
thus do not have first wave data). Independent samples $t$-tests revealed no significant differences between the waves in each of the four conditions for all of the dependent variables ($p$ values > .07) and for age and BMI ($p$ values > .19), with one exception: participants in the second wave AVR group reported a significantly lower BMI ($M = 19.87, SD = 1.55$) compared to participants in the first wave AVR group ($M = 23.59, SD = 3.02$), $p = .048$. Because this one difference should not bias the results (the ability to perform simple plank exercises should not differ between participants who are underweight and participants who are in a normal weight range), each of the waves for each condition was combined.

This created unequal-sized groups (all conditions but the IC had 30 participants each). Therefore, the analyses that utilized a control group (i.e., difference scores, RPE, SE, enjoyment, and intention to exercise) were also run with a sub-group of 30 IC participants randomly drawn from the larger control group in place of the 40 person control group. For all analyses, the patterns of results were the same. Below, the analyses conducted with the full sample are reported.

**Main Analysis**

To control for individual differences in strength and fitness, difference scores were calculated between the two sets of exercises by using Block 1 performance scores as a baseline measure. Thus, the primary dependent variable was the Block 2 – Block 1 difference in performance as measured in seconds, which represents participants’ Block 2 scores relative to their own Block 1 scores. This approach has been used in the research previously (Kerr et al., 2007; Kerr et al., 2012). Descriptive statistics are presented in Tables 3 and 4 and the correlations between all dependent variables are presented in Table 5. A preliminary one-way
ANOVA of Condition on Block 1 scores revealed no significant differences ($p = .51$), suggesting the groups were similar at baseline.

A 7 (Condition: VR, AVR, SG, PI, AI, DI, IC) x 2 (Gender: Male, Female) ANOVA was conducted on the block difference scores. Results indicated significant main effects for Condition, $F(6,205) = 7.14, p < .001, \eta^2 = .17$ and Gender, $F(1,205) = 5.95, p = .016, \eta^2 = .03$.

The ANOVA was supplemented with three planned contrasts that tested the hypotheses. The first contrast compared the control group with the remaining conditions and was significant ($p < .001$), supporting Hypothesis 1. The second contrast compared the VR group with the other partner conditions (AVR, SG, PI) and was significant ($p = .001$), supporting Hypothesis 2. The third contrast compared the partner conditions (VR, AVR, SG, PI) with the other two imagery conditions (AI, DI) and was not significant ($p = .40$), not supporting Hypothesis 3. (Although difference scores control for initial individual variations in pre-existing strength and ability and are easy to interpret, there are some shortcomings with the approach (Edwards, 2001; Knapp & Schafer, 2009; Senn, 2006). Alternative analyses are presented in Appendix N).
Table 3.

**Overall Means and Standard Deviations for All Dependent Variables.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference score (sec)</td>
<td>-6.73</td>
<td>55.08</td>
</tr>
<tr>
<td>RPE Block 1</td>
<td>14.09</td>
<td>1.73</td>
</tr>
<tr>
<td>RPE Block 2</td>
<td>14.48</td>
<td>1.93</td>
</tr>
<tr>
<td>SE Time 2 (sec)</td>
<td>216.24</td>
<td>119.94</td>
</tr>
<tr>
<td>SE Time 3 (sec)</td>
<td>166.06</td>
<td>129.15</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>4.32</td>
<td>1.01</td>
</tr>
<tr>
<td>Intention to exercise</td>
<td>1.69</td>
<td>1.82</td>
</tr>
</tbody>
</table>

*Note.* Rating of perceived exertion was measured with a 6-20 scale, with higher mean scores representing greater exertion. Enjoyment was measured on a 1-7 scale, with higher mean scores indicating greater task enjoyment. Intention to exercise was measured on a -3 to +3 scale, with higher mean scores indicating greater intention to exercise the following day.

Student-Newman-Keuls (SNK) post-hoc tests revealed a number of significant differences. Participants in the VR condition \((M = 31.76, SD = 58.51)\) had significantly larger difference scores (indicating greater persistence during Block 2 relative to their own Block 1 performance) than participants in the AI \((M = -3.43, SD = 58.10)\), DI \((M = -2.83, SD = 54.31)\), AVR \((M = -9.77, SD = 41.89)\), SG \((M = -18.40, SD = 43.40)\), and IC \((M = -42.58, SD = 55.07)\) conditions. Additionally, participants in the PI \((M = 11.37, SD = 39.21)\), AI, DI, and AVR conditions had significantly larger difference scores than participants in the IC condition. The difference between the SG and IC conditions trended towards significance \((p = .061)\), while the difference between the VR and PI conditions did not approach significance \((p = .114)\). Regarding gender, female participants \((M = 1.50, SD = 52.26)\) had larger difference scores compared to
male participants ($M = -14.88, SD = 56.80$). All condition means are plotted in Figure III. Additional post-hoc tests are presented in Appendix N.

*Figure III.* Mean Block 2 – Block 1 difference scores (in seconds) with 95% confidence intervals. Negative difference scores indicate reduced performance compared to baseline, while positive difference scores indicate greater performance compared to baseline.
Table 4.

*Means and Standard Deviations for All Dependent Variables by Condition.*

<table>
<thead>
<tr>
<th>Measure</th>
<th>VR</th>
<th>AVR</th>
<th>SG</th>
<th>PI</th>
<th>AI</th>
<th>DI</th>
<th>IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>score (sec)</td>
<td>(58.51)</td>
<td>(41.89)</td>
<td>(43.40)</td>
<td>(39.21)</td>
<td>(58.10)</td>
<td>(54.31)</td>
<td>(55.07)</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(1.61)</td>
<td>(1.64)</td>
<td>(1.97)</td>
<td>(1.31)</td>
<td>(2.08)</td>
<td>(2.05)</td>
</tr>
<tr>
<td>RPE Block 2</td>
<td>15.10</td>
<td>15.01</td>
<td>14.41</td>
<td>14.93</td>
<td>14.53</td>
<td>13.77</td>
<td>13.91</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(2.13)</td>
<td>(1.66)</td>
<td>(2.09)</td>
<td>(1.84)</td>
<td>(1.84)</td>
<td>(2.06)</td>
</tr>
<tr>
<td>SE Time 2</td>
<td>241.31</td>
<td>181.06</td>
<td>179.35</td>
<td>243.69</td>
<td>256.75</td>
<td>188.41</td>
<td>214.64</td>
</tr>
<tr>
<td>SE Time 3</td>
<td>178.95</td>
<td>141.67</td>
<td>148.12</td>
<td>167.57</td>
<td>182.55</td>
<td>138.13</td>
<td>198.23</td>
</tr>
<tr>
<td>(sec)</td>
<td>(17.75)</td>
<td>(18.07)</td>
<td>(17.74)</td>
<td>(17.44)</td>
<td>(17.74)</td>
<td>(17.43)</td>
<td>(15.59)</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>3.97</td>
<td>4.34</td>
<td>3.98</td>
<td>4.49</td>
<td>4.43</td>
<td>4.61</td>
<td>4.36</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(0.86)</td>
<td>(1.00)</td>
<td>(1.07)</td>
<td>(1.30)</td>
<td>(0.95)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>Intention to</td>
<td>2.00</td>
<td>1.73</td>
<td>1.63</td>
<td>1.83</td>
<td>1.43</td>
<td>1.63</td>
<td>1.59</td>
</tr>
<tr>
<td>exercise</td>
<td>(1.49)</td>
<td>(1.91)</td>
<td>(2.04)</td>
<td>(1.68)</td>
<td>(2.05)</td>
<td>(1.88)</td>
<td>(1.79)</td>
</tr>
</tbody>
</table>

*Note.* Rating of perceived exertion was measured with a 6-20 scale, with higher mean scores representing greater exertion. Enjoyment was measured on a 1-7 scale, with higher mean scores indicating greater task enjoyment. Intention to exercise was measured on a -3 to +3 scale, with higher mean scores indicating greater intention to exercise the following day.
Table 5.

*Correlations Between All Dependent Variables.*

<table>
<thead>
<tr>
<th></th>
<th>Difference score</th>
<th>RPE Block 1</th>
<th>RPE Block 2</th>
<th>SE Time 2</th>
<th>SE Time 3</th>
<th>Enjoyment</th>
<th>Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference score</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPE Block 1</td>
<td>-0.01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPE Block 2</td>
<td>0.09</td>
<td>0.82**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE Time 2</td>
<td>-0.13</td>
<td>-0.01</td>
<td>0.04</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE Time 3</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.03</td>
<td>0.68**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>-0.04</td>
<td>0.10</td>
<td>0.05</td>
<td>0.15*</td>
<td>0.12</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intention</td>
<td>-0.10</td>
<td>-0.07</td>
<td>-0.11</td>
<td>0.26**</td>
<td>0.21**</td>
<td>0.29**</td>
<td>1</td>
</tr>
</tbody>
</table>

* = p < .05; ** = p < .01

Secondary Analyses

Rating of Perceived Exertion (RPE)

RPE scores were averaged across each Block, thus creating an average RPE score for Block 1 and an average RPE score for Block 2. A preliminary one-way ANOVA revealed no significant differences between conditions during Block 1 (p = .487). To analyze RPE scores, a 7 (Condition) x 2 (Gender) x 2 (Time: Block 1, Block 2) ANOVA was conducted with repeated measures on Time. Results indicated a significant Time x Condition interaction, F(6,198) = 3.12, Wilks’ Λ = .91, η² = .09, p = .006. To determine where the significant differences occurred, seven follow-up dependent samples t-tests were conducted between both RPE measures for each condition. (The alpha level was adjusted by dividing .05 by 7, for a new alpha level of .007). Results showed significant differences for the VR (p = .003), AVR (p < .001), SG (p = .003), and PI (p = .005) conditions, suggesting that participants in these conditions perceived themselves as
working harder during Block 2, whereas participants in the AI ($p = .820$), DI ($p = .884$), and IC ($p = .161$) conditions did not perceive themselves as working harder during Block 2, partially supporting Hypothesis 4. (Refer to Table 4 for means and standard deviations).

