Investigating the Requirements and Establishing an Exercise Habit in Gym Members

By

Navin Kaushal
MSc (Dist.), Memorial University, 2011
BHSc (Joint Honours), Western University, 2009

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In the School of Exercise Science, Physical and Health Education

Navin Kaushal, 2016
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author
Investigating the Requirements and Establishing an Exercise Habit in Gym Members

By

Navin Kaushal
MSc (Dist.), Memorial University, 2011
BHS (Joint Honours), Western University, 2009

Supervisory Committee:

Dr. Ryan E. Rhodes, PhD (School of Exercise Science, Physical & Health Education, University of Victoria)
Supervisor

Dr. John T. Meldrum, PhD (School of Exercise Science, Physical & Health Education, University of Victoria)
Department Member

Dr. John C. Spence, PhD (Faculty of Physical Education and Recreation, University of Alberta)
Outside Member
Abstract

Supervisory Committee
Dr. Ryan E. Rhodes, PhD (School of Exercise Science, Physical & Health Education, University of Victoria)
Supervisor
Dr. John T. Meldrum, PhD (School of Exercise Science, Physical & Health Education, University of Victoria)
Department Member
Dr. John C. Spence, PhD (Faculty of Physical Education and Recreation, University of Alberta)
Outside Member
Dr. John C. Spence, PhD (Faculty of Physical Education and Recreation, University of Alberta)

Background: Exercise behaviour has largely been studied via reflective social cognitive approaches over the last thirty years. Emerging findings have shown habit to demonstrate predictive validity with physical activity. Habit represents an automatic behaviour that becomes developed from repeated stimulus-response bonds (cued and repetitive action) overtime. Despite the correlation with PA, the literature lacks research in understanding habit formation in new exercisers and experimental evidence of this construct. Hence, the purpose of this dissertation was to: i) understand the behavioural and psychological requirements of habit formation in new gym members, ii) investigate how regular gym members maintain their exercise habit, and iii) incorporate these findings to design a randomized-controlled trial (RCT) to test the effectiveness of an exercise habit building workshop in new gym members. In particular, the RCT sought to test if the habit group would develop greater exercise improvement over a control condition and another intervention group that employed a variety-based approach. Methods: Participants for all three studies were healthy adults (18-65) who were recruited from local gym and recreation centres in Victoria, BC. Studies I and III included only new gym members who were not meeting the Canadian Physical Activity guidelines upon recruitment while study II were a sample of gym members who have been exercising for at least one year. The first two studies were prospective, observational designs (twelve and six weeks respectively) while the third was a CONSORT based experimental study. Results: The first study found that exercising for at least four bouts per week for six weeks was the minimum requirement to establish an exercise habit. Trajectory change analysis revealed habit and intention to be parallel predictors of exercise in the trajectory analysis while consistency of practice revealed to be the best predictor. The second study highlighted the distinction between the preparatory and performance phases of exercise and further found intention and preparatory habit to be responsible for behaviour change across time. This study also found consistency to be the strongest predictor for habit formation. The intervention found the habit group to increase in exercise time compared to the control (p<.05, d=.40) and variety (p<.05, d=.36) groups. Mediation analysis found habit to partially mediate between group and behaviour. Contextual predictors revealed cues and consistency to mediate habit formation and group type. Conclusions: This dissertation provided significant novel contributions to the literature which included: i) calculating the behavioural and psychological
requirements for establishing an exercise habit, ii) distinguishing two behavioural phases of exercise and iii) conducting the first exercise habit-based RCT. These findings demonstrate the effectiveness of the proposed habit-based worksheet which could be helpful for trainers and new gym members in facilitating an exercise habit.
Table of Contents

Supervisory Committee .................................................................................................................. ii
Abstract ........................................................................................................................................ iii
Table of Contents .......................................................................................................................... v
List of Tables ................................................................................................................................ viii
List of Figures ................................................................................................................................. ix
Acknowledgement .......................................................................................................................... x
Chapter 1: General Introduction ................................................................................................. 1
  Background .................................................................................................................................. 1
  A Brief History of Habit ............................................................................................................... 5
  Dissertation Objective .................................................................................................................. 7
Chapter 2: Exercise Habit Formation in New Gym Members- A Longitudinal Study ..................... 9
  Abstract ..................................................................................................................................... 10
  Introduction ................................................................................................................................. 11
  Method ....................................................................................................................................... 16
  Analysis Plan ............................................................................................................................... 20
  Results ....................................................................................................................................... 23
  Discussion ................................................................................................................................... 28
  Tables and Figures ....................................................................................................................... 33
Chapter 3: The Role of Habit in Different Exercise Phases ............................................................ 37
  Abstract ..................................................................................................................................... 39
  Introduction ................................................................................................................................. 40
  Methods ...................................................................................................................................... 44
  Results ....................................................................................................................................... 49
  Discussion ................................................................................................................................... 51
  References ................................................................................................................................... 55
  Tables and Figures ....................................................................................................................... 59
Chapter 4: Establishing an Exercise Habit: A Randomized-Controlled Trial .................................. 64
  Abstract ..................................................................................................................................... 65
  Introduction ................................................................................................................................. 66
  Methods ...................................................................................................................................... 70
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 1</td>
<td>Habit Models in Physical Activity: A Systematic Review</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Context</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>Evidence Acquisition</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>Evidence Synthesis</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>Psychology Models</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>Economic Models</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>156</td>
</tr>
<tr>
<td>Appendix 2</td>
<td>Research Methods of Measuring Physical Activity Habit: A Systematic Review</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>Background</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>190</td>
</tr>
<tr>
<td>Appendix 3</td>
<td>The Home Physical Environment and its Relationship with Physical Activity and Sedentary Behavior: A Systematic Review</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>Context</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
<td>206</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Descriptive Data ........................................................................................................................................ 36
Table 2. Bivariate Coorelations of Habit Antecedents with MVPA and Habit ....................................................... 37
Table 3. Baseline and Trajectory Analysis: Antecedents as Predictors of Habit ....................................................... 38
Table 4. Descriptive Data ........................................................................................................................................ 62
Table 5. Bivariate Correlations of Habit antecedents with Exercise and Habit ....................................................... 63
Table 6. Descriptive Data ........................................................................................................................................ 102
Table 7. Habit Model: Bivariate Correlations ......................................................................................................... 104
Table 8. Primary Outcome: Behavior Change between Habit, Variety and Control Groups at 4 Weeks ....... 106
Table 9. Primary Outcome: Behavior Change between Habit, Variety and Control Groups at 8 Weeks .......... 107
Table 10. Habit Model- The Change of Habit and its antecedents between Habit, Variety and Control Groups at Week 4 ........................................................................................................................................ 108
Table 11. Habit Model- The Change of Habit and its antecedents between Habit, Variety and Control Groups at Week 8 ........................................................................................................................................ 109
Table 12. Mediation Models of Behavior, Habit, and Action Control Models at Week 4 ........................................ 110
Table 13. Mediation Models of Behavior, Habit, and Action Control Models at Week 8 ........................................ 111
Table 14. Psychology Habit Models ......................................................................................................................... 159
Table 15. Economic Models ....................................................................................................................................... 161
Table 16. Data of Extracted Studies ........................................................................................................................ 195
Table 17. Description of Instruments that measured PA Habits ................................................................................ 198
Table 18. Data Extraction of Experimental Studies: Participant and Study Characteristics ................................ 225
Table 19. Data Extraction of Experimental Studies: Instruments and Analysis ........................................................ 230
Table 20. Data Extraction of Observational Studies: Participant and Study Characteristics ................................ 234
Table 21. Data Extraction of Observational Studies: Instruments and Analysis ........................................................ 238
Table 22. Evaluation of Experimental Studies ......................................................................................................... 245
Table 23. Evaluation of Observational Studies ........................................................................................................ 251
List of Figures

Figure 1. Habit Scores between high and Low Frequency Groups ........................................ 39
Figure 2. Dual Process Model .......................................................................................... 64
Figure 3. Habit Preparation Model ............................................................................... 65
Figure 4. Habit Preparation Model ............................................................................... 66
Figure 5. Randomization Process .................................................................................. 105
Figure 6. Mediation Model: Habit mediating Group and Accelerometry MVPA (Baseline-Week 8) .................................................................................................................. 112
Figure 7. Mediation Model: Habit mediating Group and MVPA (Baseline-Week 4) ........ 113
Figure 8. Mediation Model: Habit mediating Group and MVPA (Baseline-Week 8) ...... 114
Figure 9. Mediation Model: Consistency and Cues Mediating Group and MVPA (Baseline-Week 4) .................................................................................................................. 115
Figure 10. Mediation Model: Consistency and Cues Mediating Group and MVPA (Baseline-Week 8) .................................................................................................................. 116
Figure 11. Flow Diagram for the literature search............................................................. 163
Figure 12. Triandis (1977) Theory of Interpersonal Behaviour ........................................ 165
Figure 13. Bargh's (1994) Four Horsemen Model .......................................................... 166
Figure 14. Verplanken et al., (1997) Habit Model ................................................................ 167
Figure 15. Aarts et al., (1997) Habit Model ...................................................................... 168
Figure 16. Ouellette & Wood's (1998) Behavior Prediction Model .................................... 169
Figure 17. Grove & Zillich's (2003) Habit Requirement Model ........................................ 170
Figure 18. Anshel & Kang's (2007) Disconnected Values Model ...................................... 171
Figure 19. Lally & Gardner's (2011) Four Antecedent Habit Model ................................. 172
Figure 20. Flow Diagram for the literature Search ............................................................. 259
Figure 21. The Dual Process Approach ......................................................................... 269
I would like to thank my supervisor Dr. Ryan Rhodes for giving me the opportunity and flexibility to pursue my research questions. His guidance and feedback through theory based research has helped me develop a variety of methodological and statistical skills. I enjoyed our thought stimulating chats which have led to several novel research ideas. I would like to thank my committee members Dr. Meldrum and Dr. Spence for their valuable feedback on my dissertation. Additionally I appreciate Dr. Meldrum’s insight and skills that have helped me along the way. I would like to thank Dr. Spence and Dr. Naylor for lending me their accelerometers to collect the data for my final study. I also appreciate the help of Ms. Rebecca Zammit for promptly addressing my questions over the past four years. I appreciate the time of every participant who volunteered in my research. I would finally like to thank my family for their continued support throughout this journey.
Chapter 1: General Introduction

Background

The mortality rate from chronic diseases across the world continues to escalate yearly despite media awareness and advancements in medicine (Manchester, 2009). For instance, lung cancers caused 1.6 million (2.9%) deaths in 2012, up from 1.2 million (2.2%) deaths in 2000. Similarly, diabetes resulted in 1.5 million (2.7%) deaths in 2012, an increase from 1.0 million (2.0%) deaths in 2000. Heart disease and stroke continue to remain the leading cause of mortality with approximately 7.4 million and 6.7 million deaths respectively in 2012 (WHO, 2015). In addition, a noteworthy disturbing trend is the prevalence of depression which affects more than 350 million individuals and is now the leading cause of disability worldwide (Manchester, 2009).

Perhaps even more alarming is the increased prevalence of diseases in developed nations. The Centres for Disease Control and Prevention (CDC) found that total number of deaths in the United States are from preventable illnesses which include: heart disease and stroke (25%), cancer (20%) and diabetes (15%) (CDA, 2014). Similarly in Canada, about 2 in 5 Canadians will develop cancer in their lifetimes (CCS, 2014) and nine out of ten individuals over the age of 20 years have at least one risk factor for heart disease or stroke (H&SF, 2014). However the risks of these diseases can be significantly reduced by preventive measures which include a healthy diet, living a smoke free lifestyle, and regular physical activity (PHAC, 2014). Of these behaviour changes, arguably the one which has the most profound effect across various diseases and illnesses is physical activity (Warburton, et al., 2007).
Incorporating physical activity (PA) regularly is a preventive measure for more than 25 chronic diseases and illnesses which includes stroke, heart diseases, diabetes and various cancers and depressive symptoms (Warburton et al., 2007). It’s been recommended that adults should include at least 150 minutes of moderate-to-vigorous intensity physical activity (MVPA) per week to maximize the health benefits (Garber et al., 2011; Warburton et al., 2007). Despite these convincing findings, the majority of Canadian adult population struggles to achieve these requirements. Although self-reported data found that 53% of Canadian adults are regularly active, objective measurement has revealed that an alarming 95% of are not meeting the moderate-to-vigorous physical activity (MVPA) guidelines to reap the health benefits (Colley et al., 2011).

Although the ideal goal is for individuals to incorporate 150 minutes of MVPA, this term loosely refers to any activity at that intensity, such as running up the stairs in an office building. However, a more particular categorization of MVPA is exercise which is defined as “a subcategory of physical activity that is planned, structured, repetitive, and purposeful in the sense that the improvement or maintenance of one or more components of physical fitness is the objective” (WHO, 2014). The difficulty of successfully maintaining a regular exercise routine can be attributed to the distinctive characteristics when compared to other health practices such as brushing your teeth, flossing, and wearing a seatbelt (Rhodes & Nigg, 2011). Exercise is a complex health behaviour that requires an individual to remove him/herself from a stable physiological state (Ekkekakis, Hall, & Petruzzello, 2008), create a protected time slot (Rhodes & De Bruijn, 2010) and perform at least two distinctive behaviour phases which include preparation and performance aspects (Rhodes & De Bruijn, 2010; Verplanken & Melkevik, 2008). In addition, exercise is predicted through a large number of correlates which can also be
viewed as hurdles for maintaining behavioural adherence (Bauman et al., 2012). Overall, the challenge of understanding behavioural maintenance can be reflected from ongoing research over the past thirty years (Rhodes & Nasuti, 2011).

In an attempt to predict exercise, various conscious regulatory theories have been applied. Essentially, these theories propose that behaviour can be predicted by reflective motivational processes which include: the Self-Determination Theory (E. Deci & Ryan, 2002), Protection Motivation Theory (W. Rogers, 1974), Health Belief Model (Rosenstock, 1974), Theory of Reasoned Action (Fishbein, 1975), Social Cognitive Theory (Bandura, 1986) and the Theory of Planned Behaviour (Ajzen, 1991). For instance, the Theory of Planned Behaviour (Ajzen, 1991) suggests that three constructs (attitude, subjective norm, and perceived behavioural control) predict intention, which is the proximal predictor of behaviour. Attitude has been defined as an individual’s overall evaluation for the behaviour; subjective norm measures the individual’s perceived social influence (e.g., family, friends, physician, etc.) regarding the behaviour; and perceived behavioural control represents an individual’s perceived belief of performing the behaviour (Ajzen, 1991). Most other social cognitive theories follow a similar functionality of using reasoned action approaches to predict behaviour (Hagger, 2010; Head & Noar, 2014; Linke, Robinson, & Pekmezi, 2014; Rhodes & Nasuti, 2011). These theories assume that: i) behaviour can be entirely predicted by conscious processes and ii) behaviour is entirely volitional. These models have contributed to identifying key correlates to PA such as intention, affective attitude, and perceived behavioural control, (Rhodes & De Bruijn, 2013).

Although these theories helped identify some key correlates to PA, these models were not specifically designed to predict PA, rather they were intended to predict all human behaviour (Rhodes, 2014a). It is important to note that some of these theories place intention as the most
proximal conscious determinant of behaviour. In line with this theorizing, intention has demonstrated a reliable correlation with PA that acts as a central mediator between other conscious motives and behaviour (Armitage & Conner, 2001; Hagger, Chatzisarantis, & Biddle, 2002; McEachan, Conner, Taylor, & Lawton, 2011). Still, the relationship suggests that intention accounts for approximately 23% of the variance in PA and 5% of the variance in PA change (McEachan et al., 2011). This has led theorists to reevaluate the validity of these models for understanding PA.

The recent emergence of several reviews that outline various shortcomings of reasoned action methods, combined with emerging proponents of alternative frameworks, suggest that a movement beyond reasoned action approaches could be an insightful approach (Ekkekakis, Hargreaves, & Parfitt, 2013; Rhodes, 2014a, 2014b; Rhodes & Nigg, 2011; Sheeran, Gollwitzer, & Bargh, 2013; Sniehotta, Presseau, & Araújo-Soares, 2014). Following this rationale, one direction to consider are models that also incorporate non-conscious processes (Sheeran et al., 2013). Although non conscious processes are not clearly understood, several leading theorists agree that this process possesses characteristics of being automatic, rapid, and high capacity (Bargh, 2011; Dijksterhuis & Nordgren, 2006; Evans & Stanovich, 2013; Schneider & Shiffrin, 1977). Perhaps the most prominent theorizing to propose the functionality of unconscious system is the dual process approach (see review in appendix). The dual process approach proposes that behaviour is the result of both conscious and unconscious systems. With regards to predicting PA, habit is an unconscious construct which could play a significant role (Sheeran et al., 2013; Verplanken & Melkevik, 2008).
A Brief History of Habit

Various models (see review paper on models in appendix) and assessment methods (see review paper on methods in appendix) have been used to decipher the functionality of habit. Early theorists conceptualized habit as learned sequences of acts that become automatic responses to certain situations which could assist in achieving certain goals such as “morning routine”, “going to work”, or “doing dishes” (Hull, 1943; James, 1890; Triandis, 1977, 1980). Though habit is a psychological construct, it develops based on certain behavioural requirements such as repeating the behaviour in the same context (Verplanken, 2006). One of the earliest theorists proposed that the number of repeated pairings between a situation (e.g. travel location) and response (e.g. travel mode) is positively correlated with the strength of that association or habit (Hull, 1943). Habits initially develop from a particular goal or purpose and hence, there is a goal-directed component in habit which can be consciously instigated (Bargh, 1989). This theorizing was further refined by Verplanken and colleagues (1997) in which they proposed that the process of a goal activation followed by automatic behaviour is what distinguishes habit from other body reflexes such as dodging a projectile object, or catching/throwing during a sport (Verplanken, Aarts, van Knippenberg, 1997). Following this rationale, the authors concluded that habit can be defined as a goal-directed automaticity (Verplanken et al., 1997). The consistency of repeatedly performing a particular behaviour in the same context becomes paired with the expected outcome which then results in two changes of cognitive processing. First, the behaviour becomes an automatic script that is linked to a goal such as “go to work” or “morning routine”. Second, the expected outcome from the scripted behavior minimizes the cognitive justification of performing, or altering the behaviour routine (e.g. I should ride my bike to work to get in my physical activity, or I should brush my teeth because it’s hygienic). However a
change of context, which changes the script (such as a snow storm), would then prompt a re-evaluation of the behaviour (Aarts, Paulussen, & Schaalma, 1997; Verplanken et al., 1997).

The stability of the environment and context have been documented as key components for habit formation or establishing a habit script (Bargh, 2011; Wood, Quinn, & Kashy, 2002; Wood, Tam, & Witt, 2005). In most cases, when an individual enters the familiar habit script then the cognition related to the task becomes less salient (Verplanken et al., 1997). The behaviour becomes automatic and the thoughts may not correspond to the present activity at hand. The shift of conscious cognition into a background process has become a key identifier of habit process and, hence, Wood and colleagues (2002) define habit as a discrepancy between thought and behaviour. The conscious mind does not become burdened at the present task when the behaviour is habitual as the actions come under control of the environmental contexts or cues (Aarts & Dijksterhuis 2000a; Bargh, 1990; Ouellette & Wood, 1998; Sheeran et al., 2005; Verplanken & Aarts, 1999; Wood et al.2005; Danner et al., 2008). Based on this theorizing, habit has also been defined as a predisposition to automatically enact behaviours in specific contexts based on previous context-behaviour associations (Ouellette & Wood, 1998) and that it can be established from repeated stimulus-response associations (Gardner, 2014). Due to the complexity of this construct, one definition alone may not be sufficient to encompass the meaning of this variable and hence, each of these definitions provides an important element in defining this construct. Despite some variability, all of the proposed definitions converge into the concept that habit allows the behaviour to be performed more easily than if it was consciously regulated. This could be a key characteristic that would help individuals facilitate a regular exercise routine. Despite the numerous events and campaigns that promote exercise, behavioural changes usually
do not last. It has been proposed that exercise is not automated and thus not established in an individuals’ daily activities (Verplanken & Melkevik, 2008). Though the theoretical rationale of using habit to incorporate the regularity of exercise is sound (Rothman, Sheeran, & Wood, 2009), the functionality of this construct requires further investigation.

Though the reasoning of turning exercise into a habit is convincing in theory, the literature is currently limited in understanding the role of habit in exercise. Notably, there is a lack of longitudinal and experimental designs which assess exercise habit formation. To my knowledge, only one study has used a longitudinal design to investigate the habit of various health-related behaviours (Lally, Van Jaarsveld, Potts, & Wardle, 2010). The researchers found that it took on average 66 days to develop a health related habit (healthy eating, drinking and exercise) among a small student sample. While this is a compelling finding, it warrants replication and extension with other samples with a focus on exercise habit. Moreover, Lally and colleagues (2010) used predictive regression analysis to determine exercise habit formation. In addition, a literature search revealed that there are currently no experimental studies focused on exercise habit formation.

**Dissertation Objective**

The purpose of this dissertation is to understand the role of habit in facilitating a regular exercise routine. The first step was to conduct a series of literature reviews to develop an understanding on habit models, methods of measuring habit, the role of the physical built environment in predicting behaviour and the dual process approach. The primary aim of the first study will be to understand behavioral and psychological requirements of habit formation. New gym members will be tracked over twelve weeks to understand the psychological and behavioral requirements of establishing an exercise habit. The second study will observe a group of
experienced exercisers (those who have been exercising for at least one year at a gym). The purpose of this study is to investigate how habit functions in maintaining regular exercise behavior by measuring habit in different phases of exercise. Finally, the last study will be a randomized-controlled trial which will implement an exercise habit intervention in new gym members by incorporating the findings of the previous two studies.
Chapter 2: Exercise Habit Formation in New Gym Members- A Longitudinal Study
Abstract

Reasoned action approaches have primarily been applied to understand exercise behaviour for the past three decades, yet emerging findings in nonconscious and dual process research show that behavior may also be predicted by automatic processes such as habit. The purpose of this study was to: i) investigate the behavioral requirements for exercise habit formation, ii) understand how the dual process approach predicts behaviour, and iii) what predicts habit by testing a model (Lally & Gardner, 2013). Participants (n=111) were new gym members who completed surveys across 12 weeks. It was found that exercising for at least four bouts per week for six weeks was the minimum requirement to establish an exercise habit. Dual process analysis using Linear Mixed Models (LMM) revealed habit and intention to be parallel predictors of exercise behavior in the trajectory analysis. Finally, the habit antecedent model in LLM showed that consistency (β=.21), low behavioral complexity (β=.19), environment (β=.17) and affective judgments (β=.13) all significantly (p<.05) predicted changes in habit formation over time. Trainers should keep exercises fun and simple for new clients and focus on consistency which could lead to habit formation in nearly six weeks.

Keywords: habit, dual process, exercise, MVPA, longitudinal
Introduction

Incorporating 150 minutes of moderate-to-vigorous intensity physical activity (MVPA) a week has been associated with the prevention of at least 25 chronic health diseases and conditions (Garber et al., 2011; Warburton, Katzmarzyk, Rhodes, & Shephard, 2007); however, most adults do not meet these recommendations (Colley et al., 2011; Troiano et al., 2008). Thus, understanding factors that contribute to regular MVPA is paramount. Research in the past decades has investigated this issue primarily through reasoned action approaches (Hagger, 2010; Head & Noar, 2014; Linke et al., 2014; Rhodes & Nasuti, 2011), that assume behavior is a volitional and reflective process (Sheeran et al., 2013). However, a combination of several recent reviews outlining the shortcomings of reasoned action approaches, combined with emerging proponents of alternative frameworks, have suggested that a movement beyond reasoned action approaches could be insightful (Ekkekakis et al., 2013; Rhodes, 2014a, 2014b; Rhodes & Nigg, 2011; Sheeran et al., 2013; Sniehotta et al., 2014). In line with this reasoning, one direction to consider are models that also incorporate unconscious processes (Sheeran et al., 2013). It has been proposed that conscious intention and unconscious processes operate parallel on behavior which is known as a dual process approach (see Evans, 2008 for review). Based on previous conscious rational models, social cognitive theorists propose intention to be the strongest predictor of behavior, thus suggesting intention as the primary conscious motive for behavior (Ajzen, 1991; W. Rogers, 1974; Rosenstock, 1974). By contrast, research in unconscious processes have ranked habit as possibly the strongest unconscious determinant of behavior (Sheeran et al., 2013).

Habit can be defined as “a learned sequence of acts that have become automatic responses to specific cues, and are functional in obtaining certain goals or end-states”
Habit is thought to have a reciprocal relationship with behavior (Gardner, 2014), where habit affects behavioral repetition but that repetition also strengthens habit formation. Overall, habit has demonstrated predictive validity in the physical activity domain; for example, a recent meta-analysis found it to correlate $r = .43$ with behaviour which is similar to the magnitude of the intention-behavior relationship (Gardner, de Bruijn, & Lally, 2011).

Despite the importance of habit outlined in these reviews, there are still several limitations in the contemporary habit literature. For example, the majority of the studies on exercise habit are cross-sectional (Gardner et al., 2011). Given that habit is a dynamic construct, longitudinal studies would provide stronger support for understanding habit formation (Gardner, 2014; Lally et al., 2010). To the authors’ current knowledge only one study has used a longitudinal design to understand habit development (Lally et al., 2010); the researchers found that it took on average 66 days to develop a health related habit (healthy eating, drinking and exercise) among a small student sample. Though this is a compelling finding, it warrants replication and extension with other samples with a focus on exercise habit. Exercise is a type of physical activity that is planned, structured, and repetitive (WHO, 2015). However, it is important to note that 95% of Canadian adults fail to achieve the recommended physical activity guidelines (Colley et al., 2011) with the majority of unsuccessful adopters ranking time as the largest barrier to their exercise (Salmon, Crawford, Owen, Bauman, & Sallis, 2003). With these findings in perspective, simply prescribing the general population to exercise every day for over two months is not a realistic goal. It would be helpful to understand the minimum exercise frequency and time required to successfully establish an exercise habit. Behavioral frequency or repetition is a necessary component for habit formation (Ouellette & Wood, 1998), thus it would
stand to reason that habit formation is partly dependent on time and frequency. Currently, no study has examined a time x frequency effect on habit formation.

A second shortcoming in the habit and exercise literature is the limited understanding of the antecedents required for habit formation. Several models have been proposed to predict habit formation (Aarts, Paulussen, et al., 1997; Bargh, 1994; Grove & Zillich, 2003; Lally & Gardner, 2013; Triandis, 1977; B. Verplanken et al., 1997). Despite some differences in antecedents or the process of habit formation, these models share the importance of behavior repetition based on consistent situational cues or context. One of the most recent models (Lally & Gardner, 2013) suggests there are four antecedents that are conducive for habit formation: reward, consistency, environmental cues, and low behavioral complexity. The researchers theorize the reward component to be intrinsic, which in exercise research could be interpreted as positive affective responses to a behavior (Ekkekakis et al., 2013) or affective judgments (Rhodes, Fiala, & Conner, 2009) about the behavioral experience. Affect has been proposed as having both effects on behavior that are conscious and unconscious (Custers & Aarts, 2005; Williams & Evans, 2014; Zajonc, 1980).

Behaviors that are perceived as complex or have not been sufficiently practiced likely require conscious processes (Verplanken & Melkevik, 2008; Wood, Quinn, & Kashy, 2002) which would consequently prevent automaticity. Building from this research and the habit model proposed by Lally and Gardner (2013), we theorize that behavioral complexity represents the level of challenge of performing a task, independent from motivation or planning. The use of conscious process can also be reduced depending on cues present in the environment. The environment plays a critical role that can prompt or disrupt automatic behavior (Orbell & Verplanken, 2010; Rothman et al., 2009; Wood & Neal, 2009). Environmental cues, such as
mirrors (Sentyrz & Bushman, 1998), lights (Kasof, 2002), or cue cards (Fabio A. Almeida et al., 2005) have predicted behavior in past research. Additionally, close proximity to recreation facilities have also been shown to predict behavior which could act partly via ease of access but also via environmental cues (Kaushal & Rhodes, 2013; Moudon et al., 2007; Rhodes, 2006; Rhodes, Courneya, Blanchard, & Plotnikoff, 2007). In addition to facilitating habit, we theorize that if an individual does not feel comfortable in a particular environment due to the presence of any negative cues (e.g. safety concerns, social physique anxiety, ), then the automaticity process would be interrupted. Hence we theorize that an environment that provides discomfort functions as a distraction that would consequently increase the level of conscious awareness and prevent habit formation.

Consistency is arguably the most unique of the four antecedents as it is a practice rather than feedback (e.g. perceived affect, complexity, or environment). Although the measurement of temporal consistency in exercise is scarce, it has been hypothesized that temporal consistency helps create a protected time for exercise habits (Rhodes & G. J. De Bruijn, 2010). Hence, we define temporal consistency as performing the behavior at a particular time or after a particular activity such as exercising regularly at 6 am or after supper. The closest proposed construct involving consistency is patterned action (Grove & Zillich, 2003; Grove, Zillich, & Medic, 2014). The purpose of this study was to understand habit formation in new gym members. This was a relevant population for this study as the enrolment spike during the New Year followed by a large drop-out of gym members is a well-known trend but is not clearly understood. The objectives of the present study were trifold with a focus on understanding: i) exercise behaviour, ii) habit formation, and iii) habit predictors.
i) The first objective was to test the dual process approach by investigating how habit predicts exercise behavior over 12 weeks while controlling for intention. It was hypothesized that habit and intention would both be required to work in synergy to predict exercise.

ii) The second objective was to further understand habit formation by: i) determining how long it takes to develop an exercise habit, ii) discerning the cut-off score for habit, and iii) testing for a time X behavior interaction. The time required for habit formation would be found by conducting survival analysis. This analysis determines when the changes of habit scores would no longer be significant across time (Bland & Altman, 1998; Greenhouse, Stangl, & Bromberg, 1989; Luke & Homan, 1998). Habit cut-off score would be revealed by Receiver Operating Characteristic (ROC) analysis with habit being the test variable and exercise requirement as state variable; the cut-off score was identified from having the highest sensitivity and lowest specificity values (Greiner, Pfeiffer, & Smith, 2000; Kraemer et al., 1999). We hypothesized that habit formation depended on frequency (Gardner, 2014) and time; hence, this can be represented with the following equation: time X frequency = habit strength. Previous analyses which would identify the time required for habit formation and cut-off score was substituted as time and habit strength respectively in the equation to determine the frequency requirement.

iii) The final objective was to test the multivariate model by Lally and Gardner (2013) to predict habit development. We hypothesized that habit formation first depended on affective judgements about exercise, as a repeated behavior without reward would require conscious evaluation. We also expected that complexity would be a strong antecedent as it could determine if the behavior is consciously directed or automatically brought to attention. Finally, we expected that practice consistency would be a strong predictor of habit formation to reinforce stimulus-
response (S-R) (environment-affect) as well as operant response (O-R) (exercise-affect) systems (Skinner, 1954).

Method

Participants and Procedure

One hundred and forty four adults showed interest in participating in our study by requesting a consent form and, of these individuals, 77% (n=111) signed the consent and completed baseline measures. Participants were excluded if they indicated that they did not meet one of the following inclusion criteria: i) being in the age of 18-65 years, and ii) being a recent gym member, which was defined as someone who has joined a gym/recreation centre within the past two weeks. Thirteen gyms and recreation centres were randomly contacted in the Greater Victoria region in British Columbia, Canada. Eleven of the 13 facilities granted permission to advertise this study. Methods of advertising included: posting wall posters in high traffic areas (i.e., main lobby, water fountain, change rooms), placing information sheets at the main desk, and on-site recruitment which was performed by the primary investigator. Potential participants who were interested contacted the primary investigator to receive the consent form via e-mail along with a web link to the baseline survey. Consent was implied if participants clicked on the link and completed the baseline survey. Follow-up questionnaires were sent at week six, nine, and twelve. We used a 12-week longitudinal design based on the average time required to develop habit in a prior research study (66 days) (Lally et al., 2010). All questionnaires measured the same constructs described under the Instruments section. The questionnaires and study protocol were approved by Human Research Ethics at the University of Victoria.
Instruments

The participants were instructed to consider the definition of “exercising regularly” as performing 30 minutes of moderate-to-vigorous in duration five times per week (CSEP, 2011). They were advised to only count exercise that was done during free time (i.e. not occupation or housework).

Exercise

Exercise was measured by administering the Godin Leisure Time Exercise Questionnaire (GLTEQ) (Godin, Jobin, & Bouillon, 1986a). The questionnaire consists of three open-ended questions of time and frequency spent on type of physical activity (mild, moderate and strenuous). The 2-week test–retest reliability of the measures of total physical activity and the frequency of activity have been estimated to be 0.74 and 0.80, respectively (Godin, Shephard, & Colantonio, 1986). For the purpose of this study, only moderate and strenuous values were used to calculate the exercise behavior. These categories reflect the definition of MVPA provided by recommended guidelines (Garber et al., 2011; Warburton et al., 2007).

Exercise Habit

Exercise habit was assessed by administering the Self-Report Behavioral Automaticity Index (SRBAI) (B. Gardner, 2012; Gardner, Abraham, Lally, & de Bruijn, 2012a). This scale has been modified from the Self-Report Habit Index (SRHI) which was developed by Verplanken and Orbell (2003). The SRBAI consists of 4 items on a 5-point Likert scale with 1 being strongly disagree to 5 being strongly agree. The question stem stated “When I exercise...” which was then followed by four items on the scale: “I do it without having to consciously remember”, “I do it automatically”, “I do it without thinking”, and “I start before I realize I am
doing it”. The internal consistencies of this measure were high across baseline ($\alpha=.84$), week 6 ($\alpha=.92$), week 9 ($\alpha=.91$), and week 12 ($\alpha=.95$).

**Intention**

Intention was used as the proximal measure of reflective, conscious motivation to enact exercise. This construct was assessed by using a continuous open measure worded, “I intend to engage in regular exercise ____ times per week for the next twelve weeks” (Courneya, 1994). Continuous open measurement of intention preserves scale correspondence with our measure of behavior and has been shown to be a superior predictor of behavior over dichotomous closed measures of intention (Courneya, 1994; Courneya & McAuley, 1994; Rhodes, Matheson, & Blanchard, 2006).

**Reward**

A modified version of the Subjective Exercise Experience Scale (SEES) (McAuley & Courneya, 1994) was used to measure exercise reward in the form of affective judgments about exercise. This instrument has been shown to be a valid and reliable measure of affect in a variety of exercise settings (Lox & Rudolph, 1994; McAuley & Courneya, 1994). Items that did not convey a sense of reward were removed from the scale a priori which were: drain, exhaust, fatigue, tired, and strong. These terms reflect energy levels which could be independent from affective reward. For instance, an individual can experience a very enjoyable run (intrinsically rewarding) but feel tired after. The remaining items included: great, positive, terrific, and reverse-scored items of awful, crummy, discourage and miserable. The Cronbach alphas across each measurement period were: baseline ($\alpha=.84$), week 6 ($\alpha=.84$), week 9 ($\alpha=.86$), and week 12 ($\alpha=.90$).
Consistency

Temporal consistency had not been assessed in previous research at the time of the study. Hence, a measure was created to assess this construct. The item read, “How consistently did you exercise at the same time each day (e.g., every morning at 7 am, or exercising daily after supper)?” The options ranged on a 5-point Likert scale with 1= not consistent, always at a random time to 5= very consistent.

Environment

Asking participants to recall an object or context which functions as a cue has been shown to be problematic (Gardner & Tang, 2013). The researchers proposed that individuals may not be able to accurately recall particular cues as they influence behavioral responses on an unconscious level. It is likely that a distinct stimuli/change in the environment would be consciously processed and disrupt automaticity such as encountering a construction site while driving or the presence of an uncomfortable object on the driver’s seat. We theorized that an individual would not be in an automatic state if he/she felt threatened in the environment as this would trigger conscious sensory awareness. Herein, an item worded “How comfortable do you feel in your exercise environment” which was scored on a 5 point Likert scale (1=not very comfortable to 5= very comfortable) was used to assess if the environment supported the process of behavior.
Behavioral Complexity

Similar to consistency, a measure to assess behavioral complexity of performing exercise has not been used in previous research. A behavior that an individual finds difficult would require conscious deliberation to perform and consequently hinder automaticity. The original Self-Report Habit Index (SRHI) (Verplanken & S. Orbell, 2003) recognized this importance and incorporated related items. The present study applied these items to function as antecedents to automaticity based on the proposed model (Lally & Gardner, 2013). Hence two items were adapted from the SRHI which included: “Exercise is something that i) requires effort to do, and ii) I find hard to do” (Verplanken & S. Orbell, 2003). In addition, an individuals’ physical ability could also reflect behavioral complexity. For instance, a novice exerciser would not be as fluent exercising compared to an experienced individual. An item adapted from Rhodes et al., (2006) was also incorporated into this scale which was worded “I have good athletic ability”. All three items were measured on a 5-point Likert scale with 1= strongly disagree to 5= strongly agree. The internal consistencies in the present study were: baseline ($\alpha=.80$), week 6 ($\alpha=.76$), week 9 ($\alpha=.73$), and week 12 ($\alpha=.77$).

Analysis Plan

i) Dual Process Approach

Linear Mixed Model (LMM) in SPSS 20.0 (IBM, 2011) was used to understand how intention and habit predicted exercise behavior across time (Field, 2009; Shek & Ma, 2011; B. T. West, 2009). LMM provides strong methodological advantages over traditional repeated measures analysis of variance which includes: i) maintaining precision with multiple time waves,
ii) examining intra- and inter-individual differences in the growth parameters (e.g., slopes and intercepts), iii) selecting an appropriate covariance structure for the growth curve model (this helps reduce error variance as researchers can choose the correct model that reflects the patterns of change over time), and iv) handling missing data (for further explanation see Field, 2009; Shek & Ma, 2011). LMM allowed for simultaneous assessment of the effects of within-person variation in predictor variables (level 1) across each time measurement (level 2). Before any analysis was conducted, the time parameters were grand mean centered to reduce multicollinearity (UCLA: Statistical Consulting Group, 2014). The next procedure involved a series of steps to determine appropriate model fit (Field, 2009). This consisted of first determining if a random intercept would provide a significant difference based on Chi-squared values. A random intercept in a longitudinal model tests the assumption that each participant can have his or her own starting point. The next step consisted of calculating the Intraclass Correlation Coefficient (ICC) on the baseline model. The ICC describes the amount of variance in the outcome from differences between individuals. A high ICC value indicates the stability of the dependent variable over time. The last step involved conducting a slope analysis to identify which time polynomial would provide a suitable fit for the model (Field, 2009; Shek & Ma, 2011; B. T. West, 2009).

Once the model demonstrates appropriate fit parameters (Field, 2009) then LMM/multilevel analysis can be performed by selecting the Restricted Maximum Likelihood for estimation method (Field, 2009). Two sets of multilevel analysis were performed which consisted of testing intention and habit as predictors of exercise behavior at baseline and at trajectory/across time.
**ii) Habit Stabilization, Cut-off Score, and Required Frequency**

LMM was also used to determine if length of time for habit formation would be moderated by frequency of behavior. The interaction can be represented by the following equation: \( \text{time} \times \text{frequency} = \text{habit strength} \). Identifying the values for this equation is a multi-step process which involved finding how long it takes for habit to develop and identifying the interaction value. Survival analysis was used to understand the stability of habit formation; in particular, this analysis determined when the changes of habit scores were no longer significant across time (Bland & Altman, 1998; Greenhouse et al., 1989; Luke & Homan, 1998). The next step involved calculating a cut-off score for habit formation. Determining the cut-off score was performed by Receiver-Operating Characteristic (ROC) analysis (Greiner et al., 2000; Kraemer et al., 1999). ROC curves were constructed by plotting true-positive rates (sensitivity) against false-positive rates (1-specificity). “Habit” was the test variable and “exercise requirement” was the state variable. The cut-off values for each time period were then averaged to find the overall cut-off score for the measure. Cut-off values were determined by identifying points on the curve which demonstrated maximum sensitivity and minimal specificity. The area under the curve was also calculated with 95% confidence interval (Greiner et al., 2000; Kraemer et al., 1999). Finally, the time requirement for habit formation and cut-off values were then substituted as “time” and “habit strength” respectively in the interaction equation to determine the required minimum “frequency” to achieve habit formation. This would then be tested by first grouping participants into meeting, or not meeting the required frequency values then using those groups to predict habit formation in LMM.
**iii) Habit Antecedents as Predictors of Habit Formation**

LMM was used to test if the antecedents (affect, consistency, complexity, and environment) predicted habit formation. This was a similar procedure to the Dual Process Approach which first involved testing the antecedents as predictors of habit at baseline followed by a time-varying model. Four time measurements of each variable were used to test if the change of each of antecedent predicted change of habit in the trajectory LMM.

**Results**

*Descriptives*

The mean age of participants was 47.7 (SD = 13.5 years), 70% were female, and the BMI was 25.8 (SD = 4.63), suggesting an overweight sample (NIH, 2011). The majority of the participants completed post-secondary education with 59% of the sample having a university degree. Approximately 40% had a household income >$75 000. The participants reported an average of 186 (SD= 158) minutes of total physical activity (light, moderate and vigorous) but 72% were not meeting the recommended exercise guidelines at baseline (Garber et al., 2011; Warburton et al., 2007). All participants were within their first two weeks of enrolling in their gym or recreation centre and reported being a new member in a gym or recreation facility with the intention to develop a regular exercise routine. Descriptive data for the participants are displayed in Table 1. Bivariate correlations of the antecedents with habit and exercise are presented in Table 2.
Dual Process Approach

i) Model Setup and Baseline Analysis

Habit and intention were placed in LMM to compare the model with and without a random intercept. The Chi squared difference was not significant [$\chi^2 (1, N = 111) = 1.96, p=.37$]. Thus, participants’ random starting points did not significantly change the model (Field, 2009). The baseline model did not find habit $F (1, 101)= .22, p = .64$; or intention $F (1, 101)= .90, p = .34$ to be a significant predictors of exercise. The ICC intercept/ (intercept + residual) = .76, suggesting that about 76% of total variation from the predictors was due to individual differences. ICC values were in acceptable range for model fit (> .25) and allowed us to proceed with testing independent growth curves (Shek & Ma, 2011).

ii) Trajectory Analysis

Analysis of independent growth curves (IGC) was used to understand which polynomial value of time would demonstrate the best fit for changes in exercise. The 2-log likelihood was used to calculate the chi squared difference which was significant between all three models. Since all three time slopes showed significance, the Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) were compared. From these results, the cubic polynomial was selected as smaller statistical values reflect stronger model fit to the data (Shek & Ma, 2011). The trajectory model showed habit habit, $\beta=.23 (p=.001)$ and intention, $\beta=.23, (p=.007)$ to be equivalent in strength for predicting exercise behavior across time.
**Habit Stabilization, Cut-off Score, and Required Frequency**

LMM was used to test how frequency and time interacted to predict habit. Testing of random intercepts revealed that the Chi squared difference was not significant $[\chi^2 (1, N = 111)=.06, p=.68]$. Thus, a random intercept model did not improve fit (Field, 2009). The baseline habit model found habit to significantly predict exercise $F (1, 99)= 8.78, p = .004$. The ICC value was .38, which means that 38% of total variation from exercise was due to individual differences. This was also in the acceptable range to continue testing IGC.

Test for IGC found a significant chi squared difference between linear and quadratic models, $\chi^2 (1, N=111) = 14.1, p = .03$. It was optimal to proceed with the quadratic time value for further analysis as: i) the study consists of four measurement points and a valid polynomial can be a maximum of one less than the number of time points (Field, 2009), and ii) it has been theorized that habit develops non-linearly (Lally et al., 2010). A quadratic polynomial for time was then used to test habit change across 12 weeks, which was found to be significant $F (1, 233)= 14.96, p = .001$.

The next step was to perform Kaplan Meir survivor analysis to investigate interaction values at each of the time slopes. The Kaplan Meir survival curve showed a significant difference ($p < .001$) between each time curve over the three tests: Log Rank, Breuslow and Trone-Ware. Each of these tests compares the differences between curves (Breslow= first third of the curve, Trone= middle section, and Log rank = last third of curve). Pairwise comparisons were used to further determine the significant differences among the three sections of the curve. This showed that the second curve (week 6) was significantly different ($p<.001$) than baseline but not with the other curves (week 9) and (week 12). This stability suggests that the majority of
habit formation in the sample was by week 6 with an interaction value of 12.16 (lower bound of 95% Confidence Interval).

**Habit Cut-off Score**

Four separate ROC analyses were performed for habit scores at each time point. The baseline cut-off was 2.91 with a sensitivity of 0.70 and 1-specificity of .18. The AUC value was 0.76 (95% CI: 0.667–0.858, p<.001). Cut-scores for week 6, 9 and 12 were 2.52, 2.76, and 3.01 respectively which averaged a cut-off score of 2.80. The AUC values ranged from .63-.76 and were considered in acceptable range (Akobeng, 2007; Fischer, Bachmann, & Jaeschke, 2003).

**Frequency Required for Habit Formation**

Previous analysis found habit stabilized at week 6, with the interaction value of 12.16 (lower bound of 95% Confidence Interval). We substituted this value with the habit cut-off score of 2.8 in the equation to solve for minimal frequency of exercise bouts required to achieve habit formation and found that a frequency of approximately four days per week was required to achieve an interaction score of 12. This finding was tested by first determining if behavioral frequency predicted habit across time. The LMM analyses found behavioral frequency to predict habit over twelve weeks (β=.24, p<.001). The next step involved separating values based on high frequency (≥ 4 days/week) and low frequency (< 4 days/week) groups. When these groups were then tested as predictors of habit, the low frequency group did not predict habit (β=.09, p=.42) but the high frequency group was significant (β=.24, p<.001). A descriptive plot was produced to depict how the frequency groups affected habit scores across time (Figure 1). The figure shows that those in the high frequency group demonstrated stability of habit scores and maintenance of habit (≥ 3/5) scores across the 12 weeks. In particular, week 6 shows that 61.5% of participants
achieved habit in the high frequency group compared with 44.8% in the low frequency. By week 12, the values for high and low frequency groups were 63.8% and 22.6% respectively.

**Habit Antecedents as Predictors of Habit Formation**

*Baseline*

The four baseline antecedents (affect, consistency, complexity, cues) were placed in LMM to compare two variations of the model: with and without a random intercept. The Chi squared difference was not significant [$\chi^2 (1, N = 111) = 3.82, p = .12$]. Thus, a random intercept model did not improve fit (Field, 2009). The ICC intercept/ (intercept + residual) = .64, suggesting that about 64% of total variation from the antecedents was due to individual differences. LMM analysis of the baseline habit model found that affective judgments (reward) predicted habit with a medium-large effect size $\beta = .47, F (1, 106)= 31.56, p < .001$ followed by consistency $\beta = .45, F(1, 106)= 13.36, p = .001$; behavioral complexity and cues were not significant (Table 3).

*Trajectory Analysis*

The following trajectory analysis revealed if the antecedents contributed a significant change to habit scores across 12 weeks. A quadratic polynomial was used for the trajectory analysis as the results from the previous IGC found this time slope to be a suitable fit for a model with habit as the DV. When time was added in the trajectory analysis, consistency demonstrated the largest effect size for predicting habit formation, $(\beta = .21, p < .001)$, followed by low behavioral complexity $(\beta = .19, p < .001)$, environment $(\beta = .17, p = .008)$ and affective judgements $(\beta = .13, p = .003)$ (Table 3).
The primary purpose of this study was to understand the process of habit formation in new gym members over 12 weeks. The secondary purpose was to investigate how the dual process approach predicts exercise and how the antecedents in the habit model predict habit formation. The present study found the SRBAI (Gardner, Abraham, Lally, & de Bruijn, 2012b) to have a cut-off score of 2.80/5. With regards to behavioural requirement for habit formation, it was found that participants who exercised for at least four bouts per week for six weeks successfully established an exercise habit. Dual Process tests showed that intention and habit were not significant at baseline but they became equal predictors of exercise in the trajectory analysis. Finally, the habit model found that affect and consistency were the largest predictors for people starting a habit; however, trajectory analyses revealed that consistency was the most important predictor followed by low behavioral complexity, environment, and affect.

It was hypothesized that habit and intention would both be significant predictors of exercise, commensurate with the Dual Process approach. Though habit significantly predicted exercise behavior during baseline; however, both constructs became significant predictors over time with equal effect sizes ($\beta=.23$, respectively) in support of our hypothesis. The non-significant finding of intention at baseline could be attributed to the sample being new gym members with already high intentions. Intention-based approaches have been criticised in this particular situation and it represents a practical application of the intention-behavior gap (Rhodes & De Bruijn, 2013). However, as time progressed, the change of intention and habit scores predicted change of exercise over the 12 weeks. Overall, the results add support to a small literature on Dual Process approach applied to exercise behavior (Calitri, Lowe, Eves, &
In terms of the time required to establish an exercise habit, exercise habit plateaued on the 6th week (42-49 days) of the study with 48% of the sample achieving habit formation. Previous work has found that it took an average of 66 days to establish a health related habit (Lally et al., 2010). However the differences in methodologies do not warrant much comparison. For instance, Lally et al., (2010) used a combination of data and projected analysis to determine exercise habit formation. The present study also found the cut-off score of the SRBAI to be 2.8/5 using ROC analyses. This indicates that 2.8/5 is the minimal score to detect that the behavior is not entirely controlled by conscious processes. Scores ≥ 2.8/5 would suggest that automaticity is significantly involved in the behavior. The score is fairly low on the measure, suggesting that automaticity may be a continuum where low scores still represent predictive values. Scores that are very low on this continuum would reflect high cognitive process with minimal automaticity (e.g. controlling air traffic) and the other end of the continuum would indicate the opposite (e.g. sleeping). These findings and theorizing satisfy both perspectives of habit research; the results support theorizing that exercise is not completely automatic (Maddux, 1997) yet it demonstrates that habit may be critical for exercise continuance (Rhodes & De Bruijn, 2013).

Although, the present study estimated a similar time required for habit formation to Lally et al. (2010), we also hypothesized that time would be moderated by performance frequency. The results clearly supported this conjecture, with a time X frequency interaction. A large drop (44.8% to 22.6%) in habit was noticed from week 6 and to 12 in the low frequency group (less than 4X/week); however, those in the high frequency group maintained habit across time (61.5%
to 63.8%). Theoretically, this pattern aligns with several models that propose establishing a habit requires repeated behavioral practice across time (Hall & Fong, 2007; Ouellette & Wood, 1998; Rhodes & De Bruijn, 2013; Triandis, 1977; R. West, 2006). Fortunately, these findings are also aligned with public health guidelines suggesting that an exercise habit can be achieved in 4-5 bouts with 30/40 minutes per session (Garber et al., 2011; Warburton et al., 2007).

We also hypothesized that habit formation would depend on the presence of the antecedents theorized by Lally and Gardner (2013), with affective judgments and complexity predicting habit in the initial phases but consistency predicting habit formation over time. We had some support for this hypothesis. Affective judgments about the exercise experience were found to be the primary predictor of habit formation at baseline but consistency became the strongest predictor in the trajectory analysis. This supports prior theorizing on the foundation of habits. Affect has been investigated in understanding general unconscious goals (Custers & Aarts, 2005) and habit of fruit consumption (Wiedemann, Gardner, Knoll, & Burkert, 2014) but not for exercise. It is likely that negative feelings which stem from unfavourable experiences could prompt conscious deliberation for the individual before performing the behavior. On the other hand, a positive reward would not require evaluative process; the presence of positive affect may drive behavior at an unconscious level (Custers & Aarts, 2005; Zajonc, 1980).

In terms of consistency, our results support our conjecture that it may be a pillar in establishing both the stimulus-response (S-R) (environment-affect) as well as operant response (O-R) (exercise-affect) conditions as the behavior becomes more familiar. The significant effect of consistency also helps establish a potentially different antecedent for habit formation than motivation. This construct suggests that how, rather than why one practices may be more
important to forming habits. Hence, these results suggest that initiating an exercise routine that is enjoyable and consistent can help in habit formation.

Behavioral complexity was found to predict change of habit across time which aligns with previously theorized research on the importance of low cognitive load for habit formation (Verplanken & Melkevik, 2008; Wood et al., 2002). Although exercise is a complex behavior, it was likely that practicing consistently eventually eased the challenges of the behavior across time thus allowing for the facilitation of habit. A comfortable environment that does not stimulate more conscious thinking was also shown to predict habit over time. Assessment of environmental cue from traditional methods may not be clear due to variability in the type of cues and the method of measurement (Gardner & Tang, 2013). Hence the present finding could provide a novel approach to assess if the environment supports the development of habit.

Despite the longitudinal design, analyses, methods, and novel approach to understanding habit and its antecedents, the present study still has limitations that are important to address. For instance, although the sample consisted of new gym members, there was some variability in their exercise history. Since habit formation occurred by week 6, this suggests that the majority of variation of habit occurred within this period. Assessing habit scores more frequently within the first six weeks could provide a more detailed scope of the habit formation phase. Second, the habit model proposed by Lally and Gardner (2013) presents a strong case of four antecedents of habit which have individually been found to correlate with habit in various studies (Gardner et al., 2011). However, the authors did not provide suggestions on measuring these predictors. The present study used a mixture of previous validated scales and customized items to this model. Although these scales predicted change of habit across time, other measurements of these constructs may yield different findings and this warrants sustained research. Finally, future
research should also employ objective measurement to yield a stronger interpretation of exercise behavior and habit formation.

In summary, the study found support for the dual process approach as intention and habit both predicted exercise over time. Exercising for at least four times per week for approximately six weeks was required to establish an exercise habit. Although affect was found to be the strongest predictor at baseline, consistency was the most important factor for predicting changes in habit. The environment and low behavioral complexity played a significant role in changing habit across time. Exercise promoters should focus on setting a consistent exercise schedule and keeping the workouts fun and skill appropriate to increase the likelihood of habit formation. In addition, the environment should be comfortable and welcoming for new clients. The first 6 weeks appear critical for habit formation and new exercisers should strive to workout at least four times per week.
Tables and Figures

Table 1. Descriptive Data.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household Income</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;$50 000</td>
<td>27</td>
</tr>
<tr>
<td>$50 001-$75 000</td>
<td>32</td>
</tr>
<tr>
<td>$75 001-$100 000</td>
<td>19</td>
</tr>
<tr>
<td>$100 001-$150 000</td>
<td>12</td>
</tr>
<tr>
<td>&gt; $150 000</td>
<td>10</td>
</tr>
<tr>
<td><strong>Job Status</strong></td>
<td></td>
</tr>
<tr>
<td>Homemaker</td>
<td>7</td>
</tr>
<tr>
<td>Temporary unemployed</td>
<td>3</td>
</tr>
<tr>
<td>Part-time employed</td>
<td>23</td>
</tr>
<tr>
<td>Full-time employed</td>
<td>51</td>
</tr>
<tr>
<td>Retired</td>
<td>16</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Less than highschool</td>
<td>1</td>
</tr>
<tr>
<td>Highschool diploma</td>
<td>23</td>
</tr>
<tr>
<td>College diploma</td>
<td>17</td>
</tr>
<tr>
<td>University Degree</td>
<td>29</td>
</tr>
<tr>
<td>Graduate or professional degree</td>
<td>30</td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
</tr>
<tr>
<td>Never Married</td>
<td>13</td>
</tr>
<tr>
<td>Married/common law</td>
<td>76</td>
</tr>
<tr>
<td>Separated/divorced/widowed</td>
<td>11</td>
</tr>
</tbody>
</table>
Table 2. Bivariate Correlations of Habit Antecedents with MVPA and Habit

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>.20*, .48**</td>
<td>.20*, .30**</td>
<td>.10, .48*</td>
<td>.27*, .31**</td>
</tr>
<tr>
<td>Reward</td>
<td>.33*, .59**</td>
<td>.29**, .46**</td>
<td>.14, .21*</td>
<td>.12, .26**</td>
</tr>
<tr>
<td>Behavioural Complexity</td>
<td>.38*, .38**</td>
<td>.25*, .63**</td>
<td>.27**, .48**</td>
<td>.22*, .64**</td>
</tr>
<tr>
<td>Environment Cue</td>
<td>.19*, .44**</td>
<td>.26**, .23*</td>
<td>.17, .18</td>
<td>.28**, .23*</td>
</tr>
</tbody>
</table>

Note. * p<.05, M= MVPA, H= habit
### Table 3: Baseline and Trajectory Analysis: Antecedents as Predictors of Habit Formation

<table>
<thead>
<tr>
<th>Source</th>
<th>B</th>
<th>$\beta$</th>
<th>S.E</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency</td>
<td>.22***</td>
<td>.45**</td>
<td>.13</td>
<td>.14</td>
</tr>
<tr>
<td>Reward</td>
<td>.36***</td>
<td>.47***</td>
<td>.08</td>
<td>.30</td>
</tr>
<tr>
<td>Complexity</td>
<td>.03</td>
<td>.02</td>
<td>.08</td>
<td>-.18</td>
</tr>
<tr>
<td>Environment Cue</td>
<td>.16*</td>
<td>.09</td>
<td>.08</td>
<td>-.06</td>
</tr>
<tr>
<td><strong>Trajectory Analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency</td>
<td>.26***</td>
<td>.21***</td>
<td>.07</td>
<td>.11</td>
</tr>
<tr>
<td>Reward</td>
<td>.05**</td>
<td>.13**</td>
<td>.04</td>
<td>.07</td>
</tr>
<tr>
<td>Complexity</td>
<td>.09*</td>
<td>.19**</td>
<td>.05</td>
<td>.01</td>
</tr>
<tr>
<td>Environment Cue</td>
<td>.14**</td>
<td>.17**</td>
<td>.07</td>
<td>.08</td>
</tr>
</tbody>
</table>

Note. ***p<.001, **p<0.01, * p < .05; $\beta$ = standardized beta, ns = non-significant variability in sample to predict trajectory change
Figure 1: Habit Scores Between High and Low Frequency Groups

Note. High frequency (=>4 times/week), Low Frequency (<4 times/week)
Chapter 3: The Role of Habit in Different Exercise Phases
The Role of Habit in Different Phases of Exercise

Navin Kaushal, Ryan E. Rhodes, John T. Meldrum, John C. Spence
Abstract

Social cognitive theories have dominated research in understanding physical activity (PA) behavior, but constructs that underlie more automatic antecedents of behavior may augment these approaches. One such construct is habit, defined as a system of actions which become automatic responses to certain situations. Due to the complexity of exercise, it has been suggested that there is at least more than one behavioral phase and this may also extend to habits. The primary purpose of this study was to investigate how habit strength in a preparatory and performance phase predicts exercise behavior while accounting for intention. The secondary purpose was to determine the strength of potential habit antecedents (affect, consistency, cues, complexity) in both exercise phases. Participants (n=181) were a sample of adults (18–65) recruited across 11 recreation centers who completed baseline and follow-up questionnaires after six weeks. When predicting exercise behavior, intention (β=.27, p< .001) and habit preparation (β= .18, p< .001) predicted change of exercise behavior across six weeks but not habit performance. The habit models found consistency to be the strongest predictor in predicting habit (β=.28, p<.001) in both phases. The present study highlighted the distinction between the two phases and the importance of preparatory habit in predicting behavior change. Focusing on a consistent preparatory routine could be helpful in establishing an exercise habit that translates to changes in behavior.