**Self-efficacy Beliefs (SE)**

SE was measured at three times during the experiment, and for each measurement, SE responses were summed, creating three SE scores. According to SE theory, SE scores for an upcoming task should correlate to actual performance on that task (Bandura, 1997). To test for this, two correlations were conducted. The first was between the first SE measure (prior to Block 1) and Block 1 performance, which was significant ($r = .56, p < .001$). The second was between the second SE measure (prior to Block 2) and Block 2 performance, which was also significant ($r = .62, p < .001$).

A preliminary one-way ANOVA suggested no significant differences between conditions on the first measurement point ($p = .751$). To control for individual differences in initial SE, the first SE sum was used as a covariate in a 7 (Condition) x 2 (Gender) x 2 (Time: SE prior to Block 2, SE following Block 2) ANCOVA with repeated measures on Time. Results indicated a significant main effect only for Time, $F(1,198) = 37.50$, Wilks’ $\Lambda = .84$, $\eta^2 = .16$, $p < .001$. SE prior to Block 2 ($M = 215.03$) was significantly larger than SE following Block 2 ($M = 165.03$). There were no significant differences between the conditions, supporting Hypothesis 5. Descriptive statistics are presented in Table 4.

**Enjoyment and Intention to Exercise**

Responses on the PACES were averaged to obtain an exercise enjoyment score. A reliability analysis was conducted on the eight items of the PACES. An adequate Cronbach’s $\alpha$ was obtained ($\alpha = .87$). The 7 (Condition) x 2 (Gender) ANOVA on this score yielded no
significant interaction or main effects ($p$ values > .11), supporting Hypothesis 5. The overall mean score of 4.31 ($SD = 1.01$) indicated a general enjoyment of the task, as this value was significantly above the scale’s midpoint ($p < .001$).

The 7 (Condition) x 2 (Gender) ANOVA on the intention to exercise item yielded a significant main effect only for Gender, $F(1,214) = 5.54, \eta^2 = .03, p = .020$, supporting Hypothesis 5. Results showed that males ($M = 1.99, SD = 1.59$) indicated a greater intention of exercising for at least 30 minutes the following day compared to females ($M = 1.40, SD = 1.99$). The overall mean of 1.69 ($SD = 1.82$) indicated a general intention to exercise the following day, as this value was significantly above the scale’s midpoint ($p < .001$). Refer to Table 4 for means and standard deviations for enjoyment and intention to exercise by condition.

**Uncanny Valley**

Two questionnaires were used to assess emotional responses to one’s partner. Because three of the conditions did not have partners, the following analyses only took place with the four partner conditions (VR, AVR, SG, PI). For both questionnaires, averages were calculated for each individual subscale. Refer to Table 6 for means and standard deviations for each Uncanny Valley measure by condition and to Table 7 for correlations between the secondary measures and difference scores.

**Godspeed Indices.** Reliability analyses were conducted with the five subscales from the Godspeed Indices. An adequate Cronbach’s $\alpha$ was obtained for Anthropomorphism ($\alpha = .90$), Animacy ($\alpha = .95$), Likeability ($\alpha = .93$), and Intelligence ($\alpha = .92$). The Safety subscale demonstrated poor reliability, however ($\alpha = .45$). Next, correlations between the subscales were calculated (see Table 8). A significant strong positive correlation was observed between Anthropomorphism and Animacy ($r = .90$), suggesting the two subscales do not measure unique
constructs. The remaining correlations were weak to moderate in magnitude ($r$ values ranged from .21 to .61). These results are similar to the findings by Ho and MacDorman (2010), who found that the Godspeed Indices had poor divergent validity. Therefore, the Godspeed Indices may not be psychometrically adequate, and thus, no further analyses were conducted with this measure.

**Alternative Godspeed Indices.** Reliability analyses were conducted with the three subscales from the Alternative Godspeed Indices. An adequate Cronbach’s $\alpha$ was obtained for the Humanness ($\alpha = .96$), Eeriness ($\alpha = .78$), and Attractiveness ($\alpha = .91$) subscales. Next, correlations between the subscales were calculated (see Table 9). Results indicated significant weak to moderate correlations ($r$ values ranged from .34 to .52), suggesting the sub-scales measured related but distinct constructs. Following this, a 4 (Condition) x 2 (Gender) MANOVA on the subscales revealed significant main effects for Condition, $F(3,98) = 6.86$, Wilks’ $\Lambda = .57$, $\eta^2 = .17$, $p < .001$ and Gender, $F(1,98) = 7.29$, Wilks’ $\Lambda = .81$, $\eta^2 = .19$, $p < .001$.

SNK post-hoc tests indicated that the significant differences for Condition were observed on the Humanness and Attractiveness subscales. Specifically, for Humanness, participants in the PI ($M = 4.03$, $SD = 0.92$) condition rated their partners as significantly more humanlike than participants in the VR ($M = 3.27$, $SD = 1.19$), AVR ($M = 2.81$, $SD = 0.84$), and SG ($M = 1.98$, $SD = 1.06$) conditions. Additionally, participants in the SG condition rated their partners significantly less humanlike than participants in the remaining conditions. For Attractiveness, participants in the PI ($M = 3.97$, $SD = 0.71$) condition rated their partners as significantly more attractive than participants in the VR ($M = 3.32$, $SD = 0.58$), AVR ($M = 3.27$, $SD = 0.67$), and SG ($M = 2.93$, $SD = 0.79$) conditions.
Table 6.

*Means and Standard Deviations for the Partner Conditions for Emotional Responses to their Partners and Attitudes Toward their Partners.*

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Subscale</th>
<th>VR</th>
<th>AVR</th>
<th>SG</th>
<th>PI</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godspeed</td>
<td>Anthropomorphism</td>
<td>2.99</td>
<td>3.04</td>
<td>2.25</td>
<td>3.99</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.96)</td>
<td>(.85)</td>
<td>(1.05)</td>
<td>(.78)</td>
<td>(1.11)</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>3.11</td>
<td>2.97</td>
<td>2.39</td>
<td>4.07</td>
<td>3.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.03)</td>
<td>(.89)</td>
<td>(1.00)</td>
<td>(.97)</td>
<td>(1.16)</td>
</tr>
<tr>
<td></td>
<td>Likeability</td>
<td>3.36</td>
<td>3.55</td>
<td>3.48</td>
<td>3.84</td>
<td>3.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.61)</td>
<td>(.56)</td>
<td>(.70)</td>
<td>(.95)</td>
<td>(.74)</td>
</tr>
<tr>
<td></td>
<td>Intelligence</td>
<td>3.59</td>
<td>3.46</td>
<td>3.41</td>
<td>4.00</td>
<td>3.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.65)</td>
<td>(.52)</td>
<td>(.78)</td>
<td>(.79)</td>
<td>(.74)</td>
</tr>
<tr>
<td></td>
<td>Safety</td>
<td>3.67</td>
<td>3.65</td>
<td>3.50</td>
<td>3.66</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.54)</td>
<td>(.48)</td>
<td>(.54)</td>
<td>(.60)</td>
<td>(.60)</td>
</tr>
<tr>
<td>Alternative</td>
<td>Humanness</td>
<td>3.27</td>
<td>2.81</td>
<td>1.98</td>
<td>4.03</td>
<td>3.05</td>
</tr>
<tr>
<td>Godspeed</td>
<td></td>
<td>(1.19)</td>
<td>(.84)</td>
<td>(1.06)</td>
<td>(.94)</td>
<td>(1.26)</td>
</tr>
<tr>
<td>Indices</td>
<td>Eeriness</td>
<td>2.60</td>
<td>2.70</td>
<td>2.50</td>
<td>2.72</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.35)</td>
<td>(.53)</td>
<td>(.67)</td>
<td>(.84)</td>
<td>(.64)</td>
</tr>
<tr>
<td></td>
<td>Attractiveness</td>
<td>3.32</td>
<td>3.27</td>
<td>2.93</td>
<td>3.97</td>
<td>3.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.58)</td>
<td>(.67)</td>
<td>(.79)</td>
<td>(.71)</td>
<td>(.79)</td>
</tr>
<tr>
<td>Team</td>
<td></td>
<td>4.24</td>
<td>4.37</td>
<td>4.12</td>
<td>5.37</td>
<td>4.54</td>
</tr>
<tr>
<td>Perception</td>
<td></td>
<td>(1.34)</td>
<td>(1.36)</td>
<td>(1.58)</td>
<td>(2.44)</td>
<td>(1.81)</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td>3.01</td>
<td>2.97</td>
<td>2.76</td>
<td>3.49</td>
<td>3.06</td>
</tr>
<tr>
<td>Identification</td>
<td></td>
<td>(.49)</td>
<td>(.68)</td>
<td>(.70)</td>
<td>(.96)</td>
<td>(.77)</td>
</tr>
</tbody>
</table>

*Note.* The Godspeed Indices and Alternative Godspeed Indices were rated on a 1-5 scale, with higher mean scores indicating stronger feelings on that subscale. Team Perception was rated on a 1-9 scale, with higher mean scores indicating a greater sense of teamwork. Group Identification was rated on a 1-5 scale, with higher mean scores indicating a stronger degree of identification with the exercise group.
Table 7.

Correlations Between Difference Scores and Emotional Responses to their Partners and Attitudes Toward their Partners by Partner Condition.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Subscale</th>
<th>VR</th>
<th>AVR</th>
<th>SG</th>
<th>PI</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Godspeed Indices</strong></td>
<td>Anthropomorphism</td>
<td>.22</td>
<td>-.05</td>
<td>.01</td>
<td>.11</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>.25</td>
<td>.04</td>
<td>-.08</td>
<td>.14</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>Likeability</td>
<td>.04</td>
<td>.03</td>
<td>-.21</td>
<td>.13</td>
<td>-.02</td>
</tr>
<tr>
<td></td>
<td>Likeability</td>
<td>.08</td>
<td>-.07</td>
<td>-.20</td>
<td>.31</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Safety</td>
<td>.02</td>
<td>.21</td>
<td>-.00</td>
<td>.06</td>
<td>.07</td>
</tr>
<tr>
<td><strong>Alternative Godspeed Indices</strong></td>
<td>Humanness</td>
<td>.14</td>
<td>.00</td>
<td>.01</td>
<td>.02</td>
<td>.20*</td>
</tr>
<tr>
<td></td>
<td>Eeriness</td>
<td>.17</td>
<td>-.03</td>
<td>.25</td>
<td>.06</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>Attractiveness</td>
<td>.22</td>
<td>-.02</td>
<td>-.01</td>
<td>.12</td>
<td>.16</td>
</tr>
<tr>
<td><strong>Team Perception</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.34</td>
<td>.18</td>
<td>.09</td>
<td>.44*</td>
<td>.26**</td>
</tr>
<tr>
<td><strong>Group Identification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.28</td>
<td>.23</td>
<td>-.11</td>
<td>.32</td>
<td>.20*</td>
</tr>
</tbody>
</table>

* = p < .05; ** = p < .01

Table 8.