Keywords: habit, automaticity, exercise, MVPA, preparation, phases
Introduction

The health benefits of performing 150 minutes of moderate-to-vigorous physical activity (MVPA) regularly has been well documented (Garber et al., 2011; Warburton et al., 2007). However, a large proportion of people are not meeting these recommendations (Colley et al., 2011; Troiano et al., 2008). Hence, the importance of physical activity (PA) promotion cannot be overemphasized.

The use of sound theories has been recommended for understanding PA behavior (Rhodes & Nigg, 2011). For the past two decades, most theoretical applications have applied conscious regulatory models in PA research (Rhodes & Nasuti, 2011). These theories propose that consciously controlled cognition will lead to behavioral enactment as explained in prominent theories such as Protection Motivation Theory (Rogers, 1974), Health Belief Model (Rosenstock, 1974), Theory of Reasoned Action (Fishbein, 1975), and Theory of Planned Behavior (Ajzen, 1991). These theories place intention as the most proximal conscious determinant of behavior. In line with this theorizing, intention has a reliable correlation with PA that acts as a central mediator between other conscious motives and behavior (Armitage & Conner, 2001; Hagger et al., 2002; McEachan et al., 2011). Still, the relationship between intention and behavior is modest. Intention accounts for approximately 23% of the variance in PA and 5% of the variance in PA change (McEachan et al., 2011). Thus, there is considerable unexplained variation in PA.

Recently, some theorists have suggested that intention may act on behavior parallel with unconscious processes; this is known as the dual process approach (Evans, 2008). The dual process approach proposes that behavior is influenced by a combination of both unconscious (rapid, automatic) and conscious (slow, deliberative) processes. Despite the changes in nomenclature of these systems over the past thirty years which include automatic and controlled...
(Schneider & Shiffrin, 1977), stimulus bound vs. higher order (Toates, 2006), Type 1 and Type 2 systems (Evans & Stanovich, 2013), the agreement on definitions and distinction between two systems have remained relatively consistent (Evans, 2008b). These systems can work synergistically, which is known as goal-derived automaticity (Aarts & Dijksterhuis, 2000b; Bargh & Ferguson, 2000; Wood & Neal, 2007), such as driving the same route to work. Alternatively, the systems can work against each other when the outcome goal is not in agreement with strong unconscious drives (Anshel, Kang, & Brinthaupt, 2010).

Although the unconscious system may consist of multiple constructs that are not clearly understood, habit, has seen predictive results for PA (Gardner et al., 2011; Sheeran et al., 2013). Habit has been defined as a process by which a stimulus cues a learned stimulus-response association (Gardner, 2014); habits are also a form of goal-directed automatic behavior (Bargh, 1989). A recent meta-analysis found habit to correlate $r = .43$ with behavior which is similar to the magnitude of the intention-behavior relationship (Gardner et al., 2011). This suggests that these two constructs may carry similar weight in predicting behavior.

Despite the noticeable increase of habit research within the past decade (Gardner et al., 2011), the functionality of habit is still not clearly understood and hence several models have been proposed to predict habit formation (Aarts, Paulussen, et al., 1997; Bargh, 1994; Grove & Zillich, 2003; P. Lally & B. Gardner, 2011; Rhodes & De Bruijn, 2013; Triandis, 1977; B. Verplanken et al., 1997; Wood & Rünger, 2015). A recent model proposed by Lally & Gardner (2013) theorizes that habit formation depends on the presence of four antecedents: reward, low behavioral complexity, consistency, and cues. These variables are a blend of constructs from social cognitive models (reward, low behavioural complexity) combined with those exclusive to habit literature (cues and consistency). Lally and Gardner (2013) define reward as intrinsic,
which in the exercise literature could align with affective judgments (expectations of intrinsic regulation, enjoyment, fun, pleasure) (Deci and Ryan, 1985). Although a limited definition was provided for behavioral complexity, the authors propose that a complicated task would require cognitive resources and hence interrupt habit formation (Kaushal & Rhodes, 2015). Kaushal and Rhodes (2015) theorized that an individuals’ level of ability, skill and control could be components that comprise the behavioural complexity construct. Cues have been well-documented in several habit models (Grove & Zillich, 2003; Lally & Gardner, 2013; B. Verplanken et al., 1997; Wood & Rünger, 2015), and has predicted behavior when manipulated (F. A. Almeida et al., 2005; Kasof, 2002; Sentyrz & Bushman, 1998). Finally, consistency could be defined as a ritualistic practice structure and has been shown as the strongest of the four in possible antecedents (Kaushal & Rhodes, 2015).

Nevertheless, the possibility of habituating physical activity has also been questioned because exercise requires scheduling, effort, which ostensibly seem volitional and mindful (Maddux, 1997). Verplanken and Melkevik (2008) acknowledged the complexity of establishing an exercise habit and have suggested a more viable approach may be the pre-exercise preparation phase. This approach has been applied by testing exercise habit as instigation and execution states (Phillips & Gardner, 2015). Specifically, when the two habit states were tested together, only habitual instigation was found to significantly predict behavior change.

An alternative approach, which further extends the work of Phillips & Gardner (2015), would be to consider exercise as two distinct behavioral phases: a preparatory and performance phase. Each of these phases involves a separate environment with a separate set of actions. We define the exercise preparation phase as behaviors conducted to deliver individuals from their home to an exercise-ready state. Although individualized, it could comprise of various tasks
such as: gathering gym materials (water bottle, mp3 player, workout gloves), wearing or packing gym clothes, or transport to an exercise facility. The preparation phase could be analogous to an individual’s morning routine before leaving for work. The preparation phase ends as soon as the individual begins exercising, which marks the commencement of the performance phase. This phase can be defined as exercising in the supportive environment (e.g., gym, recreation centre, swimming pool). Since the preparation is not as complex as the performance, establishing a habitual preparation practice could be more feasible than habituating the performance itself (B. Verplanken & Melkevik, 2008).

The primary purpose of this study was to understand fluctuations in exercise behavior over six weeks among a group of active gym members using a dual process approach with habit framed in both exercise preparatory and performance phases. Based on our contention that the execution of exercise itself is likely more volitional then preparation, it was hypothesized that intention and preparatory habit would be the strongest predictors in behavior change (Phillips & Gardner, 2015). The secondary purpose was to investigate which antecedents predict the change of each habit construct (preparation, performance) by applying the model suggested by Lally and Gardner (2013). Following the results from Kaushal and Rhodes (2015), it was hypothesized that consistency of practice routine would be the strongest predictor for both habit phases.
Methods

Participants

The mean age of participants was 43.4 ($SD = 15.3$ years), 64% were female, and 66% of the sample identified themselves as being in either “very good” or “excellent” health. The participants were averaging 301 minutes/week ($SD=144$) of MVPA. Approximately 35% of the sample had a household income >$75 000 and 54% were married or in a common law relationship. Further descriptives are presented in Table 1.

Eleven gyms and recreation facilities were approached as venues to recruit participants. Of the eleven, nine centres allowed advertising which included: posting wall posters in high traffic areas (eg. change rooms, spin class rooms, by the water fountain, main lobby/bulletin board) and providing consent forms at the main desk. Four of these centres additionally allowed on-site recruitment which was performed by the primary investigator. The inclusion criteria included adults (18-65) who were affiliated with a gym or recreation centre who have been exercising for at least one year.

Measurements and Procedure

Interested participants e-mailed the primary investigator to receive a digital copy of the consent form along with the link to the baseline questionnaire. Consent was implied if participants followed the link and completed the survey. A follow-up questionnaire was sent after six weeks. The questionnaires and study protocol were approved by Human Resource Ethics at the University of Victoria. The questionnaires defined “exercising regularly” as performing at least 150 minutes of exercise at a moderate-to-vigorous level per week (CSEP, 2011; Garber et al., 2011). Participants were advised to only count physical activity that was performed during their free time (eg. not occupation related or housework). Commensurate with
the purpose of this research, the questionnaires were further divided into exercise preparation and performance sections. For exercise preparation, the participants were provided with the following instructions, “The following questions ask how you feel about the process of preparing to exercise. The exercise preparatory phase includes all of the activities you would regularly perform before each exercise session (e.g., packing gym bag, equipping mp3 player, having a pre-workout snack)”. For the exercise performance section it stated, “The following questions ask how you feel during the exercise phase. The exercise phase includes the time from your first exercise to the end of your last routine/set in the gym”.

**Exercise Behavior**

The Godin Leisure Time Exercise Questionnaire (Godin, Jobin, & Bouillon, 1986b) was used to measure exercise after six weeks. The questionnaire assesses the time and frequency of each type of physical activity by intensity (mild, moderate, and strenuous). This instrument has demonstrated stable test-retest reliability which ranges from .74-.80 (Godin, Shephard, et al., 1986). For the purpose of the present study, only moderate and strenuous values were used to estimate exercise behavior as these categories reflect the definition of exercise provided by recommended guidelines (CSEP, 2011; Garber et al., 2011).

**Exercise Preparation and Performance Habit**

Habit was measured by administering the Self-Report Behavioral Automaticity Index (SRBAI) (B. Gardner, 2012; Gardner et al., 2012a). The SRBAI is a modified version of the Self Report Habit Index (B. Verplanken & S. Orbell, 2003) and consists of 4 items on a 5-point Likert scale with 1 being strongly disagree to 5 being strongly agree. The scale was provided in both the preparatory and performance sections of the questionnaire. The question stem was slightly
modified in each section to assess the respective process (i.e. When I prepare to exercise/When I exercise) which was then followed by four items on the scale: “I do without having to consciously remember”, “I do it automatically”, “I do without thinking”, and “I start before I realize I am doing it”. The SRBAI demonstrated strong internal consistencies of, α = .86 and α = .88 for exercise preparatory and performance habit respectively.

Intention

Intention served as the assessment of conscious motivation exercise over the six weeks. This construct was assessed by using two items (Courneya, 1994; Rhodes, Blanchard, Matheson, & Coble, 2006). These items included, “I plan to engage in regular exercise over the next 6 weeks” which was rated on a 5-point likert scale (strongly disagree to strongly agree) and “I intend to engage in regular exercise ____ times per week over the next 6 weeks”.

Affect/Reward

Following previous work on the habit model (Kaushal & Rhodes, 2015), expected affective judgment was used to reflect reward. Three validated items from previous research were used to measure affect during exercise (Courneya, Conner, & Rhodes, 2006; Lawton, Conner, & McEachan, 2009). These items assessed how “enjoyable”, “exciting”, and “pleasant” the participant felt on a seven point semantic differential scale. A literature review revealed the absence of a validated scale to measure affect/reward in preparing to exercise. Hence, items were created that may reflect feeling states during the exercise preparatory phase. Participants were asked to rate the degree of “annoyed”, “hassled”, and “pleasant” they felt while preparing to exercise on a five-point Likert scale. Annoyed and hassled were reverse scored. The internal consistencies for the performance and preparation scales were .82 and .83 respectively.
Consistency

Temporal consistency was used to measure consistency in the preparatory phase. The item was worded, “How consistently do you exercise at the same time each day? (e.g., exercising every morning at 7 am, or exercising daily after supper)” (see Kaushal & Rhodes, 2015). Performance consistency was also assessed in this manner, “How consistent is your exercise routine?” Both items were scaled from 1= not consistent at all to 5= very consistent.

Environmental Cues

Preparatory cues were assessed with the following item, “I have things from my home or work that remind me to prepare for exercise (e.g. gym bag when I open my car trunk, water bottle on my desk)”. Two items were used in the exercise performance section that included, “In my exercise environment, there are things that remind me of my exercise routine (i.e., certain gym equipment, running trail, etc.)”, and “In my exercise environment, I use cues to gauge my progress (i.e., looking at the next set of weights to lift, passing by familiar landmarks while running)”. The internal consistency for the performance cue items was strong (α = .86).

Behavioral Complexity

Since a validated scale for behavioral complexity is yet to be established (Kaushal & Rhodes, 2015), a combination of previous validated and custom items were aggregated to create the behavioral complexity scales. Preparation is conceptually about the barriers an individual faces prior to performance and thus perceived behavioral control measures may be a good proxy. Items that were used to assess this construct included: “I have complete personal control over exercise if I really wanted to”, “engaging in regular exercise is mostly up to me if I wanted to do so”, and “engaging in regular exercise if I wanted to, would be” (Rhodes & Courneya, 2003,
2004). By contrast, exercise performance reflects an individual’s skill and ability to enact the task. The following three items were used to measure this variable, “I have good athletic ability”, and “I have considerable skill when it comes to doing physical activities” (Rhodes, Blanchard, & Matheson, 2006; Rhodes & Courneya, 2003, 2004). All items were on a five-point Likert scale; the first two items on each scale ranged from “strongly disagree to strongly agree” and both third items ranged from “extremely difficult” to “extremely easy”. The Cronbach alphas for preparation and performance items were .80 and .73 respectively.

**Analysis Plan**

Prior to any hypothesis testing, the dataset was first screened to determine the pattern of missing data. This involved conducting Little’s MCAR (Missing Completely at Random) test, which included all variables involved in the model along with descriptive data (age, BMI, sex, marital status, income, and education). Confirmatory Factor Analysis (CFA) was then performed to determine if habit preparation and performance constructs were sufficiently distinct (Anderson & Gerbing, 1988). A significant Chi square value would indicate that the two phases are significantly distinct to be treated as separate constructs.

Following descriptives of the sample at baseline, bivariate correlations for both the dual process and habit models were analyzed using SPSS 20.0 (IBM, 2011). Analysis of Moment Structures (AMOS) version 20.0 was then used to create the structural equation models to predict the outcome change. A model was designed to test each of the hypotheses which included: the Dual Process approach, habit preparation, and habit performance. The Dual Process approach investigated how intention, habit of preparation, and performance predict change in exercise behavior duration across six weeks. The two habit models consisted of placing the four
antecedents (reward, consistency, behavioral complexity, and environment) as predictors of change in preparation and performance habit respectively. The change (or residual) in exercise behavior construct was calculated by regressing the values at 6-weeks over baseline.

**Model Specification and Fit**

For all structural equation models, appropriate fit was assessed by the Comparative Fit Index (CFI) of equal to or greater than 0.90 (Bentler, 1990), Root Mean Square Error of Approximation (RMSEA) of less than or equal to 0.08 (Steiger, 1990), and a chi-square (CMIN) value less than 5 (Arbuckle, 2006).

**Results**

Missing analysis found 28% of the data to be incomplete. Further investigation using Little’s MCAR test revealed that the data was missing completely at random (Chi-Squared=83.41, DF=65, p>.05). Given the non-significance of the Chi Squared value, the hypothesis of data not missing at random was rejected and it was safe to proceed with multiple imputations (Rubin, 1996; Schafer & Graham, 2002).

The chi-square difference test was found to be significant ($\Delta \chi^2=61.8$, DF=1 $p<.001$) which indicates that the two habit constructs demonstrate discriminant validity (Segars, 1997) and provides rationale to proceed with a model test with two habit constructs.

Bivariate correlations were conducted for the Dual Process approach and each of the Habit models. The dual process model found intention ($r = .28$, $p < .001$), preparatory habit ($r = .22$, $p = .001$), and performance habit ($r = .20$, $p = .003$) to all correlate with changes in exercise behavior at six weeks. Preparatory habit was only found to correlate with consistency ($r = .20$, $p < .001$). However, performance habit was found to correlate with consistency ($r = .29$, $p < .001$),
behavioral complexity (r = .20, p < .001), environmental cues (r = .16, p < .05) and affect (r = .16, p < .05).

The data was found to fit the proposed Dual Process model within the range of acceptability (CMIN= 2.96, CFI= .96, RMSEA= .080). Overall the model consisted of strong factor loadings for intention (.76 - .85), habit preparation (.85 - .96), and habit performance (.80 - .93). The averaged variances for intention, preparation, performance were .82, .89, and .75, respectively. The structural model found intention (β=.27, p<.001) and habit preparation (β=.18, p<.01) to be significant in predicting change of exercise behavior across six weeks. However, habit performance did not predict exercise. Overall, the model explained 17% of its variance (Figure 1).

The habit preparatory model (Figure 2) showed a modest fit with the observed data (CMIN =1.7, CFI= .96, RMSEA = .06). The factor loadings for affect (.66 - .87), complexity (.59 - .87) and preparatory habit change (.80 - .89) ranged from satisfactory to strong. The respective variances for these latent variables were .61, .57, and .69. Analysis of the structural model, however, only found consistency as a significant (p< .05) predictor of change in habit across six weeks (β = .28, p<.001), explaining 8% of its variance.

The SEM showed a borderline adequate fit of the observed data (CMIN =2.82, CFI= .91, RMSEA = .09) when predicting change of performance habit. The factor loadings for affect (.63-.97), complexity (.70 - .84), environmental cues (.84 - .91) and performance habit change (.82 - .89) also ranged from satisfactory to strong. Their respective variances were .66, complexity .60, environment .77, and habit change .74. The structural model found consistency (β = .28, p<.001) to be the strongest antecedent for predicting change of performance habit over six weeks. Behavioral complexity (β = .12, p<.05) and environmental cues (β = .11, p<.05) were also
significant predictors. However, affect was not found to predict performance habit. Overall the model explained 11% of the variance in performance habit (Figure 3).

Discussion

The challenge of adopting a regular exercise routine could stem from its complexity compared with some other health behaviors (Rhodes & Nigg, 2011). For instance, exercise requires a scheduled time and a series of sequential actions unlike hygiene or safety practices (e.g., brushing teeth or wearing a seatbelt). Hence, when considering the concept of an exercise habit, it should not be limited to just the performance phase. Rather, the complexity of exercise and habit would be better investigated as more than one phase (B. Verplanken & Melkevik, 2008). One interpretation of this suggestion would be to test exercise as consisting of a decision (instigation) and execution component which each could be habitual (Phillips & Gardner, 2015). An alternative approach that extends the instigation state would be to treat exercise as a process that consists of at least two distinct behaviors, preparatory and performance—each which carry potential to have habitual processing. Hence, the purpose of this study was to further understand the role of habit in these two exercise phases by testing which phase predicted exercise behavior, and which antecedents predicted habit.

It was hypothesized that a multivariate Dual Process model would work in synergy and thus predict fluctuations in exercise behavior via intention and habit. However, the novel aspect of this study was the separation preparation and execution phases in the exercise process that may contribute to habit. The model revealed that only intention and preparatory habit predicted change in exercise behavior at six weeks. This supports previous research on the Dual Process approach (Conroy et al., 2010; Kaushal & Rhodes, 2015; R. E. Rhodes & G. J. De Bruijn, 2010; Rhodes, de Bruijn, & Matheson, 2010; B. Verplanken et al., 1997). However, this is the first
study to provide evidence for the importance of preparatory over execution actions as habits. Given that exercise performance is separate from exercise preparatory behaviors, this complements the work by Phillips and Gardner (2015) on the importance of habituating the state that precedes the exercising/performance state. When both results are taken together, the findings support the conceptualization that states and behaviors preceding exercise may be the best characterization for how habits influence such a complex behavior like exercise. Further, the emerging results regarding the importance of preparatory habit over execution provide some connection between the past divide between habit theorists and exercise theorists who have dismissed the role of habit in physical activity. Where motivation would be required to perform the exercise behavior, preparatory habit may function as a bridge that transfers an individual to the exercise environment.

The secondary objective of this study was to test how the antecedents in Lally and Gardner’s (2013) habit model predicted change in habit across time. It was hypothesized that consistency would be the strongest predictor of habit in both phases. In support of this finding and past research (Kaushal & Rhodes, 2015), consistency was found to be the largest predictor in both preparatory and performance models and the only significant construct in predicting change of preparatory habit. Consistency is relatively a unique construct to habit that is not part of planning or social cognitive research (Ajzen, 1991; Bandura, 1986; Fishbein, 1975; W. Rogers, 1974; Rosenstock, 1974). In proposed habit models the concept of consistency (Lally & Gardner, 2013) has been acknowledged under various labels such as patterned action (Grove & Zillich, 2003), behavioral repetition (Gardner, 2014) and repeated action under stable contexts (Ouellette & Wood, 1998; Wood & Rünger, 2015). These findings validate the importance of this construct across various habit models.
The present study also confirms previous work which found consistency to be the strongest predictor for habit formation among new gym members (Kaushal & Rhodes, 2015). It is interesting to note that consistency remains a strong predictor of habit for both exercise adopters (Kaushal & Rhodes, 2015) and maintainers (present study). The present study did not find affect to predict change of preparatory habit. It is likely that the acts involved in preparation align more closely with other mundane tasks such as a morning routine or household chores. Given this comparison, strong affect may not be required to facilitate this habit. Additionally, since the sample consisted of regular exercisers, it is likely that their preparatory environment may no longer be consciously salient for cues.

Despite several novel contributions to the exercise habit literature, there are study limitations that warrant mention. The limited advancement of the habit model used in the research (Lally & Gardner, 2013) provided some shortcomings. For instance, although the measures were internally consistent, they were created or adapted given the limited literature at present. Furthermore, an objective measure of exercise would add additional validity to the results. Finally, the sample consisted of healthy, active adults. While this is a sensible sample to investigate a maintenance-level construct like habit, it tells us little about the process of habit formation from adoption. Research on the full process of habit formation to behavioral continuation is needed.

The conscious and deliberate effort required to perform exercise while in a habitual state has been a notable paradox (Maddux, 1997). However the predictive validity of habit in the physical activity literature (Gardner et al., 2011), combined with recent reoccurring findings (Calitri et al., 2009; Conroy et al., 2010; A. L. Hyde et al., 2010; Kaushal & Rhodes, 2015; Phillips & Gardner, 2015; Rebar, Elavsky, Maher, Doerksen, & Conroy, 2014; R. E. Rhodes &
G. J. De Bruijn, 2010) presents more evidence to support rather than refute this construct in the physical activity domain. The present study highlighted the distinction and importance of habit preparation and performance for exercise adherence. The behavioral fluctuations that resulted from changes in preparatory habit suggests that exercise habit formation may be most influential in the preparatory phase and is likely facilitated by practice consistency. Hence, a trial that focuses on establishing an exercise habit by emphasizing temporal consistency during preparation could be a viable approach in promoting sustained exercise.
References


Table 1. Descriptive Data.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household Income</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;$50 000</td>
<td>39</td>
</tr>
<tr>
<td>$50 001-$75 000</td>
<td>26</td>
</tr>
<tr>
<td>$75 001-$100 000</td>
<td>13</td>
</tr>
<tr>
<td>$100 001-$150 000</td>
<td>13</td>
</tr>
<tr>
<td>&gt; $150 000</td>
<td>9</td>
</tr>
<tr>
<td><strong>Job Status</strong></td>
<td></td>
</tr>
<tr>
<td>Homemaker</td>
<td>6</td>
</tr>
<tr>
<td>Temporary unemployed</td>
<td>7</td>
</tr>
<tr>
<td>Part-time employed</td>
<td>25</td>
</tr>
<tr>
<td>Full-time employed</td>
<td>44</td>
</tr>
<tr>
<td>Retired</td>
<td>18</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>High school diploma</td>
<td>12</td>
</tr>
<tr>
<td>College diploma or equivalent</td>
<td>32</td>
</tr>
<tr>
<td>University Degree</td>
<td>30</td>
</tr>
<tr>
<td>Graduate or professional degree</td>
<td>26</td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
</tr>
<tr>
<td>Never Married</td>
<td>30</td>
</tr>
<tr>
<td>Married/common law</td>
<td>54</td>
</tr>
<tr>
<td>Separated/divorced/widowed</td>
<td>16</td>
</tr>
<tr>
<td><strong>Overall Health</strong></td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>21</td>
</tr>
<tr>
<td>Very Good</td>
<td>45</td>
</tr>
<tr>
<td>Good</td>
<td>27</td>
</tr>
<tr>
<td>Fair</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 2. Bivariate Correlations of Habit Antecedents with Exercise and Habit

<table>
<thead>
<tr>
<th>Antecedent</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PR. Consistency</td>
<td>.29*</td>
<td>.21**</td>
<td>.02</td>
<td>.61**</td>
<td>.10</td>
<td>.15*</td>
<td>.11</td>
<td>.20**</td>
<td>.32**</td>
<td></td>
</tr>
<tr>
<td>2. PR. Reward</td>
<td>.20**</td>
<td>.09</td>
<td>.20**</td>
<td>.14*</td>
<td>.16*</td>
<td>.11</td>
<td>.03</td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. PR. Behavioural Complexity</td>
<td>.07</td>
<td>.19**</td>
<td>.26**</td>
<td>.25**</td>
<td>.16*</td>
<td>.08</td>
<td>.25**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. PR. Environmental Cues</td>
<td>.07</td>
<td>.04</td>
<td>.01</td>
<td>.11</td>
<td>.02</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. PF. Consistency</td>
<td>.10</td>
<td>.05</td>
<td>.05</td>
<td>.11</td>
<td>.29**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PF. Reward</td>
<td>.28**</td>
<td>.10</td>
<td>.19</td>
<td>.16*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. PF. Behavioural Complexity</td>
<td>.07</td>
<td>.12</td>
<td>.20**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. PF. Environmental Cues</td>
<td>.00</td>
<td>.16*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. PR. Habit</td>
<td>.43**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. PF Habit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PR = preparatory phase, PF = performance phase, * p<.05, ** p<.01
Figure 1. Dual Process Model

Figure 2. Habit Preparation Model

Figure 3. Habit Performance Model


Figure 1. Dual Process Test
Chapter 4: Establishing an Exercise Habit: A Randomized-Controlled Trial
Navin Kaushal, Ryan E. Rhodes, John C. Spence, John T. Meldrum
Abstract

Exercise behavior has largely been studied via reflective social cognitive approaches. Emerging correlational findings have shown the independent prediction of habit, which represents automatic behaviour from stimulus-response bonds (cued and repetitive action) with exercise. Preliminary research suggests that habit during the preparation of exercise may be the most important predictor of enactment. An opposing view to habit is exercise variety which proposes that flexibility increases autonomy. Currently no experimental research has tested the effectiveness of incorporating habit or variety to promote exercise. The purpose of this study was to conduct a randomized-controlled trial to examine the promotion of preparation habit formation compared with control and variety groups on exercise behavior. New gym members (n=141) were recruited across Victoria, BC for this eight-week, three-arm randomized-controlled trial. Participants in the habit and variety groups attended their respective workshops and received a phone call booster follow-up at week four. An ANCOVA controlling for baseline exercise found the habit group to increase in exercise time compared to the control and variety group for both accelerometry (control p<.05; d=.40; variety p=.07; d=.36) and self-report (control p<.05; d=.51; variety p<.05; d =.50). The habit contextual model found the habit group to engage in significantly more exercise consistency, cue use and demonstrated greater automaticity during preparation (η²=.05 to .08; p<.05). Future research is needed to replicate these findings and extend the duration of assessment to evaluate whether changes in exercise behavior are sustained across time.