Correlations Between Subscales of the Godspeed Indices.

<table>
<thead>
<tr>
<th></th>
<th>Anthropomorphism</th>
<th>Animacy</th>
<th>Likeability</th>
<th>Intelligence</th>
<th>Safety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropomorphism</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animacy</td>
<td>.90**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likeability</td>
<td>.47**</td>
<td>.57**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>.43**</td>
<td>.51**</td>
<td>.61**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>.21*</td>
<td>.24*</td>
<td>.31**</td>
<td>.19*</td>
<td>1</td>
</tr>
</tbody>
</table>

* = p < .05; ** = p < .01
Follow-up tests indicated that the significant differences for Gender were also observed on the Humanness and Attractiveness subscales. Specifically, for Humanness, male ($M = 3.27$, $SD = 1.18$) participants rated their partners as more humanlike than female ($M = 2.77$, $SD = 1.30$) participants. For Attractiveness, female ($M = 3.55$, $SD = 0.76$) participants rated their partners as more attractive than male ($M = 3.24$, $SD = 0.78$) participants.

Table 9.

*Correlations Between Subscales of the Alternative Godspeed Indices.*

<table>
<thead>
<tr>
<th></th>
<th>Humanness</th>
<th>Eeriness</th>
<th>Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humanness</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eeriness</td>
<td>.34*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Attractiveness</td>
<td>.52*</td>
<td>.37*</td>
<td>1</td>
</tr>
</tbody>
</table>

* = $p < .001$

**Attitudes Towards Partner**

Two questionnaires were used to assess attitudes towards one’s partner. Because three of the conditions did not have partners, the following analyses only took place with the four partner conditions (VR, AVR, SG, PI). For both questionnaires, averages were calculated for each scale. Refer to Table 6 for means and standard deviations for each condition for the partner and group measures and to Table 7 for correlations between these measures and difference scores.

**Team perception.** A reliability analysis was conducted on the five items of the team perception index. An adequate Cronbach’s $\alpha$ was obtained ($\alpha = .92$). A 4 (Condition) x 2 (Gender) ANOVA on the team perception index revealed a significant main effect for Condition, $F(3,104) = 3.19$, $\eta^2 = .08$, $p = .027$. SNK post-hoc tests revealed that participants in the PI condition ($M = 5.37$, $SD = 2.44$) were significantly higher in team perception than participants in
the VR \((M = 4.24, SD = 1.34)\), AVR \((M = 4.37, SD = 1.36)\), and SG \((M = 4.12, SD = 1.58)\) conditions, indicating that participants in the PI condition had stronger perceptions of being part of a team with their partners.

**Group identification.** A reliability analysis was conducted on the six items of the group identification scale. An adequate Cronbach’s \(\alpha\) was obtained \((\alpha = .90)\). A 4 (Condition) x 2 (Gender) ANOVA on group identification revealed a significant main effect for Condition, \(F(3,105) = 5.25, \eta^2 = .13, p = .002\). SNK post-hocs revealed that participants in the PI \((M = 3.49, SD = 0.96)\) condition identified significantly stronger with their exercise group than participants in the VR \((M = 3.01, SD = 0.49)\), AVR \((M = 2.97, SD = 0.68)\) and SG \((M = 2.76, SD = 0.70)\) conditions, indicating that participants in the PI condition had stronger feelings of identification with their group.
CHAPTER 5
DISCUSSION

Research to date has shown that the Köhler effect can be used in exercise settings to extend the duration at which people hold plank exercises and pedal on an exercise bike when participants are paired with supposedly real human partners (Feltz et al., 2011; Irwin et al., 2012). However, if this effect could be achieved with non-human partners, such as computer-generated or imagined partners, it would expand the range of situations in which the Köhler effect can be elicited. Exploring this question was the primary purpose of this dissertation.

Similarly, this dissertation also explored the effectiveness of different types of imagery during an exercise task. In addition to performance results (i.e., length of time holding plank exercises), secondary measures such as rating of perceived exertion, self-efficacy beliefs, enjoyment, and intention to exercise the following day were collected. The organization of this chapter mirrors the results section. Manipulation checks are discussed first, followed by discussions of the results of the main and secondary analyses (with practical implications), and then ends with an explanation of the dissertation’s strengths, limitations and future research directions.

Manipulation Checks

In the partner conditions, participants were asked two questions to gauge their understanding of the group structure. Overall, participants in these conditions understood the instructions, indicating that they were indeed working together with their partner as part of a two-person team and that the team’s score was determined by the score of the person who stopped holding the exercise first. Given that participants were told that their partner held the exercises for a longer time during the first set, it is likely participants realized the team’s score would depend on their own performance and thus felt indispensable for their team’s
performance. However, only two-thirds of participants in the PI condition reported understanding the conjunctive nature of the task, which is lower than the remaining partner conditions. Though the reason for this is unclear, it may be that participants could not quite grasp how their imagined partner would stop holding the plank exercises after they did.

In the imagery conditions, participants were asked a series of questions about their use of imagery. Overall, their responses indicated that they followed the imagery instructions closely. Participants reported creating vivid and controllable images and that they were able to keep those images in their mind throughout most of the duration of each exercise. Participants reported that these images helped them to tolerate the exertion required to perform the exercises, and they indicated a general willingness to use these imagined scenes in the future. On these measures, two significant differences emerged. First, participants in the PI condition reported perceiving more exertion than participants in the AI condition (a pattern which was mirrored in the RPE results and will be discussed later). Second, female participants started using their imagined scenes earlier than male participants. One plausible explanation for this is that participants may have started to use imagery not at the very start of the exercise, but when they started to feel some discomfort or exertion. Male participants may have tolerated the discomfort for longer until they felt the need to rely on imagery. Alternatively, female participants may have been better at following instructions and thus started using imagery closer to the beginning of each exercise compared to male participants.

The content for each of the imagery conditions was supposed to be different, and manipulation checks confirmed this was the case. Participants in the DI condition reported that their images were mostly unrelated to the task, while participants in the AI and PI conditions reported that their images were mostly related to the task. Additionally, most participants in the
PI condition reported that they stopped holding the exercise before their partner, which is in-line with the other partner conditions. Results from open-ended questions also confirmed this. Thus, it seems likely that participants understood the imagery instructions and were able to create the necessary imagery scenes.

Finally, participants in the second wave’s control group (IC) did not report using imagery at all. Instead, they used other mental strategies, such as backwards counting or nonspecific motivational thoughts (such as trying to persist). This finding would suggest that the IC group was a true no-partner, no-imagery control group; however, because this was not assessed during the first wave, it is impossible to know how representative these results are of all control group participants. Similarly, though a strategy like backwards counting itself is not imagery, the mental creation of a scoreboard or countdown clock could be considered a type of dissociative imagery. This information does add to the literature about what people may be thinking while they are exercising. Previous research suggests people will focus on either associative or dissociative thoughts while they engage in long duration exercise, such as marathon running (Morgan & Pollock, 1977). Here, the results suggest that for plank exercises, participants will engage in some goal setting (e.g., backwards counting) and motivational thoughts. These findings may generalize to other relatively short persistence tasks (e.g., performing push-ups to exhaustion). Overall, the results from the manipulation checks show that the manipulations were successful, and thus, that the results are likely to be due to the experimental manipulations.

**Main Analysis**

Before conducting the analysis on persistence scores, scores between Blocks 1 and 2 were calculated, thus creating a difference score that took into account participants’ own baseline performance. Results from an ANOVA on the difference scores were supplemented with planned
contrasts. The first planned contrast supported Hypothesis 1 – participants in the six experimental conditions (VR, AVR, SG, PI, AI, DI) improved upon their Block 1 performance during Block 2 compared to the control group (IC). To put that another way, participants in the experimental conditions had larger difference scores than participants in the control group. The second planned contrast supported Hypothesis 2 – participants teamed up with a human partner (VR) performed better compared to participants with non-human partners (AVR, SG, PI). SNK post-hoc tests generally supported these two hypotheses as well. All experimental conditions outperformed the control group with the exception of the SG condition (which was trending towards significance), generally supporting Hypothesis 1. And, participants in the VR condition did outperform participants in the non-human partner conditions, though the difference was only significantly better than participants in the AVR and SG conditions, which also generally supports Hypothesis 2. The third planned contrast, however, did not support Hypothesis 3 – there were no differences between participants in the partner conditions (VR, AVR, SG, PI) and participants in the no-partner imagery conditions (AI, DI). Post-hoc tests yielded the same pattern of results with one exception: participants in the VR condition had significantly larger difference scores compared to participants in the AI and DI conditions, but not in the PI condition. Overall, however, Hypothesis 3 was not supported.

The main analysis also revealed that female participants had significantly larger difference scores compared to male participants, a finding that is generally consistent with previous research (Feltz et al., 2012; Kerr et al., 2013; Weber & Hertel, 2007). Thus, it appears that female participants bought into the manipulations more or felt more indispensable than male participants as they performed better during the second block relative to their first block performance than did male participants. Alternatively, male participants may have worked closer
to their full capacity during Block 1, which left them more fatigued for Block 2 (though note that no significant gender differences emerged for RPE scores, a subjective rating of effort).

This dissertation replicates past research (e.g., Feltz et al., 2011, Irwin et al., 2013) on using partners to lengthen the duration at which people hold plank exercises and also extends the research to include computer-generated and imagined partners. Previous research has found that if participants do not view their partner as comparable, the Köhler effect can be weakened or even eliminated (Mussweiler, 2003; Seok, 2004). In prior exercise research with college students, working with partners who were older, heavier, or both did not moderate the Köhler effect, as performance gains were still observed (Forlenza et al., 2012). This dissertation suggests that college students will make comparisons with computer-generated and imagined partners, too, and that this will lead to a boost in motivation and performance.

These results also suggest that working with a “live” human partner leads to greater performance gains than working with computer-generated or imagined partners. Thus, the strongest motivation gain effect is achieved when working with real human partners. The results do show, however, that motivation gains can still be achieved when teamed with non-human partners, but that the improvements will not be as strong. Given that the Köhler effect was observed with non-human partners, the range of situations in which this could be applicable is expanded. For example, exergame developers should consider building partners into exergames to make the games more motivating. Many commercial exergames have a trainer to demonstrate the exercises, but few (if any) create any sense of interdependence between the trainer and player. Creating this interdependence by introducing a partner under conjunctive task demands could potentially lead players to engage with the game for longer. Increased play time ultimately benefits the game developers (as it means people are spending more time playing and enjoying
their game) and could lead to greater health or fitness benefits if played on a regular basis, a clear benefit for the player. Additionally, when exercising individually, imagining oneself exercising as part of a team with another person (or other people) could be a relatively easy way of boosting motivation, such as during the last stretch of a hard run. Indeed, participants using partner imagery persisted for longer compared to participants in the AVR and SG conditions (though this difference was not statistically significant), which suggests that allowing players to construct their own partners may lead to larger effects.

One potential reason for why the computer-generated partners did not lead to as strong a performance as the human partner is the Uncanny Valley (Mori, 1970/2005). If participants had a more negative reaction to the computer-generated partners, that could translate to a poorer performance because there would be less reason for participants to worry about letting down their partner. Interestingly, based on results from the Alternative Godspeed Indices, participants in the SG condition rated their partner as significantly less human compared to participants in the remaining partner conditions. This makes some sense, as participants working with an SG partner had the smallest performance gain (relative to individual controls). Working with a partner perceived to be much less human may have triggered participants to feel that their partner was too artificial, and thus, diminished (but not eliminated, as significant performance gains were still observed in the AVR condition and nearly for the SG condition) the social comparison mechanism of the Köhler effect.

Another potential contributing factor to why participants in the computer-generated partner conditions did not perform as well as those in the human partner condition is a face-voice mismatch. Prior research has shown that feelings of eeriness are greatest when a human voice comes from a robot, or when a synthetic voice comes from a human, versus when the agent and
voice match (i.e., human with a human voice, robot with a synthetic voice) (Mitchell, Szerszen, Lu, Schermerhorn, Scheutz, & MacDorman, 2011). In this dissertation, the AVR and SG partners used the same recording as the VR condition, and in the SG condition, the text appeared in a speech bubble next to the partner. In the SG condition, there was a clear face-voice mismatch. However, this was not reflected in eeriness ratings as there were no significant differences between conditions. One explanation for this is the study by Mitchell and colleagues (2011) used a within-subjects design, and thus their participants may have been making comparisons to previous partners during their judgments, whereas here, participants only viewed one partner at a time.

A final potential reason for why the computer-generated partners did not elicit as strong a motivation gain effect is due to the instructions used (see Appendix M). In these conditions, participants were told, “Your partner is not a real person, but is software-generated and has been programmed to be somewhat better than you at holding the plank exercises.” Additionally, participants were told, “Although Stacy/Witt is programmed to be better than you, s/he is unable to hold these exercises forever. At some point, Stacy/Witt will become ‘tired’ and stop holding the exercise.” These types of instructions are typically not given when playing games, but were given here to bring participants’ attention to the manipulation, that they would be working with a computer-generated partner. These instructions may have caused participants to prescribe robot-like characteristics to their partners (viz., being programmed to have perfect form and to be better than the participant). In turn, this may have prevented participants from prescribing human characteristics to their partners, as predicted by Media Equation (Reeves & Nass, 1996). As such, the likelihood of “teaming up” with the computer-generated partners to improve performance may have been inadvertently undermined by the instructions.
Regarding the imagery conditions, the use of imagery was shown to lead to better difference scores compared to not using imagery. Consistent with previous research (Razon et al., 2012; Razon, 2012), these results suggest that using imagery focused on oneself and the task (associative imagery) and imagery focused on something external but motivating (dissociative imagery) are both helpful for increasing exercise duration. This dissertation extends the literature by also showing that imagining oneself exercising with a superior partner is an effective imagery strategy. While there were no statistically significant differences between the three imagery conditions, participants exercising with an imagined superior partner had the largest difference scores. This potentially suggests that exercising with an imagined partner is a slightly better imagery strategy, though clearly more research is necessary to determine which strategy may be strongest. In all, these results show that using imagery is better than not having a mental strategy when the goal is to perform an exercise for as long as possible, and could be useful in applied settings to maximize performance.

Secondary Analyses

In partial support of Hypothesis 4, the repeated-measures analysis of RPE scores revealed that participants in the four partner conditions (VR, AVR, SG, PI) perceived themselves as exerting more effort during Block 2 compared to Block 1, whereas participants in the remaining conditions (AI, DI, IC) did not perceive themselves as working any harder during Block 2. Thus, it appears that exercising with a partner – whether human, computer-generated, or imagined – leads people to report greater exertion scores the second time through. This finding conflicts with some previously reported research in this area (Feltz et al., 2011; Feltz et al., 2012), though it is consistent with other research (Forlenza et al., 2012; Irwin et al., 2013; Kerr et al., 2012). Conceptually these findings make sense. Participants were asked to hold a set of plank exercises
for as long as they could, grouped with a superior partner a mere 10 min after already holding a 
set of plank exercises for as long as they could. As such, it stands to reason that the exercises 
would feel more taxing the second time through, especially when they see themselves 
outperformed continually by their partner, and these RPE scores support that notion. Participants 
in the control condition did not perceive themselves as working harder also makes sense, as their 
objective performance during Block 2 was much lower than their Block 1 performance and they 
did not have a superior partner with whom they compared themselves.

Participants in the AI and DI conditions did not perceive themselves as exerting 
significantly less effort compared to the control condition on the second block as predicted based 
on previous research (Razon et al., 2012; Razon, 2012). However, inspection of the mean RPE 
scores for these conditions shows that they stayed approximately the same (the Block 2 – Block 
1 difference for the AI condition was 0.05, while for the DI condition, the difference was -0.03). 
And consistent with the hypothesis, participants in the AI and DI conditions did not perceive any 
significant increases in exertion scores, whereas participants in all of the partner conditions did. 
Thus, these results suggest that using imagery may be an effective strategy for preventing 
increases in perception of exertion, which could be useful for workouts that involve performing 
the same set of exercises multiple times. It should also be pointed out that the small changes in 
RPE scores for the AI and DI conditions mimic the small changes in the Block 2 – Block 1 
difference scores for persistence.

Results of the self-efficacy beliefs (SE) analysis revealed that participants had 
significantly larger SE during the second measurement point (prior to Block 2) compared to the 
third measurement point (following Block 2). For the third measurement point, participants were 
asked how long they believed they could hold the exercises a third time if they had to. Thus, it
makes sense that participants had lower SE at this time point, a finding that is consistent with previous research (Irwin et al., 2013; Kerr et al., 2012). What is important about these results is that they did not differ by condition, which supports Hypothesis 5. Regardless of whether participants were the weak link in an exercise group or using imagery, SE was the same, and this finding is consistent with previous literature (Feltz et al., 2011; Forlenza et al., 2012). This means that participants’ confidence in their ability to hold plank exercises was not adversely affected by any of the interventions, supporting Hypothesis 5.

Similarly, no significant differences among conditions were observed for both enjoyment and intention to exercise for 30 min the following day, a finding that supports Hypothesis 5 and is consistent with previous literature (Feltz et al., 2011; Forlenza et al., 2012; Irwin et al., 2013). Even though participants perceived themselves to be working harder during the second block in four of the experimental conditions (VR, AVR, SG, PI), and even though participants in all of the experimental conditions held the plank exercises for longer during the second block compared to participants in the control condition, enjoyment did not suffer and it did not impact their intention to exercise the next day. This is quite encouraging, particularly for the partner conditions, as it seems participants did not enjoy the task any less despite being continuously outperformed by their partner. These findings also make the application to real world scenarios more attractive as practitioners can be reasonably certain that working with a partner will not diminish enjoyment or intention to exercise again. However, this was a single session – being continually outperformed by one’s partner across multiple sessions may eventually diminish enjoyment of the exercise, which is something future research could explore.

Participants in four of the experimental conditions (VR, AVR, SG, PI) completed the Alternative Godspeed Indices to gauge their emotional reactions to their partners (Ho &
MacDorman, 2010). As mentioned previously, participants in the SG condition rated their partner as significantly less humanlike than participants in the other partner conditions. This makes sense as the SG partner was a relatively unrefined creation, especially when compared to the VR and AVR partners. Additionally, participants in the PI condition rated their partner as significantly more humanlike and attractive compared to participants in the remaining partner conditions. This suggests that participants had stronger positive responses to partners of their own creation (PI) compared to assigned partners (VR, AVR, SG). One practical implication of this is for exergame design. Previous research suggests that allowing participants to create their own character in an exergame was associated with greater positive feelings towards the game (e.g., enjoyment, game rating) (Peng, et al., 2012). Here, the findings suggest that players should also have some control over how their partner looks.

Finally, two questionnaires were completed by participants in the partner conditions (VR, AVR, SG, PI) to assess their feelings of being on a team with their partner and feelings of how well they identified with the group. Results revealed participants in the PI condition felt that they were more of a team and identified stronger with their group than participants in the remaining partner conditions. These findings suggest that it is easier to team up with and become a group with a partner of your own creation compared to either computer-generated or prescribed partners, which could be another potential advantage to using imagined partners. Indeed, there were significant correlations between the primary dependent variable (difference scores) and responses on the team perception and group identification measures (see Table 7), particularly for participants in the PI condition. Therefore, feeling that one is a part of and identifies with a team can help explain why that leads to improvements in performance. This notion fits nicely with a meta-analysis suggesting that exercise interventions that improve group cohesion are more
effective than group exercise collectives that do not attempt to build cohesion (Burke et al., 2006). The more one identifies with the team, the more effort one will put forth to maximize the team’s performance (van Dick, Stellmacher, Wagner, Lemmer, & Tissington, 2009; though note Gockel et al., 2008). This could also be a potential explanation for why exercising with the imagined partner was effective at improving performance. Though the partner was not real, it still may have triggered feelings of togetherness and group identification in participants, thus leading them to put forth more effort and perform better.