Keywords: habit, automaticity, exercise, MVPA, RCT
Introduction

The incorporation of regular physical activity can prevent at least 25 chronic diseases and illnesses (Warburton et al., 2007). For instance, adults should engage in at least 150 minutes of moderate-to-vigorous physical activity (MVPA) per week to reap the health benefits (Garber et al., 2011; Warburton et al., 2007). Despite these convincing findings, the majority of the adult population in developed countries struggle to achieve these requirements. A national health study using accelerometers found that only 15% of Canadian adults achieve 150 minutes per week of MVPA in 10-minute bouts (Colley et al., 2011). Thus, understanding how to successfully incorporate regular MVPA into the lifestyle of adults cannot be over emphasized.

The majority of research on MVPA promotion has investigated this issue through reasoned action approaches (Hagger, 2010; Head & Noar, 2014; Linke et al., 2014; Rhodes & Nasuti, 2011), that assume PA is entirely a volitional and reflective process (Sheeran et al., 2013). In other words, these researchers propose that our behaviours are always under conscious control. This approach has been explained through the application of various social cognitive theories such as the Protection Motivation Theory (Rogers, 1974), Health Belief Model (Rosenstock, 1974), Theory of Reasoned Action (Fishbein, 1975), and Social Cognitive Theory (Bandura, 1986) and the Theory of Planned Behavior (TPB) (Ajzen, 1991). For instance, the TPB suggests attitudes, subjective norms and perceived behavioral control determine one’s intention which then predicts behavior. Hence, a person is thought to engage in PA due to volitional intent, which arises from reflection on the potential benefits, social pressure, and capability to perform PA. Recent reviews highlight the limitations of several cognitive theories when compared with alternative frameworks and proposed that investigating beyond reasoned action approaches could be insightful (Ekkekakis et al., 2013; Rhodes, 2014a, 2014b; Rhodes &
Nigg, 2011; Sheera n et al., 2013; Sniehotta et al., 2014). For instance, meta-analyses and reviews have not shown these theories to be effective in predicting PA behavior change (Prestwich et al., 2014; Sniehotta et al., 2014). This may have stemmed from the fact that these theories were adopted from the social psychology literature and hence were not designed to predict change of PA behavior (Rhodes, 2014a; Rhodes & Nigg, 2011). For instance, the complexity of PA requires the investigation of additional constructs that are absent in social cognitive models (Rhodes, 2014a).

One alternative of the deliberative or reflective approach is that behaviors can be determined by non-conscious processes (Evans, 2008b; Evans & Stanovich, 2013). For example, the dual process approach proposes that our behavior is influenced by a combination of both unconscious (rapid, automatic) and conscious (slow, deliberative) processes (Evans, 2008b). An unconscious mechanism that has been proposed to specifically link to behavior is habit (Gardner, 2014; Wood & Rünger, 2015). Habit can be defined as a behavior that occurs from stimulus response bonds forged by repeated exposure to contextual cues paired with a behavioral act (Wood & Rünger, 2015). In PA research, habit has demonstrated parallel predictive strength to the intention-behavior relationship (Gardner et al., 2011). However, habit itself could be influenced by other antecedents such as behavioral complexity, reward, cues and consistency (Lally & Gardner, 2013). One of the latest models proposes that goals facilitate habit formation by initially motivating individuals to repeat actions in the same context. Thus, habit is the unconscious process that builds from the contextual cues (Wood & Rünger, 2015).

Generally the dual process approach has been shown support in the PA literature. Out of seven studies that analyzed conscious motivation (e.g. intention) with habit, five showed support
for the Dual Process approach (Conroy et al., 2010; Kaushal & Rhodes, 2015; Rhodes & De Bruijn, 2010; Rhodes et al., 2010; Verplanken et al., 1997) and two did not (de Bruijn & Rhodes, 2011; Rebar et al., 2014). However, limited work has been conducted on the theory of antecedents.

Specifically, only one study has tested habit antecedents as proposed by Lally and Gardner (2013), to predict habit formation (Kaushal and Rhodes, 2015). The longitudinal analysis revealed that cues and consistency predicted change in habit across twelve weeks with consistency demonstrating the largest effect size. However, establishing an exercise habit should not be limited to just the performance phase or exercising phase (e.g. running, weight training, swimming). Rather, the complexity of exercise habit would be better investigated as more than one phase (Verplanken & Melkevik, 2008). One interpretation of this suggestion is to test exercise as behavior that consists of a decision (instigation) and execution component which each could be habitual (Phillips & Gardner, 2015). For instance, Phillips & Gardner (2015) found that it was the change of exercise instigation habit that predicted exercise frequency rather than execution habit. This finding mirrors with my second dissertation study that found preparatory habit (e.g. packing workout gear, driving to the gym) predicted behavior change instead of exercise performance.

Exercise Variety

Habit formation models are grounded in the concept that predictability will help facilitate a routine which in turn would establish adherence. Interestingly, exercise variety research, which is based on Self-Determination Theory (SDT) (Ryan & Deci, 2000), suggests that variety helps in exercise adherence by fuelling one’s intrinsic motivation towards exercise due to the
refreshing stimulation (Pronin & Jacobs, 2008; Silvia, 2006). The autonomy afforded by variety should help anticipated affect (enjoyment, intrinsic regulation) which in turn will foster sustained participation (Dimmock, Jackson, Podlog, & Magaraggia, 2013). In contrast, repeated experiences due to habit may become boring from the lowered autonomy and thus lower continued enactment of the behavior. Previous intervention research has tested variety in exercise process itself such as variation of type of aerobic exercises (Glaros & Janelle, 2001), equipment use (Juvancic-Heltzel, Glickman, & Barkley, 2013), or providing messages in a spin class (Dimmock et al., 2013). These trials found groups that received the variety intervention demonstrated positive outcomes which included greater enjoyment (Dimmock et al., 2013), longer duration of MVPA (Juvancic-Heltzel et al., 2013) and exercise adherence (Glaros & Janelle, 2001) compared with their respective control groups.

Currently no research has examined the effectiveness of incorporating scheduling variety in predicting exercise. Though the time variation may refresh stimulation towards exercise, the preparatory context (cues and consistency) would also vary which could possibly compromise the opportunity to perform the behavior by making it more complex (Rhodes, Blanchard, & Matheson, 2006; Rhodes & De Bruijn, 2013). Similarly, intervention in the habit literature is nonexistent.

The primary purpose of this trial was to test if participants in the habit group engaged in more MVPA compared to those in the control and variety group. Given that scheduling variety could become complicated over time, it was hypothesized that the habit group would achieve greater MVPA minutes and stronger preparatory habit compared with the other two groups.
The secondary purpose was to test if the three groups would demonstrate changes in preparatory habit. It was hypothesized that the habit group would demonstrate stronger preparatory habit scores than control and variety groups. Finally, the tertiary purpose was to test if participants in the habit group would demonstrate greater use of cues and consistency across time compared to the variety and control groups. Any significant changes between groups from the secondary hypotheses would be further tested to show if habit mediated behavior change while controlling for the type of group. If habit demonstrated at least partial mediation in change then it was also hypothesized that the contextual variables (cues and consistency) would mediate the relationship between group and preparatory habit.

Methods

Trial Design

The intervention used a three-arm, parallel design randomized controlled trial (RCT). Participants were randomised to one of three groups: 1) Habit, 2) Variety, or 3) Control condition. The primary endpoint of the study was set at week eight of the trial.

Participants

Participants were adults (over 18 years of age) who did not meet the recommended physical activity guidelines at the time of recruitment (CSEP, 2011; Garber et al., 2011). Participants were screened via the Physical Activity Readiness Questionnaire and followed the protocol established by the PAR-Q+ Collaboration (Warburton, Bredin., Jamnik., & Gledhill., 2011). Individuals who were not ready to participate in moderate intensity PA were excluded for safety reasons. A power analysis using G Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009)
found that a repeated measures test to detect a medium-small effect size $F^2$ of .20 with the alpha error probability set at .05 and Power (1 - $\beta$ error probability) adjusted to .80 would require 165 participants for identifying all potential significant effects.

**Study Settings**

The intervention workshops were delivered at the University of Victoria, Gordon Head and Westshore Recreation centres. These locations were also used for meeting with participants to distribute and collect accelerometers. All self-report data was collected via online surveys.

**Procedure**

Recruitment was conducted by both active (on-site) and passive (notice of research poster). The primary investigator (PI) recruited on-site at recreation and community centres. Passive recruitment techniques included placing posters by high traffic areas such as the change rooms and water fountains. A brief announcement regarding the study was also published at the Times Colonist newspaper. Data was collected for both self-report (online questionnaires) and objective (accelerometer) measures. Online questionnaires were distributed at baseline (pre-randomization), week four and week eight. Accelerometry data was collected at baseline and the beginning of the eighth week. Participants in the control group were instructed to simply complete the measures across the eight weeks.

After baseline assessments, participants were randomized into one of three groups: habit, variety or control. Those who were randomized to an intervention group received an e-mail invitation to attend a workshop where they followed a presentation and completed the exercise specific worksheet. The presentation for each group was designed based on the group-theme followed by worksheet instructions. The message in the habit group emphasized the importance
of consistency and the usage of cues for preparation. In contrast, the variety group was encouraged to add variability to their exercise times. Both presentations began with an introduction on the benefits of exercise and concluded with how to make their exercise performance more enjoyable.

After the presentation, participants were provided with instructions and given time to complete the worksheets over carbon copy paper. The worksheets were confirmed prior to their departure and the PI kept a carbon copy version of their plan which was later used for the phone call phone-call follow-up. Following the meeting, the participants completed an online survey that allowed them to indicate suitable times for the four-week phone call follow-up. The purpose of the phone-call follow-up was to act as a prompt for participants to adhere to their plans, review their worksheets, and help them address any questions.

**Similarity of Interventions**

The intervention was a combination of individualized and group mediated meetings. Group-mediated trials help develop group formation, identity, and establish exercise and adherence expectancies (Brawley, Rajeski & Lutes, 2000). Overall, group-mediated interventions have been shown to be an effective method for convenience of message delivery and behavior outcomes (Brawley, Rejeski, Gaukstern, & Ambrosius, 2012; Cramp & Brawley, 2009; Mailey & McAuley, 2014). A mixed approach which includes both individual and group sessions has also demonstrated to be effective(Brawley et al., 2012). The use of worksheets/workbooks (Gortmaker et al., 1999) and telephone-assisted counselling (Estabrooks et al., 2009) have also been shown to be effective elements in behavior change and retention.
Intervention

Habit formation group.

The purpose of this group was to help participants establish a preparatory exercise habit by using a Dual-Process approach and applying the habit formation model proposed by Lally and Gardner (2013). The preparatory phase was defined as the phase preceding the exercise performance. Some examples include gathering workout gear (e.g., gym clothes, water bottle, mp3 player), or changing clothes at home. After outlining the health benefits of exercising regularly, participants were then presented with the intention-behavior gap puzzle (Rhodes & De Bruijn, 2013). A brief overview was provided on the importance of intention as the most proximal determinant to behavior, yet best efforts to increase intention only had a marginal effect on behavior change (Rhodes & Dickau, 2012). This finding suggests that there could be other constructs that function within that gap. A key construct that could bridge intention and behavior is habit (Rhodes & De Bruijn, 2013), which is defined as “a learned sequence of acts that have become automatic responses to specific cues, and are functional in obtaining certain goals or end-states” (Verplanken & Aarts, 1999, p. 104). Participants were then presented with the behavioral requirements to establish a habit which consisted of performing at least four exercise sessions/week for six weeks (Kaushal & Rhodes, 2015) followed by key psychological constructs to consider during the preparatory phase (cues and consistency) (Lally & Gardner, 2013). The primary investigator (PI) first discussed the importance of maintaining a consistent preparation practice. For instance, behavioral consistency is important for both establishing a particular exercise time (Kaushal & Rhodes, 2015) and for stabilizing the preceding events that lead to the exercise preparatory phase. Following this explanation, participants were instructed to
schedule the same time of day for their workouts. Examples included “after work at 5:30 pm” or “before work at 7:00 am”.

The instructions for using cues were based on an exploratory approach proposed by the PI. This approach suggested that cues work best when they are activated to trigger the behavior, such as traffic lights or an alarm clock. Following this rationale, a suggested practice was provided. The first component highlighted the appropriate usage - the cue should be switched on and off depending on if the behavior has been performed. The second focused on properties - the cue should be specific to the activity and salient in the environment. The following example was provided to illustrate this application: During the morning you can select your favourite gym clothes from your closet and place them on your bed before you leave for work. When you return home, the clothes remain on the bed and will continue to cue you until you use them for your workout. After returning from your workout, it is critical to turn off the cue by placing the clothes in your closet. The purpose of turning off the cue is to preserve the saliency and ensure that the visible presence of the cue does not become associated with other activities. This ritual can be applied using different objects such as running shoes or a water bottle.

Finally the presentation concluded with the incorporation of reward to facilitate a positive affective judgement while exercising. The following examples were provided as suggestions for adding enjoyment while exercising: bringing an mp3 player, using gloves for better grip/reduce friction, selecting favourite machines, watching television while running on the treadmill, purchasing new gear (shoes, clothes, workout gloves) and appropriate level of intensity.
Schedule variety group.

The focus of this group was to emphasize how adding variety to one’s exercise schedule (different times and duration) could provide greater enjoyment, which in turn may lead to better adherence. Similar to the habit presentation, the participants were introduced with the benefits of exercise followed by the intention-behavior gap. However, it was explained that the lack of a rewarding feeling could be the missing link in the gap. The shortage of rewarding feeling or enjoyment could be from intrinsic motivation (Deci & Ryan, 1985). It was suggested that a method to increase intrinsic motivation in an exercise setting is incorporating variety in the scheduling practice (Dimmock et al., 2013).

The participants were instructed to incorporate variety by adding variation and flexibility into their exercise schedule. It was suggested that adding variety via their own flexibility allows them to use their autonomy which helps increase intrinsic motivation (Deci & Ryan, 1985). The following example was provided to illustrate this concept: some people have a low level autonomy of their job schedule such as being present on particular days, at a particular time and for a fixed duration. Applying this approach to exercise could make working out analogous to one’s job. Second, activities performed consistently can become monotonous (e.g. your morning routine, driving to work, cleaning. Examples of incorporating a variety schedule were given by exercising in the morning, during lunch break, after work, before bed, and randomizing the pattern. The following instructions were provided when creating a variety schedule: i) feel free to include any number of bouts but each bout should be at least ten minutes in duration, ii) select any days of your choice but ensure that the total bout time in a week amount to at least 150 minutes. Finally the presentation concluded with how to make exercise sessions enjoyable. This section was identical to the habit intervention group.
Outcomes

The primary outcome of this intervention included the comparison of behavior change between the habit and control groups at the primary end-point (week eight). The secondary outcome was to comparing the change of preparatory habit between the three groups. The tertiary outcome was to test if the habit group successfully incorporated the contextual variables from the intervention. A mediation test would be conducted for each of these outcomes if the trajectory analyses showed differences between the groups.

Outcomes Measures

Primary and Secondary Outcomes

Exercise Behavior - MVPA

MVPA was measured objectively at baseline and at week eight using GT3X+ Actigraph Activity Monitors (Abel et al., 2008). The accelerometers were attached to elastic belts and participants were instructed to wear the belt and position the device over their left hip. They were told to wear the accelerometers for seven consecutive days (five weekdays and two weekend days) from when they woke up in the morning until they went to bed. They were further instructed to wear the device for at least twelve hours/day and complete daily activity log that identified when the accelerometer was removed or exceptional circumstances (e.g. swimming, forgot to wear). Valid wear time data was followed with Troiano’s suggestions (2007), additional validation options included at least five complete days with ten valid hours following previous research with similar sample (Sallis et al., 2009). MVPA was calculated by frequency bout which is defined by MVPA performed for a minimum of 10 minutes (1 952 average
acceleration counts/minute) (Freedson, Melanson, & Sirard, 1998; Trost, Loprinzi, Moore, & Pfeiffer, 2011).

Self-report MVPA was measured by administering the Godin Leisure Time Exercise Questionnaire (Godin, Jobin, et al., 1986b). The questionnaire consisted of three open-ended questions of time and frequency spent on type of physical activity (mild, moderate and strenuous). The 2-week test–retest reliability of the measures of total physical activity and the frequency of activity have been estimated to be 0.74 and 0.80, respectively (Godin, Shephard, et al., 1986). For the purpose of this study, only moderate and strenuous values were used to calculate the total time of MVPA. These categories reflect the definition of MVPA provided by recommended guidelines (CSEP, 2011; Garber et al., 2011). MVPA bouts for each participant was performed by dividing the total moderate time by 30 and vigorous time by 20 as per recommendation from the American College of Sport Medicine (ACSM, 2015).

**Habit- Preparation**

The Self-Report Behavioral Automaticity Index (SRBAI) was used to assess preparatory habit (Gardner et al., 2012b). The SRBAI consists of 4 items on a 5-point Likert scale with 1 being strongly disagree to 5 being strongly agree. The question stem stated “When I prepare to exercise…” which was then followed by four items on the scale: “I do it without having to consciously remember”, “I do it automatically”, “I do it without thinking”, and “I start before I realize I am doing it”. The present study found high Cronbach alphas for preparatory habit throughout the study: baseline ($\alpha=.90$), week 4 ($\alpha=.91$), week 8 ($\alpha=.94$).
Tertiary Outcomes

*Environmental Cues*

Specific to the instructions in the habit workshop, the following item was used to assess if participants implemented the use of cues: “I use cues at home to remind me to exercise (e.g. placing a water bottle on my desk or gym clothes on the bed). The item was rated on a five point Likert scale which ranged from “1=strongly disagree” to “5= strongly agree”.

*Consistency*

Temporal consistency was measured by using an item worded “How consistently do you exercise at the same time each day? (e.g., exercising every morning at 7 am, or exercising daily after supper)” (Kaushal & Rhodes, 2015). The question was asked on a 5-point Likert scale which range from “1=not consistent at all” to “5=very consistent”.

*Randomization and Blinding*

A randomization procedure was used in Microsoft Excel to allocate participants to one of three groups. After the completion of baseline measures, participants were informed if they were either in the intervention or the control group. Those in an intervention group received a link to the survey to select a time and location to attend a presentation. All participants were only informed they would receive either an intervention or control condition and were not aware that there were two different intervention groups.
Statistical Methods

Missing Outcomes

Multiple imputations (MI) was used to handle missing data (Enders, 2011; Karahalios, Baglietto, Carlin, English, & Simpson, 2012). MI has been used for handling missing data in physical activity RCTs (Chalder et al., 2012; Jeffery et al., 2009; Shapiro et al., 2012) and has been recommended for addressing missing accelerometer data (Catellier et al., 2005). Prior to conducting MI, Little’s Missing Completely at Random (MCAR) test was used to determine if the missing data was not significantly correlated with the dependent variables and descriptive data (age, BMI, sex, marital status, income and education).

Non-adherence to the Protocol

Non-adherence to the protocol was recorded by identifying the number of participants who were assigned to an intervention group but did not attend and those who could not be reached during the phone call follow-up.

Additional Analysis: Outcomes and Estimations

All hypotheses were tested with their respective trajectory analyses. If a difference was found between groups then those constructs would be tested in a mediation model (Hayes, 2009, 2013). Mediation analyses was performed by using the Process macro for SPSS (Hayes, 2013).

Primary and Secondary Objectives

Basic descriptives of the sample and bivariate correlations were computed for the study constructs. The primary and secondary hypotheses were investigated by testing if the habit group
demonstrated a significant increase in MVPA (accelerometry and self-report data) and preparatory habit across the intervention when compared with the control group. This was tested by conducting a Repeated Measures (RM) ANCOVA by using the baseline as a covariate and follow-up measures at weeks four and eight as the repeated observations (Norman & Streiner, 2008). The difference in these constructs were further tested to investigate if preparatory habit mediated the change of MVPA (accelerometry and self-report) between the habit and control groups. All mediation models were conducted based on change from baseline to week four and baseline to week eight. The time difference of the constructs were calculated by using residual change values (Preacher & Kelley, 2011) which is a well-documented approach (Buffart et al., 2014; Plotnikoff, Pickering, Flaman, & Spence, 2010; Rogers et al., 2014; Vallance, Courneya, Plotnikoff, & Mackey, 2008). When reporting the results, key findings were highlighted if they demonstrated statistical significance (p<.05) or a meaningful effect size: \( d=.41, \beta=.20, \text{OR}=2.0 \) (Ferguson, 2009). Finally, logistic regression was conducted by using accelerometry data to determine the odds ratio of the habit group achieving the guidelines compared with the control and variety groups.

**Tertiary Objectives**

The manipulation check was performed by testing if preparatory cues and consistency significantly changed across time when compared with the habit, control and variety groups. A MANCOVA procedure was used to examine the change of these variables across time by setting the group variable as a fixed factor, placing the baseline values as covariates (cues and consistency) and their respective follow-up measures (week four and eight) as dependent variables (Norman & Streiner, 2008). This method determined the overall change of these
constructs for each group by controlling for their baseline values. These constructs were then tested in a mediation model to determine if the change of cues and consistency mediated the change of preparatory habit between the two groups.

**Ethical Aspects**

Ethical approval for *Understanding Exercise Adherence* was obtained from the Human Research Ethics Board at the University of Victoria. Written informed consent was obtained from all participants during recruitment.

**Results**

**Participant flow**

One hundred seventy six adults showed interest in the study by contacting the primary investigator. Twenty eight either declined to participate (n=16) or were excluded as per the eligibility criteria (n=12). Participants (n=122) were randomized into either variety (n=37), habit (n=41) or control (n=44) group. A total of 13 participants dropped-out from the study which included seven from the control and three each from the habit and variety groups (Figure 1).

**Missing Outcomes**

Missing analysis found 12.9% of the data to be incomplete. Further investigation using Little’s MCAR test revealed that the data was missing completely at random ($\chi^2=90.05, \text{DF}=76, p=.13$), hence it was appropriate to proceed with multiple imputations (Rubin, 1996; Schafer & Graham, 2002).
Non-Adherence to Protocol

Four participants could not be reached for the phone call follow-up (two from each intervention group). Upon recruitment, participants were only accepted if they were not meeting the Canadian Physical Activity Guidelines (Garber et al., 2011; Warburton et al., 2007). In-between the recruitment phase and launch of the study, 24% of the sample began exercising and reported meeting the Canadian Physical Activity Guidelines on the baseline survey. However, chi-squared test found that the groups did not differ in meeting PA guidelines at baseline ($\chi^2=.40$, DF= 2, p=.82). Given that baseline data revealed homogeneity across groups, this was not a considered a validity threat to the intervention.

Baseline data

The mean age of participants was 40.51 ($SD = 15.3$ years), 77% were female and the BMI of the participants was 27.9 ($SD=5.9$) suggesting an overweight sample (NIH, 2011). Participants were averaging 108 minutes ($SD=84.49$) of MVPA per week at baseline with 76% of the sample not meeting the physical activity guidelines (Garber et al., 2011; Warburton et al., 2007). Chi-squared test found that the groups did not differ in meeting PA guidelines at baseline ($\chi^2=.40$, DF= 2, p=.82) and were homogenous in their intention to achieve the guidelines ($\chi^2=.43$, DF= 2, p=.71). Further descriptives are presented in Table 1. Of the 122 participants enrolled in the study, 107 completed the study, representing an 88% retention rate. Of those who were randomized into an intervention group (n=72), 92% of the participants completed the study; these values land within the range of a strong trial (80-100% retention) (Jackson & Waters, 2005).
**Bivariate Correlations**

Correlations for the sample can be found in Table 2. Baseline accelerometer MVPA correlated with the follow-up measure on week 8 (r=.34, p<.01). Follow-up accelerometry correlated with intention (.17, p<.01), habit (.18, p<.05), consistency (.19, p<.05), cues (.28, p<.01), complexity (.18, p<.05), behavioral-regulation (.22, p<.05), and self-report MVPA (.33, p<.01). The strength of habit with behavior (self-report: .33, p<.01, accelerometer: .18, p<.05), consistency (.33, p<.01) and cues (r=.40, p<.01) were significant by week eight.

**H1. Comparing Behavior Change Between Groups**

At week four (Table 3a), neither the habit or variety were significantly different from the control group in self-reported MVPA F (2,117) = 3.56, p > .05 (η²=.03). However, by week eight (Table 5b), the habit group had larger changes in MVPA from baseline for both accelerometry [F (2,117) = 2.67, p = 0.07 (η²=.05)] and self-report [F (2,117) = 3.64, p < .05 (η²=.05)]. Accelerometry found the habit group to be significantly greater than the control [mean change=+35.8; 95% CI= (34.6, 41.0); d=+.40; p<.05] but not variety [mean change=+35.6; 95% CI= 33.8, 37.2); d=+.36; p>.05] group. Whereas self-report found the habit group to be significantly greater than control [mean change=+.59; 95% CI= (.43, .55); d=+.41; p<.05] and variety [mean change=33; 95% CI= (.31, .35); d=+.26; p<.05] groups. Binary logistic regression further showed that those in the habit group were 2.85 times [OR: 2.85, 95% CI (1.48, 6.45), p<.05] more likely than the control and 2.63 times [OR: 2.63, 95% CI (1.45, 5.67), p<.05] more likely than the variety group to achieve PA guidelines at week eight.
H2a. Comparing Preparatory Habit Change Between Groups

The ANCOVA on preparation habit (Table 3a) did not reveal noticeable differences at week four (d=.06-.14; when comparing all three group combinations). However, those from the habit workshop reported greater change in preparatory habit by week eight $F(2,117) = 3.30, p < 0.05 (\eta^2=.05)$ over control [mean change=+.49; 95% CI= (.43, .55); d=+.41; p<.05] but not significantly greater than the variety [mean change=+.33; 95% CI= (.35, .31); d=+.26; p>.05] group (Table 3b).

H3a. Comparing Automaticity Predictors Between Groups

A MANCOVA test which examined the change in use of cues revealed a significant treatment effect with the habit group using more cues $F(2,117) = 6.02, p < 0.01 (\eta^2=.09)$ and consistency $F(2,117) = 5.79, p < 0.01 (\eta^2=.08)$ than the variety and control groups at four weeks (Table 3a). Cues continued to show significant differences at eight weeks $F(2,117) = 5.96, p < 0.01 (\eta^2=.08)$ with pairwise comparisons revealing cues were higher in the habit group when compared with the variety [mean change=+.64; 95% CI= (.66, .61); d=+.56; p<.001] and control [mean change=+.76; 95% CI= (.69 to .83); d=+.58; p<.001] groups (Table 3b). The test for consistency was also significant at eight weeks with participants in the habit group reporting stronger consistency over the other two groups, $F(2,117) = 3.27, p < 0.05 (\eta^2=.05)$. Specific comparisons revealed participants in the habit group demonstrated the greatest change of incorporating a consistent routine over the control [mean change=+.67; 95% CI= (.59 to .75); d=+.38; p<.01] and variety groups [mean change=+.49; 95% CI= (.52, .46); d=+.47; p<.05].
**H2b: Habit Mediating Group and MVPA (week eight)**

The first mediation analysis tested if habit mediated objective MVPA (Figure 2 and Table 5b). The mediation analysis found that group type (habit vs. control) significantly predicted change in habit from baseline to week eight (path a = .46, p<.05) but change in habit did not predict change in accelerometry MVPA. The direct effect of group type to MVPA remained significant, (path c=.40, p<.05). Hence, habit was not found to mediate the effect of group change in accelerometry MVPA.

The second mediation analysis tested the same model but investigated if habit mediated self-report MVPA (Figure 4 and Table 5b). Similar to the previous model, group type (habit vs. control) continued to significantly predict change in habit from baseline to week eight (path a = .46, p<.05), however the change in habit subsequently predicted an increase of MVPA (path b=.24, p<.01). A bias-corrected bootstrap confidence interval for the indirect effect of group predicting MVPA (path ab=.11) based on 10 000 boot strap samples was above zero (.034 to .814). However the direct effect of group type to MVPA remained significant, path c= .43, (p<.05). Thus, habit only partially mediated the effect of group on change in MVPA.