**Strengths**

There were multiple strengths to this dissertation. First, this dissertation’s research questions and methodologies were based on a strong foundation of theory and prior research. Studies over the past 20 years have explored the Köhler effect and variables that potentially moderate this effect (Weber & Hertel, 2007). Several recent studies have focused on applying this effect in exercise contexts (Feltz et al., 2011; Forlenza et al., 2012; Kerr et al., 2013) and this dissertation built on those well-designed experiments. However, to knowledge, no prior studies had investigated the effects of explicitly non-real partners (viz., computer-generated and imagined partners) in exercise contexts. Thus, these results extend the group dynamics and exercise psychology literature to suggest that people will team up with and make social comparisons to unreal partners, in turn improving exercise duration.

Going along with this first point, this dissertation integrated theory and research from several different areas of study, including exergames, group dynamics, human-computer interactions, imagery, and social psychology. As there are numerous reasons for the society-wide decline in physical activity, many different avenues need to be explored in order to improve activity levels. Alleviating this problem will require that people work together from different
fields to create interdisciplinary solutions. This dissertation was an attempt in this direction and its results suggest that incorporating group dynamics principles with computer-generated and imagined partners can lead to statistically significant extensions in exercise duration. More research of this nature should be conducted to develop innovative methods to address the multifaceted issue of physical inactivity.

A third strength of this dissertation was its applied nature. While this project extended theory, the results are also potentially useful in applied settings. Exergame designers could use these results to form basic guidelines for how to successfully build interdependence and teamwork into a game. Exercise psychologists or practitioners could use the imagery results as a basis to help their clients reach their goals, particularly those goals centered on duration (e.g., jogging for 30 minutes). Similarly, these types of imagery may be useful in a rehabilitation setting for helping people persist through physically taxing recovery exercises.

**Limitations and Future Research Directions**

As with any research project, this dissertation has limitations. One limitation of the computer-generated partners is that their appearance was rather basic. Even though the AVR partner was created by applying a filter effect to the VR partner’s video, the result was a partner who looked almost animated, and thus not very realistic. The SG partner, meanwhile, was created using 3D modeling software and did not have much detail. The SG partner’s mouth did not move during the introduction, and the partner’s movements were extremely rigid and not very humanlike. Though significant motivation effects did occur, the effects may be even larger with more realistic and humanlike partners. Indeed, previous research has shown that people are more willing to cooperate with a realistic-looking human partner compared to an unrealistic human partner in a social dilemma game (Parise et al., 1999). Extended to exercise and the
Köhler effect, it is plausible then that people may be more willing to team up with a realistic human partner, and therefore, put forth greater amounts of effort.

Another limitation is that participants may have felt that they could never catch or surpass their computer-generated partner. When paired with a human partner, it is implicitly understood that humans cannot hold plank exercises forever and will fatigue at some point. However, this same point cannot be made for computer-generated partners. Though participants were told their computer-generated partner would eventually become tired and stop holding the exercise, they may not have believed it fully as computer-generated people do not have “real” muscles that can experience fatigue. As a result, the social comparison mechanism could have been undermined slightly, which would have led to participants quitting on the exercises earlier. Future research could examine this issue by testing repeated trials and testing over several days to determine if consistent improvement relative to one’s partner strengthens the Köhler effect. Previous research has explored this with two different tasks (i.e., front plank and a wall-sit) and found that being superior to one’s partner on one task does not modify the Köhler effect on the other task (Kerr et al., 2013). However, this study was with human partners and used two different tasks; the results may be different with computer-generated partners whose relative ability can be adjusted and a single task.

Regarding the imagery interventions, one limitation was the relatively short time amount of time participants had to practice their imagery. On average, participants practiced their imagery for roughly 2.5 min. This was necessary as the practice imagery time needed to fit into the 10 min break between Block 1 and Block 2. Research suggests that practice imagery times of less than 1 min or 10 to 15 min in length are best for improving performance (Hinshaw, 1991). Though significant improvements occurred in this dissertation, more time spent practicing
imagery (whether during a single session or over multiple practice sessions) may increase its effectiveness, and therefore, further increase performance. Conversely, the effectiveness of these types of imagery may be hard to maintain over time. Participants could become tired or bored of using the same type of imagery over and over again, thus diminishing any potential motivation and performance gains. More research is necessary with these types of imagery to explore whether these performance improvements can be sustained repeatedly.

Another limitation is the generalizability of the findings. This study was performed with young, healthy college students, who grew up using electronic and internet-based devices regularly. As such, there is ample room to explore how well these findings replicate in other populations. Teenagers and young adults are much more likely to spend time on the internet and play video games than other age groups (Lenhart et al., 2008; Rainie, 2010), so they most likely have greater experience interacting with computer-generated agents compared to middle-aged or older adults. Thus, the results from this dissertation may not hold up for people who are not as familiar or comfortable with interacting with newer technologies and computer-generated people. However, an argument could be made to the contrary. Due to their previous experiences with computer-generated agents and using internet-based technologies, young adults may be more skeptical and/or have higher standards for these types of interactions, and therefore may be less likely to team up with these types of partners. Middle-aged and older adults may find the experience as novel and engaging, and thus be more likely to buy into the intervention and its procedures, which would translate to better performance. Clearly more research is needed to uncover how other populations respond to these partners.

Another limitation regarding generalizability lies with the task used. Generally, the emphasis in exercise and health-oriented research is on aerobic exercise to improve
cardiovascular health and fitness. Plank exercises tap into muscular endurance, but are not aerobic in nature and thus do not improve cardiovascular health and fitness. Previous research has explored this topic somewhat with an exercise bike and human partners (Irwin et al., 2012). The results showed that participants working with a partner under conjunctive task demands exercised for significantly longer compared to participants who exercised alongside another person (coactive task demands) or individually. However, due to the limitations of working with human partners (e.g., scheduling, ability discrepancies), future research could attempt to replicate previous research with computer-generated partners in aerobic exercise. Doing so would provide stronger evidence for the usefulness of exergames and partner-based interventions in health and exercise settings.

There are other additional future research avenues that can be explored from this dissertation. One potential research direction is exploring the combination of conditions, which may yield stronger effects compared to one condition alone. It would be important, however, to avoid undermining the effects of each by their combination (i.e., one cannot focus on a partner and use imagery at the same time, or vice-versa). An example of how this could work would be by using associative imagery as a type of “mental repetition” between Blocks 1 and 2, which would essentially act as training and therefore would not interfere with exercising conjunctively with a partner during Block 2. Similarly, using dissociative imagery between the blocks could distract participants from the discomfort they experienced during Block 1 and make them feel “fresher” for Block 2.

Another research avenue is attempting to uncover the features that made the imagined partner motivating. While participants in that condition were instructed to team up with their imagined partner, the fact remained that this scenario was not real and that there would be no
real-world consequences to not performing well. Yet, performance still improved. One possibility that could explain this is the element of choice. The only design requirement for this imagined partner was that the partner must be the same gender as the participant. This gave participants a great deal of freedom in selecting the characteristics for their partner that would be most motivating to them. Identifying these characteristics could be very helpful for developing imagery scripts that involve partners and team performances. Additionally, these characteristics could also be used for the design of computer-generated partners in order to make them as motivating as possible. Stemming from this, a specific research project could focus on comparing the motivating effects of a prescribed partner compared to a created or chosen partner.

**Conclusion**

The results from this dissertation suggest that people will make social comparisons to and team up with computer-generated and imagined partners in exercise contexts. Doing so can elicit a Köhler motivation gain effect, which expands the range of situations in which this effect can improve motivation and ultimately performance. Importantly, these performance gains do not come at the expense of self-efficacy beliefs, enjoyment, and intention to exercise in the future, which is an encouraging sign for application to different forms of exercise and other applied contexts. Additionally, this dissertation suggests that associative and dissociative imagery may be effective methods to improve performance on exercise duration tasks, too, and could provide a relatively easy way to boost motivation and performance while exercising. Overall, group dynamics principles and imagery are two strategies that can be used separately or together to extend exercise duration.
APPENDICES
APPENDIX A

IRB APPROVAL LETTER

MICHIGAN STATE UNIVERSITY

March 1, 2012

To: Deborah L. Feltz
    130 IM Sports Circle
    Dept. of Kinesiology
    MSU

Re: IRB# 11-849 Category: EXPEDITED 2-7
    Approval Date: February 29, 2012
    Expiration Date: February 28, 2013

Title: Cyber Partners: Harnessing Group Dynamics to Boost Motivation to Exercise

The Institutional Review Board has completed their review of your project. I am pleased to advise you that your project has been approved.

The committee has found that your research project is appropriate in design, protects the rights and welfare of human subjects, and meets the requirements of MSU's Federal Wide Assurance and the Federal Guidelines (45 CFR 46 and 21 CFR Part 50). The protection of human subjects in research is a partnership between the IRB and the investigators. We look forward to working with you as we both fulfill our responsibilities.

Renewals: IRB approval is valid until the expiration date listed above. If you are continuing your project, you must submit an Application for Renewal application at least one month before expiration. If the project is completed, please submit an Application for Permanent Closure.

Revisions: The IRB must review any changes in the project, prior to initiation of the change. Please submit an Application for Revision to have your changes reviewed. If changes are made at the time of renewal, please include an Application for Revision with the renewal application.

Problems: If issues should arise during the conduct of the research, such as unanticipated problems, adverse events, or any problem that may increase the risk to the human subjects, notify the IRB office promptly. Forms are available to report these issues.

Please use the IRB number listed above on any forms submitted which relate to this project, or on any correspondence with the IRB office.

Good luck in your research. If we can be of further assistance, please contact us at 517-355-2180 or via email at IRB@msu.edu. Thank you for your cooperation.