**H3b: Cues and Consistency Mediation analyses for habit vs. control group**

The first mediation analysis tested the development of habit via the incorporation of cues and development of consistency from baseline to week four (Figure 5, Table 5a). Group type (eg. habit vs. control) predicted change in the use of cues (path a1=.68, p<.01), and changes in consistency (path a2=.62, p<.01) which is commensurate with the trajectory ANCOVA results. However only changes in consistency (b2=.33, p<.01) was found to predict the development of habit (cues b2=.06, p>.05). Group type was not found to have a direct effect on habit.
development (path c=.11, p>.05). A bootstrap confidence interval for the indirect effects of intervention group (ab=.23) using 10 000 bootstrap samples was greater than zero (ab=.070 to .579). The insignificance of the direct patch (path c) in addition to the CI range existing above zero suggests total mediation.

When the same model was tested for change in habit at eight weeks (Figure 6, Table 6b), group type continued to predict change in the use of cues (path a1=.66, p<.01), and change in consistency (path a2=.57, p<.01). Changes in cues (path b1=.32, p<.01) but not consistency (path b2=.11, p>.05) predicted change in habit. Group type was not found to have a direct effect on habit development (path c=.36, p>.05). A bootstrap confidence interval for the indirect effects of group type on change in habit through change in cues and consistency (ab=.22) using 10 000 bootstrap samples was greater than zero (ab=.069 to .537). The requirements were met to indicate that total mediation was achieved.

Harms

No participants reported any experiences of harm related to the study.

Discussion

The primary purpose of this study was to test if participants in the habit group demonstrated greater behavior change compared with the control and variety groups. Habit formation has been suggested as a taxonomy for behavior change (Michie et al., 2013); however this is the RCT to test the application of a habit-based intervention in new gym members and address reoccurring suggestions to test habit as a mechanism for PA behavior change (Gardner,
The secondary and tertiary purposes were to investigate if the habit group demonstrated greater preparatory habit and contextual variables (cues and consistency) compared with the control and variety groups. If the participants in the habit group differed in these constructs compared with the other two groups then a mediation analysis was performed to determine if the change in preparatory habit mediated behavior change between the habit and the control or variety group. If habit partially mediated group and behavior then the contextual variables would be tested as mediators between group and preparatory habit.

H1. Exercise

It was hypothesized that participants in the habit group would increase in MVPA compared to the control and variety groups. The hypothesis was partially supported. Participants in the habit group showed greater mean increases in MVPA via accelerometry compared to control participants (p<.05, d=.40) but not those in the variety group (p>.05; d =.36) at week eight. However, self-reported MVPA was significantly higher for participants in the habit condition compared to control (p<.05, d=.51) and variety group participants (p<.05; d =.50). Although the variety group participants were not significantly different than participants in the habit group when measured with accelerometry, the magnitude was close to a clinically meaningful effect size (d=.36) (Ferguson, 2009) whereas the difference between the variety and control was negligible (d=.10). Finally, when odds ratios were calculated, the habit group was found to be 2.85 and 2.63 times more likely to achieve MVPA guidelines by week eight compared with the control and variety groups, respectively.
Clearly, the results of this intervention demonstrate the superiority of developing cues and context-repetition for behavior change. Although previous interventions have shown support for exercise variety approach (Dimmock et al., 2013; Glaros & Janelle, 2001; Juvancic-Heltzel et al., 2013), it’s important to note that these trials focused on adding variation to the task itself rather than the schedule. The present study tested the incorporation of scheduling variety; consequently, this approach was not found to be effective as reflected from the results and thought listing feedback.

H2. Habit

It was also hypothesized that the habit group would demonstrate a stronger preparatory habit compared with the control and variety groups. The present study found that the habit group was shown to have the greatest change in preparatory habit from baseline to week eight over the control (p<.05, d=.41) but not significantly greater than the variety group (p>.05, d=.26). Given that participants in the habit workshop were provided with instructions on how plan their exercise sessions by using consistency and cues, the findings between the habit and control group is clearer to interpret. Although the variety group were inconsistent in time and did not incorporate reminder cues, the familiarity of using objects in their daily environments may have remained the same. For instance, a participant may have used his home and office as preparatory environments before going to the gym. Though it is likely that habitual behavior may not develop as fast as one preparatory environment, the familiarity of objects in two environments could develop overtime, thus not resulting in a significant difference between the two groups in habit preparation scores.
**H3a. Automaticity Antecedents**

It was hypothesized that the habit group would demonstrate the greatest change in applying the manipulated constructs (cues and consistency) compared with the control and variety groups at week eight. The manipulation check revealed that the use of cues (d=.56 to .58) and consistency (d=.38 to .47) were significantly higher among habit participants compared with the other two groups. This is the first physical activity intervention that instructed individuals to manipulate the antecedent constructs that have been theorized to facilitate habit formation (Lally & Gardner, 2013; Wood & Rünger, 2015). It is interesting to note that the instructional investment required for participants to successfully incorporate these constructs (one presentation followed by a worksheet) was rather minimal, thus reflecting the relative ease of independent application.

Overall, these results could support the conjecture that the consistency of performing a planned behavior was less of a hassle than handling a variable exercise schedule on a regular basis. If an individual desires to incorporate flexibility in his/her exercise schedule, then it may likely require the flexibility to shift other responsibilities as well. This could particularly be challenging for adults whose schedule consists of several non-mutable responsibilities (e.g. occupation, pick-up drop off children, elderly care, etc). On the other hand, a habit requires a fixed, expected, timeslot that only necessitates a onetime negotiation of what activity to sacrifice for its place, or at least less juggling of one’s schedule. The purpose of establishing an exercise habit is to reduce the cognitive burden of performing the behavior by delegating some control to familiar contextual cues (Gardner, 2014; Wood & Rünger, 2015). Consequently, constant schedule variation could make this a challenge. Still, implementing a scheduling variety tactic
could be viable for individuals who have less consistent schedules such as those who have rotating shiftwork. With these populations, it would be interesting to test if habit could be formed from a pattern routine rather than temporal consistency. For instance, pairing a new health behavior after an existing health habit helps facilitate the new behavior into a habit (Judah, Gardner, & Aunger, 2012).

**H2b. Habit Mediating Group and Behavior**

When habit was tested as a mediator between group assignment and accelerometry MVPA, it was not found to mediate MVPA at week eight. Interestingly, the self-reported MVPA model found that habit did not mediate behavior at week four but showed partial mediation by week eight. This finding supports previous research which found habit formation to stabilize at six weeks (Kaushal & Rhodes, 2015). Given that the participants were conscious intenders and their behavior was partially mediated by unconscious drive (habit), this method supports the Dual Process approach. These results further add support to previous Dual process findings (Conroy et al., 2010; Kaushal & Rhodes, 2015; Rhodes & De Bruijn, 2010; Rhodes et al., 2010; Verplanken et al., 1997) and matches an early definition of habit as a goal-dependent automaticity (Bargh 1989; Bargh 1990) and a construct that develops from action control (Rhodes & De Bruijn, 2013).

When compared to other PA interventions, the majority of mediation tests have only used self-reported data and have generally resulted in null findings (Rhodes & Pfaeffli, 2010). Although the present study did not find habit to mediate changes in accelerometry MVPA, habit was found to partially mediate self-reported behavior. The partial mediation could have resulted
from the sample being slightly underpowered (87% of recommended sample size was achieved). The discrepancy between the self-report and accelerometry measures could be attributed to the timing of the baseline assessment. Upon launching the study, participants received a link to complete the baseline survey online and were invited to retrieve their accelerometer. The devices were picked-up the same week and began recording data on the following Monday. Given that the self-report questionnaires instructed participants to reflect behavior and psychological constructs from the previous two weeks, this creates a potential three week gap of measurement comparison. However, it’s important to note that the self-report and accelerometry data displayed similar patterns between the three groups across time. Overall, low agreement with these measures are not uncommon as a review has documented the challenges and discrepancy between accelerometry and self-report data (Prince et al., 2008). In particular, Prince and colleagues (2008) found a trend where the divergence of these methods becomes magnified when comparing data with greater intensities of PA (e.g., moderate-vigorous level).

**H3b. Contextual Predictors as Mediators of Habit**

It was interesting to find that consistency mediated change of habit at week four but cues emerged as a stronger predictor by week eight. When the two mediation models are considered together, the findings suggest that consistency can be developed faster but the effectiveness of cues requires longer time in order to establish the stimulus-response association. Hence, these constructs could function in sequence rather than a parallel system. A similar system has been proposed for habit; habit is thought to have a reciprocal relationship with behavior (Gardner, 2014), where habit affects behavioral repetition but that repetition also strengthens habit
formation. Cues have been theorized as a predictor of habit in various models (Grove & Zillich, 2003; Lally & Gardner, 2013; B. Verplanken et al., 1997), yet valid assessment of this construct has remained a notable issue (Gardner & Tang, 2013). The present study instructed participants to follow two rules when implementing cues. The first was on appropriate usage- the cue should be switched on and off depending on if the behavior needs to be performed, and the second focused on properties- the cue should be specific to the activity and salient in the environment. This is also the first habit study to treat cue in a daily manipulation manner to maintain its saliency effect. Overall, the current findings support the contextual stability theorizing for habit formation (Lally & Gardner, 2013; Wood & Rünger, 2015) and the use of the two rules for implementing effective cues.

Limitations

Given that the study was eight weeks in length, these findings should be interpreted as short-term behavior change. The purpose of this study was to test the establishment of exercise habit in new gym members who were not meeting the PA guidelines. It’s important to note that all participants reported themselves as new exercisers who were not meeting the PA guidelines upon enrolment of the study. However, behavior from the enrolled participants was not controlled while the PI continued recruiting new participants. Consequently, some participants (24%) reported that they were achieving the PA guidelines on the baseline survey.

Interpretation

The dropouts of gym and recreation centres particularly during the New Year phase is well known, yet the lack of research investigating this issue is rather alarming (Kaushal & Rhodes, 2015). This is the first intervention to focus on establishing a habit to facilitate regular
exercise behavior. The present study adds to previous finding on the importance of consistency in establishing an exercise habit (Kaushal & Rhodes, 2015) and further provides a novel interpretation of utilizing and the assessment of cues. Specifically, those who incorporated consistency and cues to facilitate habit formation were at least 2.63 times more likely to meet the PA guidelines by week eight when compared with participants who did not focus on maintaining a stable preparatory context. Overall, the present findings support the utility of a habit building workshop on short-term PA change. Gym trainers should provide a plan sheet for their clients that help them plan their exercise sessions with particular emphasis on consistency and cues; a follow-up with clients could be helpful in addressing their questions and helping them modify their plan if necessary. Future research is needed with greater sample size to replicate these findings and extend the duration of assessment times to evaluate whether PA changes are sustained across time.


### Tables and Figures

Table 1. Descriptive Data.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>$F$ or $\chi^2$</th>
<th>$p$</th>
<th>Control Mean (SD) n=44</th>
<th>Habit Mean (SD) n=41</th>
<th>Variety Mean (SD) n=36</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Profile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>$F=.59$</td>
<td>.56</td>
<td>40.30 (14.69)</td>
<td>38.21 (13.98)</td>
<td>41.90 (14.38)</td>
</tr>
<tr>
<td>BMI</td>
<td>$F=.66$</td>
<td>.51</td>
<td>28.48 (6.57)</td>
<td>27.01 (5.30)</td>
<td>27.69 (5.53)</td>
</tr>
<tr>
<td>% Female</td>
<td></td>
<td></td>
<td>81</td>
<td>83</td>
<td>74</td>
</tr>
<tr>
<td>% Completed University</td>
<td></td>
<td></td>
<td>56.3</td>
<td>58.5</td>
<td>47</td>
</tr>
<tr>
<td>% Married/Common-law</td>
<td></td>
<td></td>
<td>46.9</td>
<td>46.3</td>
<td>34.3</td>
</tr>
<tr>
<td>% $75 000 Income</td>
<td></td>
<td></td>
<td>27.5</td>
<td>31.7</td>
<td>21.9</td>
</tr>
<tr>
<td>% Currently Employed</td>
<td></td>
<td></td>
<td>39.1</td>
<td>48.8</td>
<td>45.7</td>
</tr>
<tr>
<td><strong>Health Indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Health Appraisal</td>
<td></td>
<td></td>
<td>2.98 (.98)</td>
<td>2.76 (.91)</td>
<td>2.69 (.74)</td>
</tr>
<tr>
<td>% Heart disease</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>% Cancer</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>% Diabetes</td>
<td></td>
<td></td>
<td>10.9</td>
<td>9.8</td>
<td>11.4</td>
</tr>
<tr>
<td>% High Blood pressure</td>
<td></td>
<td></td>
<td>10.9</td>
<td>12.2</td>
<td>8.6</td>
</tr>
<tr>
<td>% High Cholesterol</td>
<td></td>
<td></td>
<td>4.7</td>
<td>2.4</td>
<td>14.3</td>
</tr>
<tr>
<td><strong>Baseline Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accelerometer MVPA</td>
<td>$F=.14$</td>
<td>.87</td>
<td>131.46 (81.25)</td>
<td>128.80 (91.31)</td>
<td>123.35 (69.21)</td>
</tr>
<tr>
<td>MVPA Guidelines %</td>
<td>$\chi^2=.40$</td>
<td>.82</td>
<td>26.7</td>
<td>22.5</td>
<td>25.0</td>
</tr>
<tr>
<td>S.R. MVPA Bouts</td>
<td>$F=.06$</td>
<td>.94</td>
<td>4.47 (5.98)</td>
<td>4.17 (3.22)</td>
<td>4.44 (4.21)</td>
</tr>
<tr>
<td>S.R. MVPA Total</td>
<td>$F=.19$</td>
<td>.83</td>
<td>113.92 (87.10)</td>
<td>103.85 (87.88)</td>
<td>106.55 (72.48)</td>
</tr>
<tr>
<td><strong>Baseline Predictors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habit Preparation</td>
<td>Consistency</td>
<td>Cues</td>
<td>Intention %</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>F= .32</td>
<td>.75</td>
<td>2.28 (.98)</td>
<td>2.38 (.94)</td>
<td>2.25 (.88)</td>
<td></td>
</tr>
<tr>
<td>F= .55</td>
<td>.58</td>
<td>2.16 (1.39)</td>
<td>2.39 (1.47)</td>
<td>2.17 (1.36)</td>
<td></td>
</tr>
<tr>
<td>F= .97</td>
<td>.39</td>
<td>1.72 (.91)</td>
<td>1.93 (.91)</td>
<td>2.00 (1.05)</td>
<td></td>
</tr>
<tr>
<td>$\chi^2= .43$</td>
<td>.71</td>
<td>86</td>
<td>90</td>
<td>97</td>
<td></td>
</tr>
</tbody>
</table>

*Note. Accelerometer MVPA = Freedson’s Bouts, Chi-squared tests were not significant for any of the categorical demographic constructs. Chi-squared tests are presented for MVPA Guidelines % and Intention as they pertain to the study protocol.*
Table 2. Habit Model: Bivariate Correlations

<table>
<thead>
<tr>
<th>Construct</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. T1. Acc.</td>
<td>.05</td>
<td>.09</td>
<td>.04</td>
<td>-.09</td>
<td>.01</td>
<td>.05</td>
<td>.10</td>
<td>.20</td>
<td>.34**</td>
<td>.02</td>
<td>.01</td>
<td>.09</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>2. T1. MVPA</td>
<td>.11</td>
<td>.14</td>
<td>.13</td>
<td>.39**</td>
<td>.13</td>
<td>.23*</td>
<td>.21*</td>
<td>.08</td>
<td>.42**</td>
<td>.25**</td>
<td>.33**</td>
<td>.24**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. T1. Habit</td>
<td>.18*</td>
<td>.20*</td>
<td>.09</td>
<td>.27*</td>
<td>.15</td>
<td>.18</td>
<td>.09</td>
<td>.20*</td>
<td>.25**</td>
<td>.15</td>
<td>.19*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. T1. Consist</td>
<td>.26**</td>
<td>.14</td>
<td>.16</td>
<td>.37**</td>
<td>.15</td>
<td>.15</td>
<td>.20*</td>
<td>.21*</td>
<td>.42**</td>
<td>.18*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. T1. Cues</td>
<td>.03</td>
<td>.13</td>
<td>.15</td>
<td>.26**</td>
<td>.13</td>
<td>.14</td>
<td>.13</td>
<td>.22*</td>
<td>.33**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. T2. MVPA</td>
<td>.20*</td>
<td>.13</td>
<td>.23**</td>
<td>.17</td>
<td>.44**</td>
<td>.16</td>
<td>.16</td>
<td>.19*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. T2. Habit</td>
<td>.38**</td>
<td>.25*</td>
<td>.07</td>
<td>.16</td>
<td>.50**</td>
<td>.22*</td>
<td>.26*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. T2. Consist</td>
<td>.38**</td>
<td>.15</td>
<td>.17*</td>
<td>.27**</td>
<td>.51**</td>
<td>.27**</td>
<td>.48**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. T2. Cues</td>
<td>.11</td>
<td>.27**</td>
<td>.29**</td>
<td>.27**</td>
<td>.48**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. T3. Acc.</td>
<td>.33**</td>
<td>.18*</td>
<td>.19*</td>
<td>.26**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. T3. MVPA</td>
<td>.33**</td>
<td>.34**</td>
<td>.28**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. T3. Habit</td>
<td>.36**</td>
<td>.40**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. T3. Consist</td>
<td>.42**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. T3. Cues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Acc= Accelerometer (Freedson's bouts), MVPA= moderate-to-vigorous-physical activity, Consist= consistency
Figure 1. Randomization Process

Assessed for eligibility (n=151)

Excluded (n=29)
- Not meeting inclusion criteria at time of enrolment (n=13)
- Declined to participate (n=16)

Randomized (n=122)

Allocated to Variety Group (n=37)
- Received allocated intervention (n=37)
  - Discontinued intervention: health, (n=1)
    personal, (n=1)
    did not specify (n=1)

Allocated to Habit Group (n=41)
- Received allocated intervention (n=41)
  - Discontinued intervention: "no time" (n=1)
    new job (n=1)
    did not specify (n=1)

Allocated to Control Group (n=44)
- Received allocated intervention (n=44)
  - Discontinued intervention (time, personal, did not want to be in the control group) (n=7)

Analysed (n=36)
- Excluded from analysis: unclear data provided, n=1

Analysed (n=41)

Analysed (n=44)
Note. Drop-out Chi Square Test between habit and control was non-significant: $\chi^2=2.1, \ p=.24$.

Table 3a. Primary Outcome: Behavior Change between Habit, Variety and Control Groups at 4 Weeks

<table>
<thead>
<tr>
<th>Constructs</th>
<th>G</th>
<th>Baseline Mean (SD)</th>
<th>4 Weeks Mean (SD)</th>
<th>F Values, $\eta^2$</th>
<th>Post Hoc</th>
<th>Cohen’s d H and C</th>
<th>Mean Change H and V [95% CI]</th>
<th>Cohen’s d V and C</th>
<th>Mean Change V and C [95% CI]</th>
<th>Cohen’s d V and C</th>
<th>Mean Change [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVPA Bouts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td>4.17 (3.22)</td>
<td>7.30 (4.56)</td>
<td>F=3.56</td>
<td>.39</td>
<td>1.68</td>
<td>[1.43, 1.93]</td>
<td>.44</td>
<td>1.26</td>
<td>.09</td>
<td>.42</td>
</tr>
<tr>
<td>V</td>
<td></td>
<td>4.44 (4.21)</td>
<td>6.11 (3.84)</td>
<td>P=.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>4.47 (5.98)</td>
<td>5.71 (4.70)</td>
<td>$\eta^2=.03$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVPA Total time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td>103.85 (87.88)</td>
<td>188.05 (99.64)</td>
<td>F=2.30</td>
<td>.43</td>
<td>21.19</td>
<td>[16.40, 25.99]</td>
<td>.44</td>
<td>37.80</td>
<td>.07</td>
<td>-16.60</td>
</tr>
<tr>
<td>V</td>
<td></td>
<td>106.55 (72.48)</td>
<td>151.21 (62.23)</td>
<td>P=.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>113.92 (87.10)</td>
<td>163.94 (84.50)</td>
<td>$\eta^2=.03$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVPA Guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.Odds Ratio</td>
<td></td>
<td>OR: 1.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. Chi Square</td>
<td></td>
<td>$\chi^2=.02, \ p=.97$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. MVPA Bouts= Self-report bouts from ACSM guidelines. Group code: H= habit, V= Variety, C= Control. Guidelines: participants were binary coded as 1 meeting and 0 not meeting guidelines. Significance of post-hoc tests was determined at $p<.05$. MVPA Guidelines- Odd Ratios and Chi Square tests were binary coded as 1. Habit group with 2. Control and Variety groups. DV was coded as either 1. < 150 minutes of MVPA or 2. $\geq$ 150 minutes of MVPA. OR= Odds Ratio, CI= Confidence Intervals.
Table 4b. Primary Outcome: Behavior Change between Habit, Variety and Control Groups at 8 Weeks

<table>
<thead>
<tr>
<th>Constructs</th>
<th>G</th>
<th>Baseline Mean (SD)</th>
<th>8 Weeks Mean (SD)</th>
<th>F Values, p</th>
<th>Post Hoc</th>
<th>Cohen’s d H and C</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d H and V</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d V and C</th>
<th>Mean Change [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVPA Objective</td>
<td>H</td>
<td>131.46 (81.25)</td>
<td>185.26 (112.00)</td>
<td>F=2.67 p=.07</td>
<td>H&gt;C</td>
<td>.40</td>
<td>35.84 [34.64, 37.24]</td>
<td>.36</td>
<td>35.52 [33.81]</td>
<td>.04</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>128.80 (91.31)</td>
<td>148.50 (76.10)</td>
<td>η²=.23</td>
<td>H&gt;V</td>
<td>.51</td>
<td>2.11 [1.86, 1.92]</td>
<td>.50</td>
<td>1.84 [1.75]</td>
<td>.06</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>123.35 (69.21)</td>
<td>146.35 (77.25)</td>
<td></td>
<td></td>
<td>.47</td>
<td>50.58 [47.86, 52.25]</td>
<td>.10</td>
<td>2.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVPA Bouts-self-report</td>
<td>H</td>
<td>4.17 (3.22)</td>
<td>7.36 (4.48)</td>
<td>F=3.56 p=.13</td>
<td>H&gt;C</td>
<td>.51</td>
<td>2.11 [1.86, 1.92]</td>
<td>.50</td>
<td>1.84 [1.75]</td>
<td>.06</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>4.44 (4.21)</td>
<td>5.58 (4.25)</td>
<td>η²=.03</td>
<td>H&gt;V</td>
<td>.38</td>
<td>52.80 [47.86, 52.25]</td>
<td>.47</td>
<td>50.58 [48.90]</td>
<td>.10</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.47 (5.98)</td>
<td>5.32 (4.00)</td>
<td></td>
<td></td>
<td>.38</td>
<td>52.80 [47.86, 52.25]</td>
<td>.47</td>
<td>50.58 [48.90]</td>
<td>.10</td>
<td>2.23</td>
</tr>
<tr>
<td>MVPA Total time</td>
<td>H</td>
<td>103.85 (87.88)</td>
<td>187.10 (91.73)</td>
<td>F=2.30 p=.53</td>
<td>H&gt;C</td>
<td>.38</td>
<td>52.80 [47.86, 52.25]</td>
<td>.47</td>
<td>50.58 [48.90]</td>
<td>.10</td>
<td>2.23</td>
</tr>
<tr>
<td>(self-report)</td>
<td>V</td>
<td>106.55 (72.48)</td>
<td>137.64 (89.45)</td>
<td>η²=.03</td>
<td>H&gt;V</td>
<td>.38</td>
<td>52.80 [47.86, 52.25]</td>
<td>.47</td>
<td>50.58 [48.90]</td>
<td>.10</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>113.92 (87.10)</td>
<td>130.93 (87.40)</td>
<td></td>
<td></td>
<td>.38</td>
<td>52.80 [47.86, 52.25]</td>
<td>.47</td>
<td>50.58 [48.90]</td>
<td>.10</td>
<td>2.23</td>
</tr>
<tr>
<td>MVPA Guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OR: 2.63</td>
<td>[1.23, 5.61]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR: 95% CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>i.Odds Ratio</td>
<td>[1.23, 5.61]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. Chi Square</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>χ²=6.49 p=.03</td>
<td>[1.23, 5.61]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note. MVPA Objective= Accelerometry (Freedson’s bouts). MVPA Bouts= Self-report bouts from ACSM guidelines. Group code: H= habit, V= Variety, C= Control. Guidelines: participants were binary coded as 1 meeting and 0 not meeting guidelines. Significance of post-hoc tests was determined at p<.05. MVPA Guidelines-Odd Ratios and Chi Square tests were binary coded as 1. Habit group with 2. Control and Variety groups. DV was coded as either 1. < 150 minutes of MVPA or 2. >= 150 minutes of MVPA. OR= Odds Ratio, CI= Confidence Intervals.