Sincerely,

Harry McGee, MPH
IRB Chair

c: Brian Winn, Norbert Kerr, Karin Pfeiffer, Brandon Irwin, Samuel Forlenza
### The Borg Scale

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>No exertion at all</td>
</tr>
<tr>
<td>7</td>
<td>Extremely light</td>
</tr>
<tr>
<td>8</td>
<td>Very light</td>
</tr>
<tr>
<td>9</td>
<td>Light</td>
</tr>
<tr>
<td>10</td>
<td>Somewhat hard</td>
</tr>
<tr>
<td>11</td>
<td>Hard (heavy)</td>
</tr>
<tr>
<td>12</td>
<td>Very hard</td>
</tr>
<tr>
<td>13</td>
<td>Extremely hard</td>
</tr>
<tr>
<td>20</td>
<td>Maximal exertion</td>
</tr>
</tbody>
</table>
APPENDIX C

SELF-EFFICACY BELIEFS

For Time 1 (prior to Block 1) and Time 2 (prior to Block 2):

What is the number of seconds which you are completely confident that you can hold:

The FIRST exercise (front plank)? _______

The SECOND exercise (right side plank)? _______

The THIRD exercise (right one-legged plank)? _______

The FOURTH exercise (left side plank)? _______

The FIFTH exercise (left one-legged plank)? _______

For Time 3 (following Block 2):

If you were to hold these exercises one more time (after a similar break as before), what is the number of seconds which you are completely confident that you can hold:

The FIRST exercise (front plank)? _______

The SECOND exercise (right side plank)? _______

The THIRD exercise (right one-legged plank)? _______

The FOURTH exercise (left side plank)? _______

The FIFTH exercise (left one-legged plank)? _______
APPENDIX D

PHYSICAL ACTIVITY ENJOYMENT SCALE (PACES) AND INTENTION TO EXERCISE

Please rate how you currently feel about the physical activity you have been doing according to the following scales:

1 = I loved it, 7 = I hated it

1 = I felt bored, 7 = I felt interested

1 = I disliked it, 7 = I liked it

1 = I found it pleasurable, 7 = I found it unpleasurable

1 = I was very absorbed in this activity, 7 = I was not at all absorbed in this activity

1 = It was no fun at all, 7 = It was a lot of fun

1 = It was very pleasant, 7 = It was very unpleasant

1 = I felt as though I would rather be doing something else, 7 = I felt as though there was nothing else I would rather be doing

I intend to exercise tomorrow for at least 30 minutes.
-3 = Not at all true for me, 3 = Completely true for me
APPENDIX E
GODSPEED INDICES

Rate your partner according to the following scales:

**Anthropomorphism**

1 = Fake, 5 = Natural

- 1 = Machinelike, 5 = Humanlike

- 1 = Unconscious, 5 = Conscious

- 1 = Artificial, 5 = Lifelike

- 1 = Moving rigidly, 5 = Moving elegantly

**Animacy**

1 = Dead, 5 = Alive

- 1 = Stagnant, 5 = Lively

- 1 = Mechanical, 5 = Organic

- 1 = Artificial, 5 = Lifelike

- 1 = Inert, 5 = Interactive

- 1 = Apathetic, 5 = Responsive

**Likeability**

1 = Dislike, 5 = Like

- 1 = Unfriendly, 5 = Friendly

- 1 = Unkind, 5 = Kind

- 1 = Unpleasant, 5 = Pleasant

- 1 = Awful, 5 = Nice

**Perceived Intelligence**

1 = Incompetent, 5 = Competent

- 1 = Ignorant, 5 = Knowledgeable

- 1 = Irresponsible, 5 = Responsible

- 1 = Unintelligent, 5 = Intelligent

- 1 = Foolish, 5 = Sensible
Perceived Safety

1 = Anxious, 5 = Relaxed

1 2 3 4 5

1 = Agitated, 5 = Calm

1 2 3 4 5

1 = Quiescent, 5 = Surprised

1 2 3 4 5
Rate your partner according to the following scales:

**Humanness**

1 = Artificial, 5 = Natural

1 = Human-made, 5 = Human-like

1 = Without definite lifespan, 5 = With definite lifespan

1 = Inanimate, 5 = Animate

1 = Mechanical movement, 5 = Biological movement

1 = Synthetic, 5 = Real

**Eeriness**

1 = Reassuring, 5 = Eerie

1 = Numbing, 5 = Freaky

1 = Ordinary, 5 = Superordinary

1 = Bland, 5 = Uncanny

1 = Unemotional, 5 = Hair-raising

1 = Uninspiring, 5 = Spine-tingling

1 = Predictable, 5 = Thrilling

1 = Boring, 5 = Shocking

**Attractiveness**

1 = Unattractive, 5 = Attractive

1 = Repulsive, 5 = Agreeable

1 = Ugly, 5 = Beautiful

1 = Messy, 5 = Sleek

1 = Crude, 5 = Stylish
APPENDIX G

TEAM PERCEPTION INDEX

For each of the following statements, rate how much you agree or disagree with them.

1 = Strongly Disagree, 9 = Strongly Agree

I felt I was part of a team.

1 2 3 4 5 6 7 8 9

I thought of my partner as a teammate.

1 2 3 4 5 6 7 8 9

I felt I worked collaboratively with my partner.

1 2 3 4 5 6 7 8 9

I felt my partner and I worked together.

1 2 3 4 5 6 7 8 9

I felt I was working separately from my partner.

1 2 3 4 5 6 7 8 9
For each of the following statements, “exercise group” refers to yourself and the person you exercised with, your partner whom you met.

1 = Strongly Disagree, 5 = Strongly Agree

I considered this exercise group to be important.

I identified with this exercise group.

I felt strong ties with this exercise group.

I was glad to belong to this exercise group.

I saw myself as belonging to this exercise group.

I would be annoyed to say that I was a member of this exercise group.
APPENDIX I

PARTNER MANIPULATION CHECK

In which of the following conditions did you perform the last series of exercises?
1. Except for the experimenter, I performed these exercises alone.
2. I performed these exercises with another person through an internet connection.
3. I performed these exercises with two other persons through an internet connection.

How was your Total Score determined during the last series of exercises?
1. My score is the number of seconds I held each exercise.
2. My score is an average of how long I held each exercise and how long the other person held each exercise.
3. My score is the sum of my team’s score on each exercise, where the team’s score is the number of seconds each exercise was held by the first team member to quit.
4. My score is the sum of my team’s score on each exercise, where the team’s score is the number of seconds each exercise was held by the last team member to quit.
APPENDIX J

IMAGERY MANIPULATION CHECK

Briefly describe the practice scene you imagined:

______________________________________________________________________________

Please describe your partner. (PI condition only)

______________________________________________________________________________

While you and your imagined partner were simultaneously holding the plank exercises, who generally stopped holding the plank exercises first? (PI condition only)
- You
- Your imagined partner

Please indicate your answers to the questions below by selecting the number that corresponds best to your experience:

How clear/vivid was the image?
0 = Not clear/vivid, 10 = Very clear/vivid

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How much control did you feel you had over the image?
0 = No control at all, 10 = Very high control

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How much did you feel engaged (active or passive) with the image?
0 = Not at all, 10 = Very much

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How motivated were you to reduce the exertion during the imagery scene?
0 = Not at all, 10 = Very much

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

To what extent did you perceive the exertion during the imagery scene?
0 = Not at all, 10 = Very much

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
To what extent was the imagery scene easy or hard to generate?
0 = Very easy, 10 = Very hard

```
0 1 2 3 4 5 6 7 8 9 10
```

How long did you hold the imagery scene?
0 = For a very short time (a few seconds), 10 = For the entire session

```
0 1 2 3 4 5 6 7 8 9 10
```

How long did it take for the imagery scene to appear as an image?
0 = Immediately, 10 = Toward the end of the session

```
0 1 2 3 4 5 6 7 8 9 10
```

To what degree do you feel the imagery use helped you tolerate the physical effort you underwent?
0 = Not at all, 10 = Very much

```
0 1 2 3 4 5 6 7 8 9 10
```

How much do you believe you had the motivation to deal with the physical effort required by the plank exercises?
0 = None, 10 = Very much

```
0 1 2 3 4 5 6 7 8 9 10
```
APPENDIX K

INDIVIDUAL CONTROL MANIPULATION CHECK

What were you thinking about while you were holding the plank exercise?
______________________________________________________________________________

Why did you think about what you did while holding the exercises?
______________________________________________________________________________

Describe what was typically going on in your mind when you first felt fatigue.
______________________________________________________________________________
APPENDIX L

DEMOGRAPHIC QUESTIONS

Gender
- Male
- Female

Age
_____

Class
- 1st year
- 2nd year
- 3rd year
- 4th year
- 5th year
- > 5 years

Height (in feet and inches)
_____

Weight (in pounds)
_____

Race/Ethnicity
- Caucasian
- African American
- Hispanic
- Asian
- Native American
- Other
Virtually-real (VR) partner script.

Spoken to participants at the start of the 10 min break:

Ok. On the next series of exercises, we’ll be following a somewhat different procedure. On this series, you’ll be paired with another person and you will be able to see how well you do at each exercise compared to him/her. Unlike the last series, on this series you are both working towards a team score. Your team’s score will be defined by the score of the person who quits first. That is, while you are both exercising, when one person quits, the other person has to quit too and the team’s score will be the time of the person who quit first. The other person is in another lab, connected to us over the internet. S/he has already gone through the exercises one time, just like you. And right now, we’d like the two of you to meet each other. Our connection can be choppy at times, so you’ll just have to bear with it if that happens. It shouldn’t be too bad, but you might see and hear some jumps in the picture and sound. Just give me one second while I set up the connection.

Spoken to participants and their partner during the introduction:

Ok, so we’d like to allow the two of you a few moments to get to know each other by sharing just some basic information. Although we don’t have much time to spare here, it’s better than nothing. (Talking to virtual partner) So why don’t you start and just tell us a little about yourself: your name, where you’re from, your favorite sports team, if you have a pet, and what you like to do for fun. And _______ (name of participant), you can provide the same information about yourself afterwards.

Spoken by the virtual partner:

Ok, well, my name is (Witt/Stacy). I’m from right here in East Lansing. My favorite team? That’s easy – I’m a big Spartans fan. Do I have a pet? No, I don’t have a pet. But I have a fish on my screensaver if that counts. And for fun? Oh, I like to just hang out and play video games.