Table 5a: Habit Model- The Change of Habit and its antecedents between Habit, Variety and Control Groups at Week 4

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Baseline Mean (SD)</th>
<th>4 Weeks Mean (SD)</th>
<th>F Values, p, η2</th>
<th>Post Hoc</th>
<th>Cohen’s d H and C</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d H and V</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d V and C</th>
<th>Mean Change [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>2.38 (.94)</td>
<td>2.81 (1.00)</td>
<td>F=.41</td>
<td>.06</td>
<td>.14</td>
<td>[.07, .19]</td>
<td>.08</td>
<td>[.04, .00]</td>
<td>.14</td>
<td>[.03, .19]</td>
</tr>
<tr>
<td>V</td>
<td>2.25 (.88)</td>
<td>2.75 (.86)</td>
<td>p=.70</td>
<td>.19</td>
<td>.07</td>
<td>[.05, .16]</td>
<td>.04</td>
<td>[.02, .06]</td>
<td>.21</td>
<td>[.09, .13]</td>
</tr>
<tr>
<td>C</td>
<td>2.28 (.98)</td>
<td>2.65 (.97)</td>
<td>η2=.00</td>
<td>.01</td>
<td>.01</td>
<td>[.01, .02]</td>
<td>.00</td>
<td>[.00, .00]</td>
<td>.02</td>
<td>[.00, .00]</td>
</tr>
<tr>
<td>Consistency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>2.39 (1.47)</td>
<td>3.05 (1.35)</td>
<td>F=5.79</td>
<td>.45</td>
<td>.79</td>
<td>[.71, .78]</td>
<td>.10</td>
<td>[.24, .19]</td>
<td>.47</td>
<td>[.48, .68]</td>
</tr>
<tr>
<td>V</td>
<td>2.17 (1.36)</td>
<td>2.73 (1.26)</td>
<td>P=.01</td>
<td>.48</td>
<td>.26</td>
<td>[.23, .30]</td>
<td>.04</td>
<td>[.02, .06]</td>
<td>.34</td>
<td>[.31, .37]</td>
</tr>
<tr>
<td>C</td>
<td>2.16 (1.39)</td>
<td>2.19 (1.32)</td>
<td>η2=.08</td>
<td>.02</td>
<td>.02</td>
<td>[.01, .03]</td>
<td>.00</td>
<td>[.00, .00]</td>
<td>.02</td>
<td>[.01, .03]</td>
</tr>
<tr>
<td>Cues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1.93 (.91)</td>
<td>2.94 (1.20)</td>
<td>F=6.02</td>
<td>.58</td>
<td>.74</td>
<td>[.68, .81]</td>
<td>.59</td>
<td>[.57, .61]</td>
<td>.02</td>
<td>[.11, .28]</td>
</tr>
<tr>
<td>V</td>
<td>2.00 (1.05)</td>
<td>2.40 (1.03)</td>
<td>P=.003</td>
<td>H&gt;V, η2=.09</td>
<td>.58</td>
<td>[.57, .61]</td>
<td>.59</td>
<td>[.57, .61]</td>
<td>.02</td>
<td>[.11, .28]</td>
</tr>
<tr>
<td>C</td>
<td>1.72 (.91)</td>
<td>2.14 (1.04)</td>
<td>η2=.09</td>
<td>H&gt;C</td>
<td>.74</td>
<td>[.68, .81]</td>
<td>.59</td>
<td>[.57, .61]</td>
<td>.02</td>
<td>[.11, .28]</td>
</tr>
</tbody>
</table>

Note. Group code: H= habit, V= Variety, C= Control
Table 5b: Habit Model - The Change of Habit and its antecedents between Habit, Variety and Control Groups at Week 8

<table>
<thead>
<tr>
<th>Constructs</th>
<th>G</th>
<th>Baseline Mean (SD)</th>
<th>8 Weeks Mean (SD)</th>
<th>F Values, $p/\eta^2$</th>
<th>Post Hoc</th>
<th>Cohen’s d H and C</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d H and V</th>
<th>Mean Change [95% CI]</th>
<th>Cohen’s d V and C</th>
<th>Mean Change [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habit</td>
<td>H</td>
<td>2.38 (.94)</td>
<td>3.08 (1.01)</td>
<td>F=3.30</td>
<td>H&gt;C</td>
<td>.41</td>
<td>.49</td>
<td>.26</td>
<td>.33</td>
<td>.18</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>2.25 (.88)</td>
<td>2.71 (.85)</td>
<td>$p=.049$</td>
<td></td>
<td></td>
<td>[.43, .55]</td>
<td></td>
<td>[.35, .31]</td>
<td></td>
<td>[.08, .24]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.28 (.98)</td>
<td>2.57 (1.04)</td>
<td>$\eta^2=.05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation</td>
<td>H</td>
<td>2.39 (1.47)</td>
<td>3.30 (1.33)</td>
<td>F=3.27</td>
<td>H&gt;C</td>
<td>.38</td>
<td>.67</td>
<td>.47</td>
<td>.49</td>
<td>.09</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>2.17 (1.36)</td>
<td>2.69 (1.44)</td>
<td>$p=.04$</td>
<td>H&gt;V</td>
<td></td>
<td>[.59, .75]</td>
<td></td>
<td>[.52, .46]</td>
<td></td>
<td>[.08, .29]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>2.16 (1.39)</td>
<td>2.53 (1.45)</td>
<td>$\eta^2=.05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency</td>
<td>H</td>
<td>1.93 (.91)</td>
<td>3.01 (1.38)</td>
<td>F=5.96</td>
<td>H&gt;V</td>
<td>.58</td>
<td>.76</td>
<td>.56</td>
<td>.64</td>
<td>.06</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>2.00 (1.05)</td>
<td>2.40 (1.09)</td>
<td>$p=.00$</td>
<td>H&gt;C</td>
<td></td>
<td>[.69, .83]</td>
<td></td>
<td>[.66, .61]</td>
<td></td>
<td>[.03, .22]</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.72 (.91)</td>
<td>2.18 (1.06)</td>
<td>$\eta^2=.08$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Group code: H= habit, V= Variety, C= Control
Table 6a. Mediation Models of Behavior, Habit, and Action Control Models at Week 4

<table>
<thead>
<tr>
<th>Source</th>
<th>B</th>
<th>β</th>
<th>S.E</th>
<th>95% Confidence Lower Bound</th>
<th>95% Confidence Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior Model (self-report)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group-Habit (a)</td>
<td>.15</td>
<td>.13</td>
<td>.20</td>
<td>-.269</td>
<td>.519</td>
</tr>
<tr>
<td>Habit-Behavior (b)</td>
<td>.32</td>
<td>.08</td>
<td>.09</td>
<td>-.133</td>
<td>.235</td>
</tr>
<tr>
<td>Group-Indirect-Behavior (ab)</td>
<td>.05</td>
<td>.01</td>
<td>.02</td>
<td>-.019</td>
<td>.107</td>
</tr>
<tr>
<td>Group-Behavior (c)</td>
<td>.78</td>
<td>.15</td>
<td>.19</td>
<td>-.299</td>
<td>.443</td>
</tr>
<tr>
<td><strong>Habit Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group-Cues (a1)</td>
<td>.68</td>
<td>.68**</td>
<td>.19</td>
<td>.249</td>
<td>1.00</td>
</tr>
<tr>
<td>Group-Consistency (a2)</td>
<td>.74</td>
<td>.62**</td>
<td>.19</td>
<td>.219</td>
<td>.989</td>
</tr>
<tr>
<td>Cues- Habit (b1)</td>
<td>.12</td>
<td>.06</td>
<td>.11</td>
<td>-.075</td>
<td>.359</td>
</tr>
<tr>
<td>Consistency-Habit (b2)</td>
<td>.24</td>
<td>.33*</td>
<td>.11</td>
<td>.115</td>
<td>.539</td>
</tr>
<tr>
<td>Group-Indirect-Habit (ab)</td>
<td>.26</td>
<td>.23*</td>
<td>.13</td>
<td>.070</td>
<td>.579</td>
</tr>
<tr>
<td>Group-Habit (c)</td>
<td>.11</td>
<td>.11</td>
<td>.21</td>
<td>-.549</td>
<td>.303</td>
</tr>
</tbody>
</table>
Table 6b. Mediation Models of Behavior, Habit, and Action Control Models at Week 8

<table>
<thead>
<tr>
<th>Source</th>
<th>B</th>
<th>β</th>
<th>S.E</th>
<th>95% Confidence Lower Bound</th>
<th>Interval Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior Model (Accelerometer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group-Habit (a)</td>
<td>.36</td>
<td>.46*</td>
<td>.20</td>
<td>.116</td>
<td>.922</td>
</tr>
<tr>
<td>Habit-Behavior (b)</td>
<td>8.16</td>
<td>.06</td>
<td>.10</td>
<td>.10</td>
<td>-.140</td>
</tr>
<tr>
<td>Group-Indirect-Behavior (ab)</td>
<td>2.40</td>
<td>.02</td>
<td>.05</td>
<td>-.046</td>
<td>.300</td>
</tr>
<tr>
<td>Group-Behavior (c)</td>
<td>30.74</td>
<td>.40*</td>
<td>.21</td>
<td>-.015</td>
<td>.812</td>
</tr>
<tr>
<td><strong>Behavior Model (self-report)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group-Habit (a)</td>
<td>.36</td>
<td>.46*</td>
<td>.20</td>
<td>-.024</td>
<td>.757</td>
</tr>
<tr>
<td>Habit-Behavior (b)</td>
<td>1.22</td>
<td>.24*</td>
<td>.09</td>
<td>.085</td>
<td>.474</td>
</tr>
<tr>
<td>Group-Indirect-Behavior (ab)</td>
<td>.44</td>
<td>.11*</td>
<td>.07</td>
<td>.016</td>
<td>.309</td>
</tr>
<tr>
<td>Group-Behavior (c)</td>
<td>1.67</td>
<td>.43*</td>
<td>.20</td>
<td>.034</td>
<td>.814</td>
</tr>
<tr>
<td><strong>Habit Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group-Cues (a1)</td>
<td>.78</td>
<td>.66**</td>
<td>.20</td>
<td>.282</td>
<td>1.07</td>
</tr>
<tr>
<td>Group-Consistency (a2)</td>
<td>.68</td>
<td>.57**</td>
<td>.19</td>
<td>.140</td>
<td>.903</td>
</tr>
<tr>
<td>Cues-Habit (b1)</td>
<td>.19</td>
<td>.32**</td>
<td>.09</td>
<td>.083</td>
<td>.446</td>
</tr>
<tr>
<td>Consistency-Habit (b2)</td>
<td>.13</td>
<td>.11</td>
<td>.10</td>
<td>-.023</td>
<td>.372</td>
</tr>
<tr>
<td>Group-Indirect-Habit (ab)</td>
<td>.24</td>
<td>.22*</td>
<td>.11</td>
<td>.069</td>
<td>.537</td>
</tr>
<tr>
<td>Group-Habit (c)</td>
<td>.27</td>
<td>.36</td>
<td>.21</td>
<td>-.145</td>
<td>.690</td>
</tr>
</tbody>
</table>

Note. *= p<.05, ** = p<.01, *** = p<.001
Figure 2. Mediation Model: Habit mediating Group and Accelerometry MVPA (Baseline-Week 8)

Δ Habit

Group

Δ MVPA (accelerometer)

.46* (a)

.06 (b)

.40* (c)
Figure 3. Mediation Model: Habit mediating Group and MVPA (Baseline-Week 4)
Figure 4. Mediation Model: Habit mediating Group and MVPA (Baseline-Week 8)
Figure 8. Mediation Model: Consistency and Cues Mediating Group and MVPA (Baseline- Week 4)
Figure 9. Mediation Model: Consistency and Cues Mediating Group and MVPA (Baseline - Week 8)

Δ Consistency
Δ Cue Manipulation

Group

.57**
.66**
(a)

.11
.32**
(b)

Δ Habit

.36
(c)
DISCUSSION

Overview

The sharp influx and dropout of gym members during the New Year phase and new school semester is a well-known ongoing phenomenon. Despite the yearly predictable patterns that are both visibly noticeable and reported by the media, research in understanding gym adherence is limited. Although exercise is arguably one of the most effective preventive measures for a variety of diseases and illnesses, it has also been noted as one of the most complicated health behaviours to adopt (Rhodes & Nigg, 2011; B. Verplanken & Melkevik, 2008). It is likely that conscious regulation of this complex behaviour on a daily basis could accumulate a cognitive toll on the individual that would result in deteriorating motivation (Baumeister, Bratslavsky, Muraven, & Tice, 1998). Fortunately, the workload can be reduced by delegating some control to the non-conscious system which consists of reflexive processes such as habit (Evans 2008; Sheeran et al., 2013). Habit is defined as a goal directed automaticity that helps manage the cognitive load while performing the behaviour. Hence, the purpose of these three studies were to understand the role of habit in starting and maintaining a regular exercise routine in gym members.

The first study found new gym members who exercised for at least four sessions per week for six weeks successfully established an exercise habit. Underlying mechanics revealed consistency and affect to be the largest predictors for developing a habit; however, trajectory analyses showed consistency to be the strongest predictor for habit formation. When combined, these findings provide a minimal prescription of what is required and how individuals can
develop habit. Finally, a dual process test found both conscious (intention) and nonconscious (habit) processes provided parallel contributions in predicting exercise change across twelve weeks.

The second study sought to further investigate habit functionality in maintainers that could be applied in new exercisers. After dividing an exercise session into two distinct behaviours (preparatory and performance phase), the results found intention and preparatory habit significantly predicted behaviour change. This novel reveal suggests that a possible area where individuals could be failing to translate their intention is at the preparation phase. However, this has not been experimentally tested. Hence, the purpose of the third study was to conduct a randomized-controlled trial where participants in the habit group would receive instructions based on the previous two studies to develop an exercise habit compared to the control condition.

The experimental study randomized new gym members into a control, habit, or variety group. The RCT found participants in the habit group showed greater mean increases in MVPA over the two other groups with significant difference over the control group at week eight. A similar pattern was also found when comparing changes of habit between the three groups. Finally, the use of cues and consistency were found to be significantly higher among habit participants compared with the other two groups. Overall, this study highlights the importance of developing a preparatory habit via contextual variables (cues and consistency) to facilitate regular exercise routine in gym members. Effective use of consistency and cues can help guide individuals through the preparation process and help transition them to their exercise environment.
Practical Implications

The act of regularly going to the gym consists of necessary behavioural components that are likely similar across a large number of countries such as: a preparatory and performance phase, exercising in the same indoor built environment, and creating a protected time window. Moreover some countries such as Canada, where the majority of the nation experiences long and harsh winters for several months of the year, the gym provides a stable exercise environment which reduces the dependence of external conditions (eg. weather). The dissertation makes several novel contributions to the literature in a recognized, but ignored issue of establishing a habit of attending the gym. This dissertation provided evidence that participants should strive to exercise at least four times per week for six weeks to develop a stable habit. Specifically, a stable preparatory routine that is consistent and uses cues can be helpful in facilitating the automaticity of preparation to make the transition to the gym easier. Overall, these results provide individuals with what to do and how to do it to successfully initiate the adoption of exercise into their regular schedule. The methods applied have demonstrated to be effective for the general adult population and hence can be helpful for other new gym members and trainers alike in facilitating an exercise habit.

Future Directions

Habit was theorized to be the most proximal non conscious determinant to predict behaviour, comparable to the proposed intention-behaviour relationship (Gardner et al., 2011; Rhodes & De Bruijn, 2013; Rhodes, 2014a). However recent studies have suggested that there are likely post-intention constructs that could be effective agents for behaviour change such as behavioural regulation and affective judgments. Given that participants in the third study created
an exercise plan and were provided with methods on making their exercise sessions enjoyable, it is likely that these constructs may have also contributed to behaviour change. Hence, a comprehensive PA model that includes these constructs could provide a larger scope on how other constructs could have facilitated the change of MVPA. In addition to extending the duration and replicating this design with other population groups (eg. shift-workers), analyzing habit in a PA model could provide greater insight into other constructs that may have promoted behaviour change.
Bibliography


Appendices

Appendix 1: Habit Models in Physical Activity: A Systematic Review
Abstract

Researchers have predominantly applied conscious regulatory models to predict physical activity (PA) behavior in the past twenty years. Although these models have helped provide insight into the mediators of PA, considerable variance remains unexplained between intention and behavior. Some researchers have proposed that an automatic process which controls behavior could be a significant component to consider, however, the majority of these studies have not used a habit model to support their findings. The primary purpose of this review was to collect models and theories which predict habit formation with a focus in the psychology and health literature. The secondary and tertiary purposes were to highlight the attributes of the frameworks, and to understand how these models could help advance PA research. Studies were searched from February 2013 to March 2013 using Medline, PsychInfo, PubMed, Scopus SportDiscus, databases which yielded 160 potential articles. Papers were considered eligible if they proposed a model or theory which predicted habit formation. After screening and manual cross-referencing, 15 articles were found which met the eligibility criteria from two disciplines (eight were extracted from psychology, and seven from economic journals). Models from the psychology literature were analyzed based on: i) structure of model (antecedents, heuristics, or transference), and ii) type of repeating mechanism. Economic habit models were organized under: i) origin of habit (internal or external), and ii) goal-direction (myopic or rational). Commonalities in both disciplines were identified and the models were compared within and between disciplines. Overall, the majority of the models require more testing to further validate their effectiveness. Suggestions for using habit models to conduct research in PA are also provided.

Keywords: Habit, theory, model, physical activity, exercise
Until the 1960s, experimental psychology was dominated by behaviorism which proposed that behavioral responses to the environment were automatic and did not require conscious processes to make choices (see review Bargh and Ferguson 2000). Notably behaviorists such as Watson (1913) and Skinner (1938) ruled out cognitive, emotional, and motivational mediators from the stimulus response (S-R) model as they lacked valid measurement and a theoretical framework at the time. Although quite successful in predicting animal behavior in rats and pigeons, the S-R model was limited as it could not entirely explain complex human behaviors as critiqued by social cognitive scientists such as Chomsky (1959). For instance, in the case of verbal behavior (Skinner, 1957), it was theorized that words were spoken in the presence of their associated object which was established during childhood years (ie. parents teaching their child a language). Sentences become formulated as reinforced chains so the first word (that was spoken as a response), becomes the stimulus for the following word which was previously reinforced as a response to the precursor word and so on. The deliberate exclusion of cognitive processes hindered this hypothesis to explain complex human behavior such as language, memory and behavior (Chomsky 1959; Mower, 1960). These shortcomings in behaviorist research encouraged the advancement of conscious regulatory research. Eventually the psychology of automatic processes became overshadowed by social cognitive theories such as the Protection Motivation Theory (W. Rogers, 1974), Social Cognitive Theory (Bandura, 1986) and the Theory of Planned Behavior (Ajzen, 1991). Hence, the application of social cognitive approaches to predict behavior became the norm.
During the mid1980s when physical activity research began to rise, it seemed natural to apply a social cognitive approach to predict behavior given the flaws of the behavioral perspective. The application of conscious regulatory models to predict PA behavior became the norm for the next three decades (Rhodes & Nasuti, 2011; Rhodes & Nigg, 2011). Reviews on the advancement of social cognitive approach to predict behavior has helped contribute to the advancement of PA literature, such as the importance of affect and perceived behavioral control (Rhodes & De Bruijn, 2013; Rhodes & Kates, 2015). However, some reviews have also highlighted some limitations on this approach which have begun to question the validity of social cognitive models in PA research (Rhodes, 2014a, 2014b; Sniehotta et al., 2014). A notable limitation is the absence of a non-conscious construct (Sheeran et al., 2013). The literature appears to be approaching a familiar turning point where the emergence of limitations of an approach makes researchers gravitate towards the alternative. Both cognitive and non-conscious research have presented sufficient convincing findings and hypotheses to simply refute either approach; hence, perhaps an appropriate method would be to strike a middle ground that allows the inclusion of both domains, such as the Dual process approach (Evans, 2008b).

Dual Process approaches propose that our behavior is influenced by a combination of both unconscious (rapid, automatic) and conscious (slow, deliberative) processes (Evans, 2008b). Using our findings from each of these areas can help us substitute appropriate constructs. Although the unconscious system may consist of a bundle of constructs that are not clearly understood, a distinctive component in this system which has seen predictive results is habit (Gardner et al., 2011; Sheeran et al., 2013). Essentially the enactment of sequential automatic micro behaviors can be dated back to research from 1930s. Habit was defined as behavior based on stimulus-response associations and a set of sequential stimulus response patterns can produce
an automatic chain behavior (Hull, 1934). The strength of habit itself could be predicted based on the number of previous reinforcements (Hull, 1943). For example, a recent meta-analysis found habit to correlate $r = .43$ with behavior which is similar to the magnitude of the intention-behavior relationship (Gardner et al., 2011). This suggests that intention and habit carry similar weight in predicting behavior and thus habit could be a possible candidate to match intention when applying the dual process approach.

The complexity of this construct has resulted in several definitions, all which define a segment of the process. When several definitions are combined together, it can provide a better picture of how this construct functions. For instance habit has been defined as a goal-directed automatic behavior (Bargh, 1989; B. Verplanken et al., 1997), that can be prompted by contextual cues (Wood & Rünger, 2015). The process of habit is identified as a discrepancy between thought and behavior (Ouellette & Wood, 1998). The completion of the behavior can predict future habit through a reciprocal behavior (Gardner, 2014).

For instance, the goal to “go home from work” will follow automatic behaviors of packing up, leaving the office, walking down familiar context of hallways to the particular parking lot or bus stop where an individual could be thinking about thoughts unrelated to going to their car. The repetition of this scenario strengthens habit formation. Habit has demonstrated to be a strong predictive validity in various health related behaviors. For instance, in the physical activity domain, a recent meta-analysis found habit to correlate $r = .43$ with behavior which is similar to the magnitude of the intention-behavior relationship (Gardner et al., 2011). This suggests that intention and habit carry similar weight in predicting behavior and thus habit could be a possible candidate to match intention when applying the dual process approach.
While the construct is growing in application, we do not want to make the same mistake of abandon and aboard in the past. Rather, it should be included within a broader theoretical frame to understand development and acknowledge conscious aspects of behavior. With this method, researchers can apply habit while continuing to build with extant theoretical work from the social cognitive paradigm. Although applicable approaches such as Dual Process approach can use habit to help predict behavior, it does not explain how habit is established. Consequently, there is no review which has highlighted proposed models and theories on habit formation. A systematic review which identifies, explains, and compares habit models could be a valuable resource for researchers to further understand habit and behavior in their respective fields. Hence the primary purpose of this study was to perform a comprehensive systematic review across all psychological disciplines to identify all models on habit formation. The secondary purpose was to interpret, appraise and recognize commonalities among the models.

Methods

Evidence Acquisition

Eligibility Criteria

Studies that were published in English peer-reviewed journals were considered for this review. The journal articles were considered eligible if they: i) identified a habit theory/model and ii) if the study was on human behavior.
Exclusion Criteria

Research articles which were thesis papers, conference abstracts, or non-English were excluded. Studies were also excluded if the topic was irrelevant (e.g., ecosystem, environmental toxicology).

Search Strategy

Articles were searched from March 2013 to April 2013 using Medline, PsychINFO, PubMed, Scopus and SPORTDiscus databases. A combination of the following terms were used for the searches including: habit theory/theories, habit model, theories on automaticity, automaticity theories, habitual.

Screening

The screening process of articles by title, abstract and full article was performed based on the eligibility criteria. Figure 1 displays the screening process of the studies.

Data Abstraction and Analysis

Separate data tables were created for both psychology based, and non-psychology habit models. Initially, items for the data abstraction included: year, author, and key notes on models. After the studies were reviewed, additional categories were created to maximize key data extraction concepts in each of the respective disciplines. Psychology models further included model structure and repeating mechanism (Table 1). Non-psychology models, which were revealed to be economic habit models, included origin and goal-direction categories (Table 2).
Common characteristics were only created if the findings were present in three or more studies (Sallis, Prochaska, & Taylor, 2000).

**Results**

**Evidence Synthesis**

The initial literature search yielded 160 potentially relevant articles. After duplicates were removed (25 articles), the remaining studies were filtered from title and abstract if they did not contain the term “habit” or “habitual” in the title or abstract (26 articles). The titles were also screened to determine if they were in English or if they were categorized as abstracts/thesis, or if the study was an irrelevant topic (e.g., ecosystem, environmental toxicology, experiments on animals). This shortened the list to 83 articles that were filtered through the second round of screening which involved reading the abstract. A detailed abstract search was performed and articles were excluded if the search criteria was not met which included the terms habit/habitual and one of the following terms: model, theory, theories, equation. Studies that used performance of a particular behavior and habit interchangeably without distinction, or specifically mentioned habit measured as frequency were removed. The next step involved thoroughly reading the 32 remaining articles which resulted in finding 8 applicable studies for this review. These studies were manually cross referenced which yielded 7 more articles for a total of 15 studies with independent proposed habit models for the systematic review. Figure 1 illustrates the search process.

**Study Characteristics**

Of the 15 studies, eight belonged to psychology and health literature and the remaining were from economic journals. Figures of the psychological models were in the same order as
presented in Table 1 (Figures 2-9). None of the economic studies provided an illustration of their model and due to the abstract nature, a clear system was not provided to illustrate their properties.

Further subcategories were revealed under each of the main characteristics identified in Data Abstraction and Analysis. In the psychological models, the category of *model structure* defined the framework as antecedents, heuristic and/or (intention-habit) transference. Antecedent frameworks were models in which the authors outlined as having certain prerequisites necessary to establish habit formation. Heuristic models were those that explained habit formation in a procedural method. Transference models explained the change of intention according to habit strength. The category of *repeating mechanism* identified component(s) which would be required to repeat the behavior in the future.

Two prominent categories were found in economic models, each were divided into two subcategories. The first was *origin*, which specified if habit formation developed from an internal or external source. Internal habit source was defined as an individuals’ past behavior as the determinant predictor for future behavior (Pollak, 1970). External source proposed that individuals develop habits to keep up with societal changes (A. B. Abel, 1990). The second category was *goal-direction* which identified if habit goal was myopic or rational. Myopic goal implies that individuals formulate habits based on short-term goals or satisfaction (Pollak, 1970), on the other hand, rational goals states that individuals develop habits based on either a logical or long-term goal (Constantinides, 1990).
Psychology Models

*Model Structure*

*i) Antecedents*

Some models demonstrate that habit formation requires the presence of particular antecedents. The present study found four models that fit this category, each model consisted of four antecedents (Bargh, 1994; Grove & Zillich, 2003; Lally & B. Gardner, 2011). Notable findings include the conceptualization of automaticity and habit and role of affect/reward. For instance Bargh’s (1994) model theorized that the presence of “the four horsemen” is required to establish automaticity, whereas Grove & Zillich (2003) proposed that automaticity of the behavior is a necessary component of establishing habit. Lally & Gardner (2011) and Grove & Zillich (2003) models both share a construct which essentially is the evaluation of an individual’s affect after performing the behavior. These have been labelled as reward and negative consequence respectively. These constructs were also labeled as determinants for behavioral repetition, which is further analyzed under the repeating mechanism section.

*ii) Heuristic Models*

Three models proposed habit formation through a heuristic system (Aarts, Paulussen, et al., 1997; Anshel & Kang, 2007; B. Verplanken et al., 1997). The earliest model, developed by Verplanken et al., (1997), distinguished the pathways between a weak and strong habit. The model surmises that weak habits are prompted by goal activation, which then require cognitive processing to eventually make a choice. Although strong habits are also prompted by goal activation, the individual skips the cognitive processing steps and make a binary decision on the choice of performing the behavior. This model was used as a framework to produce a more
complex system by Aarts et al., 1997. The model by Aarts et al., (1997) shows that the need for deliberate planning gradually diminishes as habit strengthens. The authors propose that the core feature of heuristic processing consists of the memory retrieval of the outcome based from past experience. Where a negative experience would require revaluation to perform the particular behavior, the repetition of satisfying experience would create cognitive shortcuts and hence, contribute to habit formation. Finally, the Disconnected Values Model (DVM), developed by Anshel & Kang (2007) also proposes a heuristic framework; however, this model was designed with a goal of disrupting unhealthy habits. These authors propose that the evaluation of an individual’s values is the proximal determinant if the habit will be performed or discontinued. For instance, if an individual’s values align with their current habit then the behavior would be continued. On the other hand, a discrepancy would then lead into an action plan that would help replace negative habits with positive rituals.

iii) Intention-Habit Transference

Four of the eight models specifically included an intention construct in their models (Aarts, Paulussen, et al., 1997; Bargh, 1994; Ouellette & Wood, 1998; Triandis, 1977). These models essentially propose that as habit and intention exist on polar ends of the same continuum. Overtime practiced behaviors under certain conditions become driven by habit instead of intention. All four models placed intention as an immediate proximal determinant to behavior; however, the models differ with regards to how intention is relative to the other constructs. For instance, Triandis (1977) proposed that intention strength is established from various beliefs (see figure 2), however, after repetition of behavior results in the development of a new construct, “use habit”. Use Habit becomes a predictor of the six attitude and social constructs as well as a
proximal predictor to behavior (parallel to intention). With regards to functionality, the remaining three models share similar posits of intention strength and recency of behavior (Aarts, Paulussen, et al., 1997; Bargh, 1994; Ouellette & Wood, 1998). Bargh stated that the easier the experience can be accessed in memory, then the less cognitive processing is required to initiate the behavior and its attitudes. In a similar scope, Ouellette & Wood (1998) used their model to demonstrate that past behavior becomes a stronger correlate than intention to predict future behavior if the activity was performed recently and consistently. Aarts et al., (1997) also concurred and proposed that evaluation of ideas and intentions gradually diminish as habit strengthens. Eventually a smaller heuristic circuit path becomes created in which intention is no longer considered once habit route is established.

*Repeating Mechanism*

Five models consisted of a construct or demonstrated a process to determine if behavior would be repeated in the future (Aarts, Paulussen, et al., 1997; Grove & Zillich, 2003; Lally & B. Gardner, 2011; Triandis, 1977; B. Verplanken et al., 1997). Two models stated that certain outcome expectancies formulate habitual behavior, however each of the models stand on opposite ends of this valence (Grove & Zillich, 2003; Lally & B. Gardner, 2011). The framework proposed by Grove (Grove et al., 2013; 2014) views negative affect for non-performance as an outcome of habit development. In other words, it is suggested that the strength of negative affect for not performing a health behaviour will increase in-line with the extent to which that behaviour has become automated. On the other hand, Lally & Gardner (2011) posited that rewarding experiences would likely be repeated. Both models have justified their rationale based on early behavior theorists. Following early behavioral work, Verplanken et al.,
(1997) model proposes that habitual choices are essentially stimulus-response (S-R) relations. When a strong habit is developed, then the individual may not perceive certain situations with choices as the S-R bound tunnels the individual’s perception from alternative options.

The model by Aarts et al., (2007) proposes that although a favourable outcome is essential for predicting future behavior, the individual evaluates if the same behavior can be executed in the future? The likelihood of behavioral repetition is proposed to depend on the following contingencies: i) satisfaction level of the experience, and ii) ability or behavioral control. If the repetition criteria is not met, then the perceptions are reevaluated. Finally, Triandis (1977) Theory of Interpersonal Behavior postulated that once behavior is repeatedly performed, then a habit construct will develop which would become an independent predictor of behavior (see Figure 2).

**Economic Models**

*Origin of Habit Formation*

The present review found four prominent models which supported the hypothesis that habit formation is internal (Constantinides, 1990; Pollak, 1970; Ryder & Heal, 1973; Spinnewyn, 1981). Essentially researchers who presented an internal hypothesis support the importance of an individuals’ past consumption. The majority of these models proposed that habit persistence is dependent on the consistency or change of “taste” (Pollak, 1970; Ryder & Heal, 1973; Spinnewyn, 1981). In addition to past consumption, Constantinides (1990) defined habit formation and process from an internal development. He postulates that habit formation is the process when repetition of a stimulus diminishes the saliency of the stimulus and its response.
Unlike the previous three models, he proposes that the recent level of consumption carries a higher weight than the individuals’ absolute level of consumption. Hence, he justifies that habit formation can explain why a consumers’ sense of well-being is closely related to recent changes in consumption rather than absolute level of consumption.