Spoken to participants following the partner introduction:

On the coming series of exercises, you will both begin at the same time; you will be able to see his/her displays, just as he/she will be able to see yours. You will both hold each exercise for as long as you can, saying STOP when you have to quit and dropping back down to your mat. After one of you stops, the other person has to stop. The team score is the number of seconds the first team member lasts, and both teammates receive that score. Again, our internet connection is not flawless and that you may see a delay or glitch in the image of your partner. Like before, tell me your perceived exertion after
each exercise. After you have both stopped the exercise and after a brief rest period, we’ll move on to the next exercise, just as before. You will know when the next exercise begins by watching the screen and waiting for the countdown to appear. Again, your score on the next series is equal to the team score, which is how long the first group member to quit lasts at each exercise. OK. During the series of five exercises you just completed, you held each exercise an average of ______ seconds. Your partner, Stacy/Witt, held each exercise for an average of ______ seconds. Now, before we begin, answer the questions on the screen.

Spoken to participants in the final min of the break, before starting Block 2:

OK. We’re now ready to do the next series of exercises. When you hear the countdown, focus your attention on getting into the proper position. As before, try to hold each exercise as long as you can, consistent with your own physical well-being and comfort. Again, please say STOP when you can no longer continue. Then, report your exertion by saying the number from the scale on the wall that best represents how you felt during the last exercise, and wait for the next exercise to begin. Again, when one of you stops, the other has to stop. I will let you know if Witt/Stacy stops before you, in which case you should drop down to the mat and say the number from the scale. I will let you know when you have completed the last exercise, and once you do, just sit on the mat until I give further instructions. Please watch the monitor.

Altered virtually-real (AVR) partner and software-generated (SG) partner scripts, only showing sections different from VR script, changes in italics.

Spoken to participants at the start of the 10 min break:

Ok. On the next series of exercises, we’ll be following a somewhat different procedure. On this series, you’ll be paired with a partner, and you will be able to see how well you do at each exercise compared to him/her. Unlike the last series, on this series you are both working towards a team score. Your team’s score will be defined by the score of the person who quits first. That is, while you are both exercising, when one person quits, the other has to stop. I will let you know if Witt/Stacy stops before you, in which case you should drop down to the mat and say the number from the scale. I will let you know when you have completed the last exercise, and once you do, just sit on the mat until I give further instructions. Please watch the monitor.

Spoken to participants following the partner introduction:

On the coming series of exercises, you will both begin at the same time; you will be able to see his/her displays, just as he/she will be able to see yours. You will both hold each exercise for as long as you can, saying STOP when you have to quit and dropping back down to your mat. After one of you stops, the other person has to stop. The team score is
the number of seconds the first team member lasts, and both teammates receive that score. Again, our internet connection is not flawless and that you may see a delay or glitch in the image of your partner. Like before, tell me your perceived exertion after each exercise. After you have both stopped the exercise and after a brief rest period, we’ll move on to the next exercise, just as before. You will know when the next exercise begins by watching the screen and waiting for the countdown to appear. Again, your score on the next series is equal to the team score, which is how long the first group member to quit lasts at each exercise. OK. During the series of five exercises you just completed, you held each exercise an average of _____ seconds. As I mentioned earlier, your partner, Stacy/Witt, has been programmed to be somewhat better than you at holding the plank exercises. So, had s/he done the same exercises you just completed, s/he would have held each exercise for an average of _____ seconds. Although Stacy/Witt is programmed to be better than you, s/he is unable to hold these exercises forever. At some point, Stacy/Witt will become “tired” and stop holding the exercise. Now, before we begin, answer the questions on the screen.

**Associative imagery (AI) script.**

Explanation of imagery, spoken to participants at the start of the 10 min break:

Ok. On the next series of exercises, we’ll be following a somewhat different procedure. On this next series, you are going to use imagery while performing the plank exercises. Imagery is a mental technique that involves creating your own reality using your own thoughts and mental pictures. Imagery is widely used in sport and exercise for its potential effects on enhancing motivation and performance. Imagery is about training your mind to create mental patterns that help both the brain and the motor system to carry out the activity successfully.
There are some critical aspects to consider while using imagery:
1. Vividness: The brighter, clearer, and more distinct your images are, the more effective imagery will be, much like a true life like experience.
2. Control: You should be able to control the images created in your mind. You should be capable of exercising direct influence on your mental pictures, meaning that you can check, manage, and modify them as needed.
3. Using multiple senses: You should try to perceive your mental pictures with as many senses as possible, including sight, feel, hearing, taste, smell, and touch. The more comprehensive and inclusive of the senses in your images, the better your body can understand what it has to do.

Explanation of associative imagery:

On the next set of exercises, you are going to use associative imagery, which means imagery that is focused on the task, or in this case, plank exercises. So when you begin holding each exercise, you will also start using imagery and continue to use imagery until you stop holding the exercise. For this type of imagery, you should draw motivation from imagining yourself holding the exercises successfully. For example, you might imagine yourself successfully enduring at the task, your body becoming stronger as you are
holding each exercise, or seeing and feeling yourself as being focused, confident, and in control. Keep the focus of your imagery on yourself holding each plank. Remember that your images should be vivid, under your control, and utilize as many senses as possible. You will hold each exercise for as long as you can while using imagery, saying STOP when you have to quit and dropping back down to your mat. Like before, tell me your perceived exertion after each exercise, and after a brief rest period, we’ll move on to the next one. You will know when the next exercise begins by waiting for the countdown to appear.

Spoken to participants after the imagery explanations:

OK. During the series of five exercises you just completed, you held each exercise an average of _____ seconds. Now, before we begin, answer the questions on the screen.

Explanation of practice imagery:

Ok, now we will now take a short rest period, during which time you will be able to practice the imagery. Move back to the mat and have a seat. Take the next couple of minutes to practice your imagery – imagine yourself becoming stronger and successfully enduring while holding the various plank exercises. However, while you are practicing your imagery, please verbalize it. That is, say out loud what you are imagining. We’ll do that just for the practice session, which will last for a few minutes – keeping using and refining your images until we are ready to begin again, and I will let you know when that is. Try to imagine all five plank exercises, but take your time going through them. You may begin.

Spoken to participants in the final min of the break, before starting Block 2:

OK, you can stop practicing your imagery as we’re now ready to do the next series of exercises. When you hear the countdown, focus your attention on getting into the proper position. As before, try to hold each exercise as long as you can, consistent with your own physical well-being and comfort. Again, please say STOP when you can no longer continue. Then, report your exertion by saying the number from the scale on the wall that best represents how you felt during the last exercise, and wait for the next exercise to begin. Remember, as soon as you get into position, start using imagery and continue to use imagery throughout the duration of each plank. I will let you know when you have completed the last exercise.

Dissociative imagery (DI) script, only showing sections different from AI script, changes in italics.

Explanation of dissociative imagery:

On the next set of exercises, you are going to use dissociative imagery, which means imagery that is focused on something external from the task, or in this case, plank exercises. So when you begin holding each exercise, you will also start using imagery
and continue to use imagery until you stop holding the exercise. For this type of imagery, you should draw motivation from imagining yourself doing things that are not related to holding these plank exercises. For example, you might imagine yourself successfully completing a difficult project or assignment, achieving a lifelong goal, or seeing and feeling yourself as being focused, confident, and in control during a challenging situation. Keep the focus of your imagery on yourself either accomplishing or doing something that would be motivational.

Explanation of practice imagery:

Ok, we will now take a short rest period, during which time you will be able to practice the imagery. Move back to the mat and have a seat. Take the next couple of minutes to practice your imagery – imagine yourself doing something motivating but unrelated to the task while you are holding the various plank exercises. However, while you are practicing your imagery, please verbalize it. That is, say out loud what you are imagining. We’ll do that just for the practice session, which will last for a few minutes – keeping using and refining your images until we are ready to begin again, and I will let you know when that is. You may begin.

Partner imagery (PI) script.

Spoken to participants at the start of the 10 min break:

Ok. On the next series of exercises, we’ll be following a somewhat different procedure. On this next series, you’ll be paired with an imaginary partner, and you will be able to see how well you do at each exercise compared to him/her. Unlike the last series, on this series you are both working towards a team score. Your team’s score will be defined by the score of the person who quits first. That is, while you are both exercising, when one person quits, the other person has to quit too and the team’s score will be the time of the person who quit first. However, your partner is not a real person, but will be imagined. On this next series, you are going to use imagery while performing the plank exercises. Imagery is a mental technique that involves creating your own reality using your own thoughts and mental pictures. Imagery is widely used in sport and exercise for its potential effects on enhancing motivation and performance. Imagery is about training your mind to create mental patterns that help both the brain and the motor system to carry out the activity successfully.

There are some critical aspects to consider while using imagery:

1. Vividness: The brighter, clearer, and more distinct your images are, the more effective imagery will be, much like a true life like experience.
2. Control: You should be able to control the images created in your mind. You should be capable of exercising direct influence on your mental pictures, meaning that you can check, manage, and modify them as needed.
3. Using multiple senses: You should try to perceive your mental pictures with as many senses as possible, including sight, feel, hearing, taste, smell, and touch. The more comprehensive and inclusive of the senses in your images, the better your body can understand what it has to do.

136
Spoken to participants for the imagined partner introduction:

So now, we’d like for you and your imaginary partner to take a few moments to get to know each other by sharing some basic information: your name, where you’re from, your favorite sports team, if you have a pet, and what you like to do for fun. Your partner should be someone you have not met before, of the same sex, and wearing workout clothes. Ok, now imagine yourself meeting your partner for the first time, where each of you share that basic information in about 30 seconds. Take a few seconds to imagine your partner and then introduce yourself to him/her. Please let me know when you have finished.

Explanation of partner imagery:

Ok, great. Stay seated at the computer and we’ll continue. On the coming series of exercises, when you begin holding each exercise, also imagine your partner holding that same exercise. And you should imagine your partner holding each plank successfully, that is, they should persist holding each exercise with good form. Remember that your images should be vivid, under your control, and utilize as many senses as possible. You will hold each exercise for as long as you can while using imagery, saying STOP when you have to quit and dropping back down to your mat. Like before, tell me your perceived exertion after each exercise, and after a brief rest period, we’ll move on to the next one. You will know when the next exercise begins by waiting for the countdown to appear. The team score is the number of seconds the first team member lasts, and both teammates receive that score.