Some economists have designed their models with the foundation that habit is formed from an external source (A. B. Abel, 1990; J. Y. Campbell & Cochrane, 2007; Otrok, Ravikumar, & Whiteman, 2002). Abel (1990) stated that habits are created based on the changes of the society and has identified this process as “catching up with Joneses”. The other two models followed this concept and both proposed their hypothesis into respective economic pricing models. They stated that an individuals’ habit level does not depend on his/her past consumption of a commodity, rather, habits become created based on the trends of the economy’s consumption.

Goal-Directed Habit

Four of the seven economic models identified the type goal which causes habit persistence (J. Y. Campbell & Cochrane, 2007; Constantinides, 1990; Pollak, 1970; Spinnewyn, 1981). One model was based on hypothesis that habit is fuelled by myopic goals (Pollak, 1970) while the remaining three argued that the rational behavior hypothesis is better suited to understand habit in commodity consumption (J. Y. Campbell & Cochrane, 2007; Constantinides, 1990; Spinnewyn, 1981). Pollak (1970) stated that an individuals’ present satisfaction depends on his/her past habits. He further proposed that the consumer does not think about the future and labeled this as “Naive habit formation”, which implies that the consumer is also unaware of the habit formation in progress from his/her current consumption.
Although the remaining models hypothesized that habit is formulated or persists due to rational goals, there appears to be variability with regards to their justifications. For instance, Spinnewyn, (1981) stated that if a consumer tastes change through habit stocks depending on past consumption, then the rationale consumer will make adjustments of the future habit forming effects of current consumption. Some of the authors simply stated that they developed their model with this hypothesis because it helps shed insight into a particular unexplained economic condition such as the “Equity Premium Puzzle” (Constantinides, 1990), or it simplifies the analysis (J. Y. Campbell & Cochrane, 2007).

Discussion

Conscious regulatory models have demonstrated to be the most popular tools for PA research over the past twenty years (Rhodes & Nasuti, 2011). Although these frameworks have served to identify constructs which are strong predictors such as perceived behavioral control, or those that are not as strong as once hypothesized such as subjective norm (Rhodes, Courneya, & Jones, 2003; Rhodes & Plotnikoff, 2006; Rhodes, Plotnikoff, & Courneya, 2008), they currently have not explained the asymmetrical relationship between intention and behavior (Rhodes & Dickau, 2012). The emergence of habit research has noticeably expanded within the past decade and has displayed moderate-strong effect sizes for predicting various health behaviors (Gardner et al., 2011). Though habit may present itself as a missing puzzle piece of the intention-behavior discrepancy, habit research using framework is an essential component to understand and compare research progress. Consequently, it has been noted that current social-cognitive theories are not appropriate frameworks for habit research as they do not account for repetition of behavior (B. Verplanken & Melkevik, 2008). The primary purpose of this review was to
identify habit models with a primary focus in psychology and health related disciplines for researchers to consider when conducting research in PA. The secondary purpose was to highlight and then discuss about the models’ components within and between disciplines. Finally, it was to examine the results and understand the use of these models in physical activity research.

The present review revealed 15 models and theories which have been proposed to understand habit behavior; eight models belonged to the psychological literature and seven were extracted from economic journals. The majority of psychology models implemented a repeating mechanism which demonstrated how the completion of a behavioral act can lead to future repetition (Aarts, Paulussen, et al., 1997; Grove & Zillich, 2003; Lally & B. Gardner, 2011; Triandis, 1977; B. Verplanken et al., 1997). Repeating mechanisms have been noted to be essential as it is arguably one of the key components which distinguish between habit from social-cognitive models (B. Verplanken & Melkevik, 2008). With regards to the type of framework, it was found that the psychology models possessed one or more of the following types of structural characteristics: i) antecedents (Bargh, 1994; Grove & Zillich, 2003; Lally & B. Gardner, 2011), ii) heuristic (Aarts, Paulussen, et al., 1997; Anshel & Kang, 2007; B. Verplanken et al., 1997), and/or iii) intention-habit transference (Aarts, Paulussen, et al., 1997; Bargh, 1994; Ouellette & Wood, 1998; Triandis, 1977). Intention-habit transference (IHT) models will be further discussed and compared in terms of: i) their functionality, ii) their study results, and iii) comparing their results to current findings in habit-intention relationship. Models which consisted of antecedents or possessed a heuristic framework will be examined on how they can be applied to future habit research.
It was found that IHT models share a similar standpoint which was that intention is initially required at early stages of behavioral performance, however it eventually fades as habit strengthens. This has been proposed that as habit develops, intention: i) loses its predictive strength (Ouellette & Wood, 1998), or ii) becomes discarded in the establishment of heuristic shortcut (Aarts, Paulussen, et al., 1997). Bargh (1994) emphasized the importance of recency in memory and lower requirement of cognitive resources similar to Ouellette & Wood (1998), and Aarts, et al., (1997) respectively. Triandis (1977) hypothesized that the creation of habit construct develops a synergistic relation with intention as immediate predictor to behavior. In addition, habit also influences the preceding six constructs of intention.

Though the work is limited, some of these models were tested. Ouellette & Wood (1998) conducted a meta-analysis to create their prediction model; they found that for behaviors performed less frequently or inconsistently, the predicted beta values of predicted future behavior were $\beta = .12$, and $\beta = .62$ for past behavior and intention respectively. However, when behaviors were performed regularly or in stable environments, then beta values were $\beta = .45$ and $\beta = .27$ also for past behavior and intention respectively, demonstrating a negative relationship of the two predictors. Triandis (1977) model was also tested, specifically for car use habit. It was found that habitual car use demonstrated a stronger correlation with behavior ($r = .45$) than intention ($r = .31$) (Bamberg & Schmidt, 2003).

The findings of these models which have concluded a negative relationship between habit and intention are in line with other studies (de Bruijn, Kremers, Singh, van den Putte, & van Mechelen, 2009; Rhodes & G. J. De Bruijn, 2010). Interestingly, recent findings have suggested that intention strength is still strong in the presence of strong habits (De Bruijn, Rhodes, & Van
Osch, 2012). The researchers stated that conscious processing (self-regulation) and habit strength could work synergistically to bridge the intention-behavior relation. Further testing using habit models from the present review could yield insightful findings. Hence, a model which specifically identifies habit and intention constructs could be beneficial in further understanding the intention-behavior gap in PA research. However, other frameworks could also explore this discrepancy by combining an intention construct (conscious process) with a habit model (unconscious process) to essentially form the two pillars of the dual-process theory (Evans, 2008a).

This leads to the topic of selecting an appropriate framework for habit research. Models which propose necessary antecedents required to develop habit formation appear to be similar to strategy theories (ST). ST are resourceful tools which provide the researcher with components of what to evaluate in their participants, however, they do not specify guidelines for intervention. On the other hand, heuristic models could be labelled as procedure theories (PT). They may not highlight necessary tools for the researcher, however they do provide a navigation route to study how participants progress in habit formation (Glanz, 1997). As viewed from social cognitive research, both categories have demonstrated to be resourceful in progressing research, such as the Transtheoretical Model (PT), or Theory of Planned Behavior (ST) (Rhodes & Nigg, 2011).

Habit models from the economic literature demonstrated to possess two distinguishing characteristics: origin of habit (internal or external) and goal-direction (myopic or rational). Four models proposed that habit originates from internal source (individual’s preferences and past) (Constantinides, 1990; Pollak, 1970; Ryder & Heal, 1973; Spinnewyn, 1981) and three hypothesized that individuals follow societal habit trends based on an economy’s aggregate
consumption (A. B. Abel, 1990; J. Y. Campbell & Cochrane, 2007; Otrok et al., 2002). The most common combination of the two hypothesis revealed that individuals develop a habit based on their own tastes and past behavior (internal), which is influenced by knowledge-based decision (rational) (Constantinides, 1990; Spinnewyn, 1981).

Although habit models in the economic discipline were developed for the purpose of predicting individual consumption based on particular economic trends, certain elements surfaced which aligned with psychological frameworks. First, psychological theories stated that habits require a reward (Lally & B. Gardner, 2011), consequence (Grove & Zillich, 2003) or value (Anshel & Kang, 2007) to persist. Similarly, economists use the term utility which explains how an individual perceives his/her value of the purchased commodity (A. B. Abel, 1990; Constantinides, 1990). Second, although none of the psychological models proposed an exclusive external habit model, some frameworks consisted of constructs which could be argued as “external factors” such as normative beliefs (Triandis, 1977), external search (B. Verplanken et al., 1997), social pressure (Aarts, Paulussen, et al., 1997), acceptable (Anshel & Kang, 2007) and (external) cues (Lally & B. Gardner, 2011). Third, with regards to rational goal-direction, only one model emphasized importance of continual rational process (Anshel & Kang, 2007). However the purpose of this model was designed to break negative habits, hence conscious decision making would be important to prevent the activation of previous automaticity script. Finally, both disciplines acknowledge the importance of consistency when performing a behavior. Researchers in psychological domain expressed it either directly in their model or functionality such as stability of contexts (Ouellette & Wood, 1998), patterned action (Grove & Zillich, 2003) or consistency (Lally & B. Gardner, 2011). Economists relate consistency to the price of the commodity. For instance, habit has a relative preference for low-frequency volatility
over high-frequency volatility (J. Y. Campbell & Cochrane, 2007; Otrok et al., 2002). Due to the recognized importance in cross disciplinary research, these commonalities could be important traits to consider when selecting an appropriate habit model for research.

Although the majority of psychological habit models primarily consist of internal constructs and do not identify rational processing (Bargh, 1994; Grove & Zillich, 2003; Lally & B. Gardner, 2011; Ouellette & Wood, 1998; B. Verplanken et al., 1997), thus suggesting myopic habit goals, rational goals and external sources could be important constructs to consider when implementing habit framework in health research. First, individuals develop some habits based on rational goals such as wearing seatbelt or hygiene practices. Second, with regards to habit termination, Anshel & Kang’s (2007) model indicated the importance of some variables which could be labelled as external sources such as parental/coaching support, and rational goal (eg. acceptable), to help justify or eliminate a negative habit.

The present review contains some limitations which are important to address. First, only English peer-reviewed published articles were considered for this study. Therefore, potential studies which could have been relevant (eg. thesis or non-English) were not included. Second, the search criteria was limited to the terms and databases described in the method section. The combination of search terms were created to target psychology and health literature; although models from economic journals emerged which were further cross-referenced, specific terms were not formulated to search the economic literature. The notable strength is that to the author’s current knowledge, this is the first review to summarize major proposed models on habit formation.
Despite selecting an appropriate model prior to conducting a study, proper methodologies should be implemented to ensure research fidelity. Performing moderate to vigorous level of physical activity is a strenuous and complex behavior (Ekkekakis et al., 2008; Maddux, 1997); it has been argued that exercise behavior cannot become habitual (Maddux, 1997). Instead, it was later proposed that habituating the decision to exercise (B. Verplanken & Melkevik, 2008) or preparation could be a more viable approach (R. E. Rhodes & G. J. De Bruijn, 2010). These goals could be more feasible considering that the majority of models identified in the present review were designed with hypotheses that habits depend heavily on the individual (internal) and are essentially myopic. Once these considerations are accounted, then the researcher would need to select appropriate tools to measure habit; currently this could be inconvenient as a review does not exist which outlines measurement procedures when conducting this research. Hence, a review on examining various methodologies of conducting habit research in PA would be beneficial for the literature.


Table 1: Psychology Habit Models

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Model Structure</th>
<th>Repeating Mechanism</th>
<th>Tenants/key components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>Triandis</td>
<td>Transference</td>
<td>An independent habit construct will eventually develop and drive behavior</td>
<td>Habit is part of the Theory of Interpersonal Behavior</td>
</tr>
<tr>
<td>1984</td>
<td>Bargh</td>
<td>Antecedents</td>
<td>None</td>
<td>&quot;The Four Horsemen of Automaticity&quot; consists of: i) Awareness ii) Intentionality iii) Efficiency iv) Controllability</td>
</tr>
<tr>
<td>1997</td>
<td>Verplanken</td>
<td>Heuristic</td>
<td>Development of Stimulus-Response continues the repetition of behavior</td>
<td>Weak habit follows cognitive processing. For strong habit, choice is immediately prompted by goal activation.</td>
</tr>
<tr>
<td>1997</td>
<td>Aarts</td>
<td>Heuristic</td>
<td>If outcome is favourable, then repeating behavior is likely if presented in a similar situation</td>
<td>From increased practice, cognitive shortcuts will develop due to memory of past behavioral experience.</td>
</tr>
<tr>
<td>1998</td>
<td>Ouellette</td>
<td>Transference</td>
<td>None</td>
<td>Eventually past behavior becomes a stronger predictor than intention over time</td>
</tr>
<tr>
<td>Year</td>
<td>Author</td>
<td>Category</td>
<td>Antecedents</td>
<td>Avoidance of Negative consequence</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td>----------------------------------</td>
</tr>
</tbody>
</table>
| 2003 | Grove  | Antecedents | Avoidance of Negative consequence | Requirements for habit formation:  
  i) Cue-Driven S-R bonds  
  ii) Automaticity  
  iii) Patterning of Action  
  iv) Negative Consequences if not performed |
| 2007 | Anshel | Heuristic | None        | A cognitive heuristic process combined with an  
  “action plan” can help in disrupting negative habitual behavior |
| 2011 | Lally  | Antecedents | Reward      | Requirements for habit formation:  
  i) Reward  
  ii) Consistency  
  iii) Behavioral complexity  
  iv) Cues |
<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Origin Internal/External</th>
<th>Goal-Direction Myopic/Rational</th>
<th>Key Notes on Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>Pollak</td>
<td>Internal</td>
<td>Myopic</td>
<td>Habit hypothesis: 1) Past consumption influences current preferences and demand, 2) higher level of past consumption of a good implies a higher level of present consumption of that good. An individual’s consumption does not take account of the effect of current purchase on his future preferences. The goods could be “habit forming” which imply that an individual's current preferences depend on his/her past consumption patterns.</td>
</tr>
<tr>
<td>1973</td>
<td>Ryder</td>
<td>Internal</td>
<td>Not specified</td>
<td>Model developed on Duesenberry's hypothesis (1952) that “tastes” can be affected by the level of past consumption.</td>
</tr>
<tr>
<td>1981</td>
<td>Spinnewyn</td>
<td>Internal</td>
<td>Rational</td>
<td>If a consumer tastes change through habit stocks depending on past consumption, then the rational consumer will make allowance of the future habit forming effects on current consumption.</td>
</tr>
<tr>
<td>1990</td>
<td>Abel</td>
<td>External</td>
<td>Not specified</td>
<td>Habit formation consists of the following qualities for utility: 1) utility is time-separable 2) consumer's level of consumption is relative to the &quot;lagged cross-sectional average level of consumption&quot;</td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Type</td>
<td>Framework</td>
<td>Citation</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>1990</td>
<td>Constantinides</td>
<td>Internal</td>
<td>Rational</td>
<td>Utility of consumption is assumed to depend on past levels of consumption. Utility is not time separable in habit persistence. Habit formation is the process when repetition of a stimulus diminishes the perception of the stimulus and responses to it. Habit formation can explain why consumers’ reported sense of well-being is usually more related to recent changes in consumption rather than the absolute level of consumption.</td>
</tr>
<tr>
<td>1999</td>
<td>Campbell and Cochrane</td>
<td>External</td>
<td>Rational</td>
<td>Three features of habit: 1) Habit formation is external. An individual’s habit level depends on the economy’s aggregate consumption rather than the individual’s own past consumption. 2) Habit moves slowly in response to consumption. This feature explains persistent movements in volatility. 3) Habit adapts nonlinearly to the history of consumption.</td>
</tr>
<tr>
<td>2002</td>
<td>Otrok</td>
<td>External</td>
<td>Not specified</td>
<td>A habit agent has a relative preference for low-frequency volatility over high-frequency volatility. Overall, habit agents are much more averse to high-frequency volatility fluctuations than to low-frequency fluctuations.</td>
</tr>
</tbody>
</table>
Figure 1: Flow diagram for the literature search

Potentially relevant publications identified and screened from electronic databases (Medline, PsychInfo, PubMed, Scopus, SportDiscus): 160

Total Duplicates Removed: 25

Potentially relevant citations Screened: 135

Papers excluded after screening for keywords in title and abstract: 52
26- screened out if missing "habit" in title or abstract
23- Irrelevant topic (ie. ecosystem, environmental toxicology)
2- thesis/conference
1- non English

Remaining papers for detailed evaluation: 83

Papers excluded after evaluation of abstract: 51
24- Studies which used a behaviour and habit interchangeably or specially mentioned habit measured as frequency were screened out
27-Screened out if abstract search criteria not met: "habit" and one of the following terms: model, theory, theories, equation

Articles Remaining: 32

Papers excluded after reviewing: 24
3- Non-habit models, ie. instructions, teaching methods, etc
21- Paper did not propose a habit model

Articles Remaining: 8
Total Papers Reviewed: 15

Papers added after manual cross-referencing: 7
Figure 2: Triandis (1977) Theory of Interpersonal Behavior
Figure 3: Bargh’s (1994) Four Horsemen of Automaticity

Note. Bargh labels awareness, intentionality, efficiency and controllability as the “Four Horsemen of Automaticity”.
Figure 4: Verplanken et al., (1997) Habit Process Model
Figure 5: Aarts et al., (1997) Habit Model
Figure 6: Ouellette & Wood’s (1998) Behavior Prediction

Note: The models display the results of generalized least squares regressions which predict frequency of future behavior from past behavior and behavioral intentions. Figure A represents findings from studies which examined behaviors which were performed rarely (e.g., biannually) or occurred in unstable contexts. Figure B represents findings from studies which behavior was performed daily or weekly in stable contexts. *** = p < .001.
Figure 7: Grove & Zillich’s (2003) Habit Requirement Model
Figure 8: Anshel & Kang's (2007) Disconnected Values Model
Figure 9: Lally & Gardner’s (2011) Four Antecedent Habit Model
Appendix 2: Research Methods of Measuring Physical Activity Habit: A Systematic Review
Abstract

Valid methodology for producing high quality research in physical activity (PA) literature has been emphasized in recent reviews. In a similar manner, research studies in PA habit should also be carefully designed with appropriate instruments. Habit, and its role in PA is currently not fully understood, however, the advancement of understanding this construct has led to the evolution of current and creation of new instruments. The primary purpose of this review was to describe the methodology of research in PA habit, in particular to define and describe the uses of the measurements which asses PA habit. The secondary purpose was to discuss the strengths and limitations of the research instruments. Studies were searched in December 2015 using Medline, PsychINFO, PubMed, Scopus, and SportDiscus databases which yielded 545 potentially relevant articles. Papers were considered eligible if the article provided a definition of habit along with an instrument, and the behavior was a form of physical activity (which could include travel-mode choices, general physical activity, or moderate-to-vigorous physical activity). The review found a total of 59 relevant studies and nine instruments for measuring habit (six were indirect and four were direct measures). Direct measurements, which required a researcher to administer, used semantic/prospective testing procedures and indirect used episodic/retrospective techniques. The Self-Report Habit Index was found to be the most commonly used instrument to measure PA which was used in 31 studies. Direct and indirect assessments both present their respective strengths and limitations that are discussed in this review. Procedures which have measured habit in other domains are also discussed (such as qualitative and various computer programs) with the possibility of applying them for measuring PA habit. Methodological suggestions are also provided for future research in this area.

Keywords: habit, automaticity, measure, physical activity, exercise
Background

The use of appropriate measurements is critical for determining patterns in observational studies and the effectiveness of randomized-controlled trials (RCTs) (van Poppel, Chinapaw, Mokkink, van Mechelen, & Terwee, 2010). Instruments generally fall under two categories in psychological research: self-report (eg. questionnaires and diaries), or direct measurements (eg. sensors, GPS, accelerometers, etc) (Prince et al., 2008). Direct measurements can be evaluated by their accuracy (eg. test-retest reliability) and can be further tested with similar (convergent) or different (divergent validity) technological devices. In addition to satisfying these requirements, self-report (or indirect measures), carry an additional set of challenges due to the subjectivity nature of assessment which include: i) aligning items to reflect the construct, and ii) rigorous testing to validate measures (Cook & Beckman, 2006; Messick, 1995). These are critical components to consider when selecting an appropriate measure to assess a construct and hence, reviews that outline the purpose and properties of these measurements can be an invaluable resource for researchers.

The importance of measurement validity is further magnified when assessing controversial constructs, such as habit (Maddux, 1997). Habit is a versatile construct that is used to predict behavior in at least three prominent disciplines: health and consumer psychology and behavioral economics. Habit can be defined as a goal directed automaticity (Bargh, 1989; B. Verplanken et al., 1997) that can be activated under specific contexts (Ouellette & Wood, 1998; Wood & Rünger, 2015). Although habit is associated with repetition of behavior, behavioral frequency alone is not sufficient to develop and accurately measure habit (B Gardner, 2012; Kaushal & Rhodes, 2015; B. Verplanken, 2006). For instance, an individual who goes for a run
on a daily basis does not necessarily reflect habit as the behavior could be facilitated by the presence of automatic processes or could be completely conscious regulated (B. Verplanken & Melkevik, 2008). The complexity of exercise behavior (Ekkekakis et al., 2013; Rhodes & De Bruijn, 2013) combined with the uncertainties of habit formation (Sheeran et al., 2013) make this a challenging research area that some have found to be fruitless (Maddux, 1997). However, the presence of habit can be identified from reflecting the extent of which focused conscious thoughts (or lack of) were used while performing a particular behavior (Kaushal & Rhodes, 2015). Some self-report measures which follow this rationale have shown predictive validity in the physical activity literature (Gardner et al., 2011). Despite this promising finding, the majority of research in understanding physical activity habit is dominated by one instrument, the Self-Report Habit Index (Gardner et al., 2011) followed by its modified version (Gardner, Abraham, Lally, & de Bruijn, 2012). Although both of these measures demonstrating strong psychometric properties (Gardner et al., 2012b), the assessment of a construct via one instrument is not a good idea. Hence, identifying emerging measures is critical to progress research in physical activity habit.

Several recent reviews have noted the importance of habit measures as reflected in a subsection of their papers (Gardner, 2014; B. Verplanken, Myrbakk, & Rudi, 2014; Wood & Rünger, 2015) yet there has not been a clear systematic review that has focused on all potential methods of measuring habit in the PA literature. A systematic review is an efficient scientific technique that reduces random errors and bias (Mulrow, 1994) and has been recognized as a gold-standard approach of presenting a comprehensive scientific review of the literature (Stewart et al., 2015). A review which provides psychometric properties and summarizes the theoretical rationale of habit measurements would an invaluable resource for researchers across multiple
disciplines. Hence, the primary purpose of this study was to conduct a systematic review on all measurements which have been used to assess physical activity habit. The secondary purpose was to highlight the description, proposed usage, and psychometric properties of the instruments. The final purpose was to compare, and discuss the strengths and limitations of each of the measures.

Methods

Search Strategy and Selection

Studies that were published in English peer-reviewed journals were considered for this review. Journal articles were considered eligible if they: i) identified a habit theory/model, and ii) if the study was on human behavior. Research articles which were thesis papers, conference abstracts, or non-English were excluded. Studies were also excluded if the topic was irrelevant (e.g. ecosystem, environmental toxicology). Articles were searched on December 2015 using Medline, PsychINFO, PubMed, Scopus and SportDiscus databases. A combination of the following terms was used for the searches including: SRHI, habit, habitual, methods, measurement, objective, implicit, dual-process, measured, physical activity, and exercise.

Screening

The initial literature search yielded 545 potentially relevant articles. Duplicate scanning removed 138 articles leaving 407 articles which were screened out for keywords in title, and abstract. Articles that were non-English peer reviewed were removed (n=66), including obvious irrelevant articles. Next, 270 articles were excluded after abstract search of “habit” and “exercise”, “physical activity” or “MVPA”. The remaining 71 articles were then read if the author did not provide a clear definition of habit and/or used behavioral frequency instruments to
measure habit. This stage removed an additional 22 articles leaving 49 relevant articles. The remaining studies were cross-referenced which added eight more papers bringing the total number to 57 eligible papers and 59 studies. Figure 1 displays the screening process of the studies.

**Data Extraction and Synthesis**

Table 1 displays general information of relevant studies which measured physical activity habit. Data extracted from these studies included: primary author, year, study design (cross-sectional, prospective or randomized control trial), behavior measured (travel-mode choices, general physical activity, moderate-to-vigorous physical activity), administration method (self-report or researcher) and instrument (name). Behavior measures were coded with one of the following three gradients of physical activity: i) moderate-vigorous physical activity (MVPA) ii) general physical activity- when the researcher used this term without a specific goal, or did not specify vigorous exercise, iii) travel-mode choices- studies which investigated how to promote active travel habits such as walking or cycling.

**Results**

**Description of Studies**

The search process identified 59 relevant studies which spanned across two decades (1994-2016). Of these, 25 were cross-sectional, 26 were prospective (range from 1 week-12 weeks) and 8 were randomized controlled trails which ranged from 8 days to 1 year. Habit measured in these studies include: MVPA (10 studies), general physical activity (31 studies), and travel-mode choices (24 studies).
From these studies, nine different instruments were identified which measured habit, or a component of habit (Table 2). Five instruments measured habit via indirect methods, and four were direct instruments. Direct measurements assessed PA habit by a semantic/prospective procedure which required a researcher to administer the tests; three of these were computer-based assessments and one (Response Frequency Measure I) used an interview protocol which later evolved to computer administration. All indirect instruments used episodic/retrospective techniques and were self-reported.

**Measures**

Reliability, validity, and details of the instruments were extracted. Descriptions of the scales were categorized under self-report or research administered instruments.

**I. Self-Report/ Indirect Measurements**

*Self-Report Habit Index (SRHI)*

The SRHI is a 12 item questionnaire measured on a 5 point likert scale that was developed by Verplanken and Orbell (2003). This scale assess habit based on four components which include: frequency (e.g. ‘…I do frequently’), identify (‘…that's typically “me”’), and automaticity “I have no need to think about doing” which was formulated from Bargh’s three of the four horsemen of his habit model: uncontrollability, lack of awareness, and efficiency (Bargh, 1994). The testing of this instrument demonstrated high internal reliabilities with coefficient alphas of the pretest and post-test indexes of .89 and .92, respectively (Verplanken & S. Orbell, 2003). The SRHI also demonstrated strong convergent validity with the Response Frequency Measure ($r = 58, p < 0.001$).
The present review identified 31 studies which used the SRHI to assess PA habit. Of these studies, five assessed MVPA/exercise at a moderate-vigorous intensity (Chatzisarantis & Hagger, 2007; de Bruijn, 2011; de Bruijn, Rhodes, & van Osch, 2012; de Bruijn & Rhodes, 2011; Gardner & Lally, 2012), 20 measured general physical activity (Fleig, Kerschreiter, Schwarzer, Pomp, & Lippke, 2014; Fleig, Lippke, Pomp, & Schwarzer, 2011; Fleig, Pomp, Schwarzer, & Lippke, 2013; Hyde, S. Elavsky, S. E. Doerksen, & D. E. Conroy, 2012; Jurg, Kremers, Candel, Van Der Wal, & Meij, 2006; Kremers & Brug, 2008; Kremers, Dijkman, de Meij, Jurg, & Brug, 2008; Lally, Van Jaarsveld, Potts, & Wardle, 2010; Maher & Conroy, 2015; Pimm et al., 2015; Rhodes & De Bruijn, 2010; Rhodes, de Bruijn, & Matheson, 2010; Tappe & Glanz, 2013; van Bree et al., 2015; Verplanken & Melkevik, 2008), and six measured travel-mode choices (de Bruijn, Kremers, Singh, van den Putte, & van Mechelen, 2009; Gardner, 2009; Lemieux & Godin, 2009; Murtagh, Rowe, Elliott, McMinn, & Nelson, 2012; Walker, Thomas, & Verplanken, 2015).

**Self-Report Behavioral Automaticity Index (SRBAI)**

The purpose of the SRBAI was to modify the SRHI by testing automaticity in its purest form, which involved removing items that reflect identify and frequency of behavior (Gardner et al., 2012b). It was argued that although frequency is a necessary component for habit formation, frequent behaviors can also be performed with conscious deliberate actions (Gardner et al., 2012b). Second, identity would be more appropriate as a correlate rather than a component of automaticity. This modified version demonstrated a strong convergent validity of .92 with the original scale. In addition, this scale was found to have strong reliability in 34 separate datasets with $\alpha$ of .90-.97 (23 studies), $\alpha$ between .80-.89 (17 studies), and $\alpha$ between .70-.79 (four studies) and one study revealed $\alpha = .68$. The present review found 11 studies that cited using the
SRBAI, or have used the four items identical to the scale (de Bruijn, Gardner, van Osch, & Sniehotta, 2014; Evans et al., 2015; Fleig et al., 2013; Gardner & Lally, 2013; Kaushal & Rhodes, 2015; Peacock et al., 2015; Phillips & Gardner, 2015; Quinlan, Rhodes, Blanchard, Naylor, & Warburton, 2015; Rebar, Elavsky, Maher, Doerksen, & Conroy, 2014; Thurn, Finne, Brandes, & Bucksch, 2014; Vandelanotte et al., 2015).