OK. During the series of five exercises you just completed, you held each exercise an average of _____ seconds. As I mentioned earlier, your imagined partner should be somewhat better than you at holding the plank exercises, such that had s/he done the same exercises you just completed, s/he would have held each exercise for an average of about _____ seconds. Although you are imagining your partner to be better than you, s/he should be unable to hold these exercises forever. At some point, your imagined partner can become “tired” and stop holding the exercise. Now, before we begin, answer the questions on the screen.

Explanation of practice imagery:

Ok, we will now take a short rest period, during which time you will be able to practice the imagery. Move back to the mat and have a seat. Take the next couple of minutes to practice your imagery – imagine your partner holding each plank exercise in a similar setting to the one you are in right now. Again, your partner should be the same person you just met, who is of the same sex, somewhat fitter than you, and wearing workout clothes. However, while you are practicing your imagery, please verbalize it. That is, say out loud what you are imagining. We’ll do that just for the practice session, which will last for a few minutes – keeping using and refining your images until we are ready to begin again, and I will let you know when that is. Try to imagine your partner performing all five plank exercises, but take your time going through them. You may begin.
Spoken to participants in the final min of the break, before starting Block 2:

**OK, you can stop practicing your imagery as we’re now ready to do the next series of exercises. When you hear the countdown, focus your attention on getting into the proper position. As before, try to hold each exercise as long as you can, consistent with your own physical well-being and comfort. Again, please say STOP when you can no longer continue. Then, report your exertion by saying the number from the scale on the wall that best represents how you felt during the last exercise, and wait for the next exercise to begin. Remember, as soon as you get into position, start imagining that your partner is also holding the same exercise as you, and keep this image in your mind throughout the duration of each plank. Again, when one of you stops, the other has to stop. I will let you know when you have completed the last exercise.**

**Individual control (IC) script.**

Spoken to participants early in the 10 min break:

**OK. During the five exercises you just completed, you held each exercise an average of _____ seconds. In a few minutes, you will be performing a series of the same exercises again. As before, you are going to hold each exercise as long as you can, consistent with your own physical well-being and comfort. Now, please have a seat over here in front of the computer. Before we move on to the rest period, answer the questions on the screen.**

Spoken to participants in the final min of the break, before starting Block 2:

**OK. We’re now ready to do the next series of exercises. When you hear the countdown, focus your attention on getting into the proper position. As before, try to hold each exercise as long as you can, consistent with your own physical well-being and comfort. Again, please say STOP when you can no longer continue. Then, report your exertion by saying the number from the scale on the wall that best represents how you felt during the last exercise, and wait for the next exercise to begin. I will let you know when you have completed the last exercise.**
Main Analysis post-hoc Test Summaries.

Student-Newman-Keuls (SNK) post-hoc tests were used to analyze any between-group differences by condition in the main analysis. The SNK test is a fairly moderate test, but for the sake of completeness, results from more conservative post-hoc tests (viz. the Tukey and Scheffe tests) are also reported here. Recall that the overall pattern of results for the SNK post-hoc tests was as follows: VR > AI, DI, AVR, SG, IC; PI, AI, DI, AVR > IC.

Tukey’s post-hoc test revealed that participants in the VR condition had significantly larger difference scores compared to participants in the AVR, SG, and IC conditions. Additionally, participants in the PI, AI, and DI conditions had significantly larger difference scores compared to participants in the IC condition. Thus, the overall pattern for Tukey’s post-hoc test was as follows: VR > AVR, SG, IC; PI, AI, DI > IC.

Scheffe’s post-hoc test revealed that participants in the VR condition had significantly larger difference scores compared to participants in the SG and IC conditions, while participants in the PI condition also significantly outperformed participants in the IC condition. Thus, the overall pattern for Scheffe’s post-hoc test was as follows: VR > SG, IC; PI > IC. A summary of how these post-hoc tests compared to each other can be found in Table 10.

Alternative Main Analyses

The results in this dissertation were analyzed primarily through the use of difference scores. There are two alternative approaches to assessing change that are less susceptible to some of the shortcomings of difference scores (Edwards, 2001; Knapp & Schafer, 2009; Senn, 2006).
A summary of how these alternative analyses compared to the main analysis can be found in Table 10.

The first approach is to use a regression to obtain residual scores, and to then perform an ANOVA on the residuals. As such, Block 2 scores were regressed onto Block 1 scores, and the unstandardized residual scores were saved. (The model summary revealed an adjusted $R^2$ of .78). Following this, a one-way ANOVA of Condition on the residual scores was performed. Results indicated a significant main effect for Condition, $F(6,218) = 8.22, p < .001$. SNK post-hoc tests revealed that participants in the VR condition ($M = 41.10, SD = 52.99$) performed significantly better compared to participants in all of the remaining conditions: PI ($M = 14.55, SD = 32.19$), AI ($M = 4.45, SD = 43.33$), DI ($M = -1.17, SD = 42.92$), AVR ($M = -3.88, SD = 35.01$), SG ($M = -15.08, SD = 39.24$), and IC ($M = -28.94, SD = 52.99$). (Note that adjusted means were reported here for this analysis). Further, participants in the PI and AI conditions outperformed participants in the IC condition. Thus, the overall pattern was: VR > PI, AI, DI, AVR, SG, IC; PI, AI > IC. A similar pattern was observed for Tukey’s post-hoc test, with the only exception being that there were no significant differences between the VR and PI conditions. Finally, Scheffe’s post-hoc test revealed that participants in the VR condition performed significantly better than participants in the DI, AVR, SG, and IC conditions, while participants in the PI condition also performed significantly better than those in the IC condition. The adjusted mean scores are presented below in Figure IV.

The second approach, which was more conservative, was to use the Block 1 scores as a covariate in the analysis of Block 2 scores. As such, a 7 (Condition) x 2 (Gender) ANCOVA was conducted on the sum of the Block 2 scores with the sum of the Block 1 scores as a covariate. Unsurprisingly, Block 1 scores predicted Block 2 scores, $F(1,204) = 785.53, p < .001, \eta^2 = .79$. 

Results indicated a significant main effect for Condition, $F(6,204) = 8.10, p < .001, \eta^2 = .19$. Follow-up pairwise comparisons with Bonferroni’s correction factor applied revealed that the VR ($M = 323.14, SD = 101.24$) condition was significantly larger than the AI ($M = 281.33, SD = 112.68$), DI ($M = 253.73, SD = 62.36$), AVR ($M = 265.93, SD = 93.57$), SG ($M = 245.70, SD = 95.68$), and IC ($M = 268.20, SD = 138.42$) conditions (all $p$ values < .04). Additionally, scores from the PI ($M = 274.83, SD = 86.15$) and AI conditions were significantly larger than scores in the control group. Thus, the overall pattern of results was: VR > AI, DI, AVR, SG, IC; PI, AI > IC.

*Figure IV.* Adjusted means for each condition of the unstandardized residuals (from the Block 2 on Block 1 regression) with 95% confidence intervals.
Table 10.

Summary of the Main Analysis, Alternative Follow-up Tests, and Alternative Analyses.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Comparison</th>
<th>post-hoc Test Performed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANOVA on Difference Scores</strong></td>
<td>VR Significantly Better Than...</td>
<td>Student-Newman-Keuls</td>
</tr>
<tr>
<td></td>
<td>AI, DI, AVR, SG, IC</td>
<td>Tukey</td>
</tr>
<tr>
<td></td>
<td>IC Significantly Less Than...</td>
<td>Scheffe</td>
</tr>
<tr>
<td></td>
<td>VR, PI, AI, DI, AVR, SG, IC</td>
<td>Bonferroni</td>
</tr>
<tr>
<td><strong>ANOVA on Regression Residuals</strong></td>
<td>VR Significantly Better Than...</td>
<td>AI, DI, AVR, SG, IC</td>
</tr>
<tr>
<td></td>
<td>PI, AI, DI, AVR, SG, IC</td>
<td>Tukey</td>
</tr>
<tr>
<td></td>
<td>IC Significantly Less Than...</td>
<td>Scheffe</td>
</tr>
<tr>
<td></td>
<td>VR, PI, AI, DI, AVR, SG, IC</td>
<td>Bonferroni</td>
</tr>
<tr>
<td><strong>ANCOVA on Block 2 Scores with Block 1 as Covariate</strong></td>
<td>VR Significantly Better Than...</td>
<td>AI, DI, AVR, SG, IC</td>
</tr>
<tr>
<td></td>
<td>IC Significantly Less Than...</td>
<td>VR, PI, AI</td>
</tr>
</tbody>
</table>

*Note.* All significant differences for Student-Newman-Keuls (SNK) test based on homogenous subsets. All significant differences for Tukey, Scheffe, and Bonferroni tests occurred at \( p < .05 \).

Results from these supplemental analyses revealed similar patterns as the primary analysis (ANOVA on difference scores). Overall, participants in the VR condition generally performed significantly better than participants in the AI, DI, AVR, SG, and IC conditions. Additionally, participants in the IC condition generally performed significantly worse than participants in the VR, PI, and AI conditions. Previous studies in this line of research have found consistent patterns between ANOVAs on difference scores and ANCOVAs on Block 2 scores.
with Block 1 scores as a covariate (e.g., Feltz et al., 2011; Forlenza et al., 2012). Here, the ANCOVA results generally supported the pattern of results from the difference score analysis, despite some differences. One potential explanation for the differences here was that SNK tests tend to become less conservative when more than three groups are under comparison. This may have resulted in more comparisons being noted as statistically significant, viz., the IC vs. DI and IC vs. AVR comparisons, which did not occur in the ANCOVA results. Overall, however, the patterns of results were similar.
REFERENCES
REFERENCES


Hüffmeier, J., & Hertel, G. (2011). When the whole is more than the sum of its parts: Group motivation gains in the wild. *Journal of Experimental Social Psychology, 47*(2), 455–459. doi:10.1016/j.jesp.2010.12.004