Habitual Exercise Questionnaire (HEQ)

The HEQ is an 18 item questionnaire which was specifically designed to measure exercise habit. The following are the four subscales of the HEQ with their respective Cronbach alphas: i) automaticity (.77), ii) negative consequences (.78), iii) stimulus cue (.80) and, iv) patterned action (.83) (Grove & Zillich, 2003). Further research has supported criterion-related validity for specific components of the scale (Hairul A. Hashim, Freddy, & Rosmatunisah, 2012; H. A. Hashim, Jawis, Wahat, & Grove, 2014). In addition to the testing of the four antecedents above (Grove & Zillich, 2003), this instrument was used to assess general physical activity habit in two other studies (Hashim, Grove, & Whipp, 2008; Hashim, Hairul Anuar, Freddy, & Rosmatunisah, 2011).

Context Stability

The Behavior X Frequency scale estimates habit strength based on the contextual stability of past behavior. In 1998, Ouellette and Wood proposed that habit strength is dependent on the stability of the environment or context. The context stability instrument measures participants’ behavior performance on a 3 point scale: 1 (rarely/never in the same location), 2 (sometimes in the same location), and 3 (usually in the same location). The authors proposed that habit strength could be measured by multiplying the score from context stability and frequency of behavior.
The present review identified three studies which applied this method for measuring travel mode habits (Friedrichsmeier, Matthies, & Klöckner, 2013; Tappe, Tarves, Oltarzewski, & Frum, 2013; Verplanken, Walker, Davis, & Jurasek, 2008).

This scale was further modified and adapted for estimating traveling habit (Friedrichsmeier et al., 2013). The change included additional questions which consisted of: i) uniformity of trip distance, ii) starting location, iii) time of day, iv) purpose and v) weather on a five point scale from 1- “all different” to 5- “almost always the same”. The average score of these items represented the score for context stability. Tappe and colleagues (2013) also modified this scale to measure general physical activity by adding the following items: i) number of workout days per week (to measure repetitive frequency), ii) consistency of workout time each day (cuing by time), and iii) constancy of workout location (cuing by location).

**Diary and Logbooks**

Measuring habit strength from participant dairies was initially developed by Verplanken and colleagues (1998). Given the qualitative nature of this measure, the inter-rater reliability revealed a Cohen’s kappa= 0.95 between to independent judges who coded the diaries. The habit diary was designed so it would be convenient to complete based on the rationale that participants should not feel cognitively burdened when filling out the diary. Hence, the participants were only required to complete the following regarding their trips: (1) the time of the day (i.e. morning, afternoon, evening), (2) destination and (3) mode of travel. The use of diary/log books was the only method found to estimate habit strength via qualitative assessment. Four studies were found to use diaries to measure travel-mode habits (Garvill, Marell, & Nordlund, 2003; Klöckner & Matthies, 2004; Matthies, Kuhn, & Klöckner, 2002; Verplanken, Aarts, Van
Knippenberg, & Moonen, 1998). Some researchers used diaries to collect additional data to understand habit (Garvill et al., 2003). The travel modes given as alternatives were car as driver, car as passenger, bus, bicycle, walking, and other. The types of trip given as alternatives were to work/study, visiting friends/relatives, leisure activity, to holiday cottage, shopping, services, picking up/dropping off someone or something, other trips, and returning home. The authors also asked if the travel event was a trip chain. A trip chain was defined as a sequence of trips starting and ending at the home.

II. Researcher Administered/Direct Measurements

Response Frequency Measure (RFM- Interview version)

The RFM is one of the earliest instruments designed for measuring PA habit without using past behavioral frequency. This measurement was based on the hypothesis that habits are scripted behaviors (Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994). It was hypothesized that habitual behaviors rely less on attitudes and intentions and are instead guided by the automaticity of stimulus-response, such as destination and travel mode. The RFM asked participants to respond to certain scenarios in a time pressured processes. The purpose of the time pressure was to force respondents to use schemas or scripts rather than conscious deliberations. Time pressured cognitions may trigger episodic experiences from the past but may also activate schematic plans regarding that activity (Abelson, 1981). It was proposed that responses to all scenarios with lack of variability would be indicative of a strong habit between the individual and the type of behavior (Verplanken et al., 1994). The present review found two
studies which administered the RFM via face-to-face interview to assess travel-mode choices (Klöckner & Matthies, 2004; Verplanken et al., 1994).

**Response Frequency Measure (Computer version)**

In 1997, researchers translated the RFM to a computer program (Aarts, Verplanken, & Van Knippenberg, 1997; Verplanken, H. Aarts, & A. Van Knippenberg, 1997). The protocol of the test remained the same with the exception of participants selecting their choices on the computer. Test-retest correlation over a four month period revealed strong reliability of this instrument ($r = 0.92; p < 0.001$) (Aarts, Verplanken, et al., 1997). A total of seven studies administered the computer RFM to assess travel-mode choices (Aarts & Dijksterhuis, 2000a, 2000b; Aarts, Verplanken, et al., 1997; Verplanken et al., 1997).

**Information Acquisition Task (IAT)**

The IAT is a computer program which displays a map of an imaginary town that consists of a town centre, a shopping area, where participants were supposed to go to, the participant's location of his or her imaginary home, the routes of buses, train, and their stops (Verplanken et al., 1997). Participants are then presented with a travel scenario. The travel options are aligned in rows and their attributes in columns (e.g. physical effort, probability of delay, travel time, nuisance caused by other people, expected personal convenience, and freedom). The response to these attributes were delivered in verbal form (e.g. ‘very little’, ‘little’, ‘average’, and ‘much’), except for travel time, which was presented in minutes. Participants were allowed to inspect and re-inspect any piece of information randomly by clicking the mouse on the respective cells. However, the computer recorded which information items and number of times it was inspected.
which was unknown to the participants. It was found that the amount of information the participant required to select a travel choice was negatively correlated with habit strength.

The IAT was used in three studies in a paper by Verplanken and colleagues, (1997) to measure travel modes. The first study presented participants with a familiar travel scenario, the second was an unknown scenario, and the third presented participants with obstacles, such as luggage, weather, etc. Overall, the tests revealed that individuals with strong habits required less information regarding the situation than those with weak habits.

**Single-Category Implicit Association Test (SC-IAT)**

The Implicit Association Test is one of the most commonly used implicit test due to its reliability, ease of administration, and production of large effect sizes (Greenwald, McGhee, & Schwartz, 1998). The SC-IAT is a modification of the IAT that requires participants to indicate as quickly and accurately as possible, if each presented stimulus belonged to one of two categories. For instance in Conroy, et al., (2010) the task consisted of two blocks; in one block, the categories were “physical activity+ good”/“bad.” In the other block, the categories were “physical activity bad”/“good.” The stimuli were presented in the centre of the screen and were randomly sequenced. Responses faster than 300 ms or slower than 10,000 ms were discarded (Conroy et al., 2010). Internal consistency of the SC-IAT was found to average at $r = .69$ (ranging from $r = .55$ to $r = .85$) (Karpinski & Steinman, 2006). These reliability coefficients were similar to the internal consistency for IATs used in these studies (average $r = .73$; ranging from $r = .58$ to $r = .82$) (Karpinski & Steinman, 2006).

The present review identified three studies which used the SC-IAT to predict physical activity habits (Aarts & Dijksterhuis, 2000b; Conroy et al., 2010; Amanda L. Hyde, Elavsky, Doerksen, & Conroy, 2012). The SC-IAT requires participants to indicate as quickly and
accurately as possible, if each presented stimulus belonged to one of two categories. For instance in Conroy, et al., (2010) the task consisted of two blocks; in one block, the categories were “physical activity+ good”/“bad.” In the other block, the categories were “physical activity bad”/“good.” The stimuli were presented in the centre of the screen and were randomly sequenced. Responses faster than 300 ms or slower than 10,000 ms were discarded (Conroy et al., 2010).

Discussion

The progression of habit research has resulted in an evolution of habit models over time (Aarts, Paulussen, et al., 1997; Anshel et al., 2010; Bargh, 1994; Gardner, 2014; Lally & Gardner, 2013; B. Verplanken et al., 1997; Wood & Rünger, 2015) and the incorporation of a habit construct to understand various physical activity behaviors has seen an explosive growth over the last five years. Although some of the models have provided a certain criteria for assessing habit, the majority do not recommend a particular approach for the researcher. Hence, this leaves the researcher open to compare different measures of habit to test a model. Consequently comparing habit measurements is currently not a convenient process as there is no review which has outlined these instruments. Hence, the primary objective of the present study was to identify all measurements which were used to assess various physical activity habits (e.g. MVPA, general PA, and travel-mode choices). The secondary objective was to describe the use and psychometric properties of the instruments. The final objective was to compare the strengths and weakness of these measures.

Direct Measurements
The direct habit measurements were found to use a semantic assessment procedure to evaluate the strengths of established scripts and schemas regarding a particular behavior (e.g. travelling). The first direct measure designed to assess PA habit was the Response Frequency Measure which was administered via interviews (Verplanken et al., 1994). This was later adapted as a computer program that allowed researchers to implement a precise protocol to strengthen control measures (Verplanken et al., 1997). Given that this approach focuses on the speed of participants’ response, the computer adapted version helped control the time allowance for responses (at the expense of the researcher possessing the appropriate equipment). The control for time response became a prudent component in direct assessments as it was deemed that a response within a particular time window would capture non-conscious functioning. An alternative digital method was the SC-IAT, which was also a time pressured test that shares the same strength of controlling for confounding variables (Conroy et al., 2010). In contrast, the Information Acquisition Task demonstrates that time pressured responses may not be the only direct approach to measure habit strength (Verplanken et al., 1997). This instrument removes any errors regarding potential differences in computer dexterity/motor control movement. However, the use of this instrument is limited and more research is required to establish the reliability and validity values for this measurement.

Generally, computer assessments offer precise and consistent data collection without the concern of human measurement error, and inter-rater reliability issues. However, this also presents obvious challenges such as: possessing appropriate equipment, training research assistants, and participant inconvenience (e.g. attending the research lab). Though these limitations could be worth compensating for the methodological strengths they offer which include direct testing procedures and semantic testing. Direct data collection method offer a
reliability strength of assessment while semantic procedures contribute to the validity of the measure by assessing the strengths of established schemas or personal scripts (Abelson, 1981). Although frequency of the same answer in a future scenario would predict habit strength (Aarts, Verplanken, et al., 1997), frequency is a valid measure of habit if the scripts are elicited via: i) time pressured situation, and ii) the absence of extra information.

**Indirect Measurements**

The present review found the SRHI to be the most commonly used instrument for measuring physical activity habits accounting for 53% of all studies. Some authors proposed that the frequency and identity items may contaminate the validity of the questionnaire (B Gardner, 2012; Gardner et al., 2012b; Sniehotta & Presseau, 2012). These considerations prompted the development of an abbreviated version, the Self-Report Behavioral Automaticity Index (SRBAI), which was designed to measure automaticity in its purest form by removing frequency and identity items. Hence, this scale is not to be mistaken for an entire habitual process measure as it does not include antecedents or prerequisites for habit formation such as the Habitual Exercise Questionnaire (Grove & Zillich, 2003). Rather, this versatility allows the researcher to select a habit model of choice and substitute measures to assess the respective antecedents that would predict automaticity strength.

The most unique indirect measure was by Ouellette & Wood (1998) which unlike the previous three was not a Likert-scale instrument. This measure assesses habit strength based on the contextual stability by multiplying the frequency of the behavior with the stability of the context. Some researchers have used a combination of assessing the conditions along with automaticity process to estimate habit strength. For instance, Tappe and colleagues (2013) used the SRHI to assess the psychological process of automaticity along with the contextual stability.
measure from Wood and colleagues (2005) to assess the behavioral component of habit. A similar approach was used by Friedrichsmeier and colleagues (2013) but the researchers administered the RFM to estimate the psychological habit score. Finally Kaushal & Rhodes (2015) used a scale to assess each habit antecedent in the model proposed by Lally & Gardner (2013) then used the SRBAI to predict the automaticity strength.

Finally, habit has also been measured via qualitative approach which includes diary and logbooks. Since the participants carry the diaries on their person, this method offers a unique advantage of proximal assessment to the task and could potentially reduce errors of recall bias. For instance, Wood and colleagues (2002) instructed participants to provide hourly entries of ongoing behavior which included thoughts and feelings. This allowed the researchers to code for context consistency and identify characteristics of habitual behavior.

Overall the strengths and limitations of direct and indirect habit measures align with assessments of other psychological constructs. Similar to other self-report measures, indirect habit assessments have low participant burden, economical procedure, and general acceptance (Prince et al., 2008). However technology-based direct measures can maintain the procedural reliability across participants and in repeated assessments. Some researchers have attempted to cover the limitations of both approaches by administering a combination of direct and indirect measurements which was found in four studies (Friedrichsmeier et al., 2013; Garvill et al., 2003; Klöckner & Matthies, 2004; Verplanken et al., 1998). Given that some measurement procedures include assessing the process of automaticity while others incorporate antecedents with automaticity, identifying a habit model to base the research along with a definition would be needed to justify the type of measure. Hence, the importance of providing a clear definition of
habit in future studies cannot be overemphasized. As displayed in Figure 1, 21 studies were removed due unclear definition of PA habit or using PA habit interchangeably with frequency.

**Future Research- Measurement and Design**

The process of constructing the present review and analysis of the relevant studies has helped reveal some methodological issues which would be important to consider for future research. First, a clear definition of habit as a psychological process rather than behavioral frequency is essential to distinguish between habit and repeated behavior. Finally, the search found methods of assessing other habitual behaviors which have not been translated to measuring PA habit that could be considered. For instance, other procedures include Stroop tasks (Orbell & Verplanken, 2010), primed lexical decision task (Adriaanse, Gollwitzer, de Ridder, de Wit, & Kroese, 2011), qualitative interview (Lally, Wardle, & Gardner, 2011) or participant observation (Holland, Aarts, & Langendam, 2006; Lewis & Eves, 2012).

Overall, the method of administration (direct vs. indirect) and testing procedures (episodic vs. semantic) both present their respective strengths and limitations. Researchers should consider these factors in addition to their purpose (habit strength vs. new habit formation) to select adequate prospective designs for their studies. Instruments that have been used to measure other habit behaviors should also be considered for assessing PA habit.
References


Table 1: Data Extraction of Studies

<table>
<thead>
<tr>
<th>Primary Author</th>
<th>Year</th>
<th>Study Design</th>
<th>Behavior Measured</th>
<th>Self-Report or Researcher</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aarts</td>
<td>1997</td>
<td>Cross-sectional</td>
<td>Travel-mode choice</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td>Aarts</td>
<td>2000</td>
<td>Cross-sectional</td>
<td>Travel-mode choice</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td>Aarts</td>
<td>2000b</td>
<td>Cross-sectional</td>
<td>Travel-mode choices</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>Study 1: common destination</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>Study 2: travel unusual goal</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>Study 3: location and travel mode association</td>
<td>Researcher-computer</td>
<td>Single-Category Implicit Association Test</td>
</tr>
<tr>
<td>Bamberg</td>
<td>2003</td>
<td>Cross-sectional</td>
<td>Travel-mode choice</td>
<td>Self-Report</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td>Conroy</td>
<td>2010b</td>
<td>Prospective-1 week</td>
<td>General Physical Activity</td>
<td>Researcher-computer</td>
<td>Single-Category Implicit Association Test</td>
</tr>
</tbody>
</table>
| Chatzisarantis (study 1) | 2007 | Prospective- 5 weeks | MVPA                       | Self-Report               | i) Mindful Attention Awareness Scale  
<p>|               |       |                     |                            |                           | ii) Self-Report Habit Index                     |
| De Bruijn     | 2009  | Cross-sectional     | Travel-mode choice         | Self-Report               | i) Self-Report Habit Index                      |
|               |       |                     |                            |                           | ii) Items on bicycle usage for transportation and leisure |
| De Bruijn     | 2011  | Prospective-2 weeks | MVPA                       | Self-Report               | Self-Report Habit Index                         |
| De Bruijn     | 2011b | Cross-sectional     | MVPA                       | Self-Report               | Self-Report Habit Index                         |
| De Bruijn     | 2012  | Prospective-2 weeks | MVPA                       | Self-Report               | Self-Report Habit Index                         |</p>
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Study Type</th>
<th>Measure</th>
<th>Data Collection Method</th>
<th>Habit Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Bruijn</td>
<td>2014</td>
<td>Prospective-2 weeks</td>
<td>MVPA</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Fleig</td>
<td>2011</td>
<td>Prospective- 6 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Modified Self-Report Habit Index</td>
</tr>
<tr>
<td>Fleig- Study I PSE</td>
<td>2013</td>
<td>Prospective-2 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Fleig- Rehab Psych</td>
<td>2013</td>
<td>Prospective-18 months</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Modified Self-Report Habit Index</td>
</tr>
<tr>
<td>Fleig</td>
<td>2014</td>
<td>Prospective-18 months</td>
<td>General Physical Exercise</td>
<td>Self-Report</td>
<td>Modified Self-Report Habit Index</td>
</tr>
<tr>
<td>Fleig</td>
<td>2016</td>
<td>RCT- 4 months</td>
<td>Balance and strength</td>
<td>Self-Report</td>
<td>Modified Self-Report Habit Index</td>
</tr>
<tr>
<td>Friedrichsmeier</td>
<td>2013</td>
<td>Prospective- 2 weeks</td>
<td>Travel-mode choice</td>
<td>Researcher-computer</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Self-Report</td>
<td>Context Stability</td>
</tr>
<tr>
<td>Gardner (study I)</td>
<td>2009</td>
<td>Prospective-1 week</td>
<td>Travel-mode choice</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Gardner (study II)</td>
<td>2009</td>
<td>Prospective-1 week</td>
<td>Travel-mode choice</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Gardner</td>
<td>2012</td>
<td>Prospective-1 week</td>
<td>MVPA</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Gardner</td>
<td>2013</td>
<td>Prospective-1 week</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Garvill</td>
<td>2003</td>
<td>RCT-3 weeks</td>
<td>Travel-mode choice</td>
<td>Self-Report</td>
<td>Travel Diary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Researcher-interview</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Design</td>
<td>Subject</td>
<td>Methodology</td>
<td>Dataset Description</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>------------</td>
<td>--------------------------------</td>
<td>---------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Grove</td>
<td>2008</td>
<td>Cross-sectional</td>
<td>MVPA</td>
<td>Self-Report</td>
<td>Habitual Exercise Questionnaire</td>
</tr>
<tr>
<td>Hashim</td>
<td>2008</td>
<td>Cross-sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Habitual Exercise Questionnaire</td>
</tr>
<tr>
<td>Hashim</td>
<td>2011</td>
<td>Cross-sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Habitual Exercise Questionnaire</td>
</tr>
<tr>
<td>Hyde</td>
<td>2012</td>
<td>Prospective-1 week</td>
<td>General Physical Activity</td>
<td>Researcher-computer</td>
<td>Single-Category Implicit Association Test</td>
</tr>
<tr>
<td>Hyde</td>
<td>2012b</td>
<td>Prospective-4 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Jurg</td>
<td>2006</td>
<td>RCT- 1 year</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Kassavou</td>
<td>2014</td>
<td>RCT- 3 months</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Klockner</td>
<td>2004</td>
<td>Prospective- 4 weeks</td>
<td>Travel-mode choice</td>
<td>Researcher-interview</td>
<td>Response frequency measure Logbook</td>
</tr>
<tr>
<td>Kremers</td>
<td>2008</td>
<td>Cross-sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Kremers</td>
<td>2008b</td>
<td>Cross-sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Kaushal</td>
<td>2015</td>
<td>Prospective-12 weeks</td>
<td>MVPA</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Lally</td>
<td>2010</td>
<td>Prospective-12 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Lemieux</td>
<td>2009</td>
<td>Prospective-2 weeks</td>
<td>Travel-mode choices</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Maher</td>
<td>2015</td>
<td>Prospective-1 week</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Matthies</td>
<td>2002</td>
<td>Cross-sectional</td>
<td>Travel-mode choice</td>
<td>Researcher-interview</td>
<td>Response frequency measure</td>
</tr>
<tr>
<td>Murtagh</td>
<td>2012</td>
<td>Prospective- 1 week</td>
<td>Travel-mode choice</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Study Type</td>
<td>Variable(s)</td>
<td>Method</td>
<td>Index</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>------------------</td>
<td>------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Panter</td>
<td>2011</td>
<td>Cross Sectional</td>
<td>Travel-mode</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Peacock</td>
<td>2015</td>
<td>RCT- 12 Months</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Phillips</td>
<td>2015</td>
<td>Prospective- 1 month</td>
<td>MVPA</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Primm</td>
<td>2015</td>
<td>Cross Sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Quinlan</td>
<td>2015</td>
<td>RCT-6 Month</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Rebar</td>
<td>2014</td>
<td>Prospective- 2 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Rhodes</td>
<td>2010a</td>
<td>Prospective-2 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Rhodes</td>
<td>2010b</td>
<td>Prospective-2 weeks</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Tappe</td>
<td>2013</td>
<td>Cross-sectional</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Exercise Habit Survey</td>
</tr>
<tr>
<td>Vandelanotte</td>
<td>2015</td>
<td>RCT- 9 months</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Behaviour Automaticity Index</td>
</tr>
<tr>
<td>Van Bree</td>
<td>2013</td>
<td>Prospective-1 year</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self- Report Habit Index</td>
</tr>
<tr>
<td>Van Bree</td>
<td>2015</td>
<td>Prospective- 1 year</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Verplanken</td>
<td>1994</td>
<td>Cross-sectional</td>
<td>Travel- mode choices</td>
<td>Researcher-interview</td>
<td>Response frequency measure</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Design</td>
<td>Domain</td>
<td>Data Collection Method</td>
<td>Measures</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>----------------</td>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Verplanken</td>
<td>1997</td>
<td>Cross-sectional</td>
<td>Travel-mode choices</td>
<td>Researcher-computer</td>
<td>i) Response frequency measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ii) Information Acquisition Test for all three studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>study 1: known scenario</td>
<td>Researcher-computer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>study 2: unknown scenario</td>
<td>Researcher-computer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional</td>
<td>study 3: circumstances, time</td>
<td>Researcher-computer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>restriction, etc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verplanken</td>
<td>1998</td>
<td>Cross-sectional</td>
<td>Travel-mode choices</td>
<td>Researcher-interview</td>
<td>Response frequency measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Self-Report</td>
<td>Diary</td>
</tr>
<tr>
<td>Verplanken</td>
<td>2008</td>
<td>Prospective-1 month</td>
<td>General Physical Activity</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
<tr>
<td>Verplanken</td>
<td>2008b</td>
<td>Cross-sectional</td>
<td>Travel-mode choices</td>
<td>Self-Report</td>
<td>Response Frequency Measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Self-Report</td>
<td>Context stability</td>
</tr>
<tr>
<td>Walker</td>
<td>2015</td>
<td>Prospective-19 months</td>
<td>Travel-mode choices</td>
<td>Self-Report</td>
<td>Self-Report Habit Index</td>
</tr>
</tbody>
</table>

*Note. RCT = Randomized Controlled Trial, MVPA = Moderate-to-Vigorous Physical Activity.*
### Table 2: Description of Instruments that measured PA Habits

<table>
<thead>
<tr>
<th>Instrument</th>
<th>First Author, and Year</th>
<th>Type of Measure (Direct or indirect)</th>
<th>Administration (Self-Report, Researcher)</th>
<th>Testing Procedure</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFM I</td>
<td>Verplanken, 1994</td>
<td>Direct</td>
<td>Researcher - interview</td>
<td>Semantic</td>
<td>None reported</td>
<td>None reported</td>
</tr>
<tr>
<td>(Interview version)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFM II</td>
<td>Verplanken, 1997</td>
<td>Direct</td>
<td>Researcher - computer</td>
<td>Semantic</td>
<td>Test-retest reliability was $r = .92, \ p &lt; 0.001$</td>
<td>This measure correlated 0.58 (p&lt;0.001) with self-report frequency of car choices</td>
</tr>
<tr>
<td>(Computer version)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRHI</td>
<td>Verplanken 2003</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Episodic</td>
<td>Test-retest reliability was $r = .91, \ p &lt; 0.001$</td>
<td>Convergent validity of SRHI with RFM was $r = 0.58, \ p &lt; 0.001$</td>
</tr>
<tr>
<td>SRBHI</td>
<td>Gardner, 2012</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Episodic</td>
<td>Internal consistencies: $\alpha = .90-.97$ (23 studies) $\alpha = .80-.89$ (17 studies) $\alpha = .70-.79$ (4 studies) $\alpha = .68$ (1 study)</td>
<td>The SRHI and SRBAI Correlated at $r = .92 (p &lt; .001)$</td>
</tr>
<tr>
<td>HEQ</td>
<td>Grove, 2003</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Episodic</td>
<td>Test-retest reliability was $r = .71-.78, \ p &lt; 0.001$</td>
<td>The composite measure of habit strength correlated significantly with stage of change classification based on the Transtheoretical model ($r = .67, \ p &lt; .001$).</td>
</tr>
<tr>
<td>Method</td>
<td>Year</td>
<td>Type</td>
<td>Report Type</td>
<td>Frequency</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------------</td>
<td>-----------</td>
<td>--------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>MAAS</td>
<td>Brown, 2003</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Episodic</td>
<td>Test-retest reliability: difference in scores Extraversion, $r = .19, p &lt; .01$; Openness to experience, $r = .18, p &lt; .01$</td>
<td>Convergent validity: Trait Meta-Mood Scale $r = .46, p &lt; .001$ Mindfulness/mindlessness Scale $r = .31, p &lt; .001$</td>
</tr>
<tr>
<td>Context Change</td>
<td>Wood, 2005</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Episodic</td>
<td>None reported</td>
<td>None reported</td>
</tr>
<tr>
<td>Diary</td>
<td>Verplanken, 1998</td>
<td>Indirect</td>
<td>Self-Report</td>
<td>Semantic</td>
<td>None reported</td>
<td>Only (Verplanken et al., 1997) found inter-rater reliability Cohen's kappa = 0.95,</td>
</tr>
<tr>
<td>IAT</td>
<td>Verplanken, 1997</td>
<td>Direct</td>
<td>Researcher-computer</td>
<td>Semantic</td>
<td>None reported</td>
<td>None reported</td>
</tr>
<tr>
<td>SC-IAT</td>
<td>Karpinski, 2006</td>
<td>Direct</td>
<td>Researcher-computer</td>
<td>Semantic</td>
<td>Internal consistency Averaged $r = .69$</td>
<td>Reliability coefficients are similar to the internal consistency observed for the Implicit Attitude Test average $r = .73$</td>
</tr>
</tbody>
</table>

Note. The following are the abbreviations for the scales: RFM= Response Frequency Measure, SRHI= Self-Report Habit Index, SRBAI= Self-Report Behavioral Automaticity Index, HEQ= Health Exercise Questionnaire, MAAS= Mindfulness Attention Awareness Test, IAT= Information Acquisition Task, SC-IAT= Single Category Implicit Attitude Test.
Figure 1: Flow diagram for the literature search

Potentially relevant publications identified and screened from electronic databases (Scopus, PsychInfo, SportDiscus, Medline, PubMed): 583

Total Duplicates Removed: 138

Potentially relevant citations screened: 445

Papers excluded after screening for keywords in title and abstract: 66
35- Non-academic studies
19- Thesis, abstracts
10- Rats animal experiments
2- Non-English

Remaining papers for detailed evaluation: 315

Papers excluded after evaluation of abstract: 270

270-did not meet the following search criteria in abstract: habit + physical activity or exercise

Articles Remaining: 45

Papers excluded after reviewing:
21- articles that measured habit as frequency
Articles Remaining: 24

Papers added after manual cross-referencing: 14

Total papers reviewed: 38
Total studies reviewed: 41
Appendix 3: The Home Physical Environment and its Relationship with Physical Activity and Sedentary Behavior: A Systematic Review
Abstract

Reviews of neighbourhood (macro) environment characteristics such as the presence of sidewalks and aesthetics have shown significant correlations with resident physical activity (PA) and sedentary (SD) behavior. Currently, no comprehensive review has appraised and collected available evidence on the home (micro) physical environment. The purpose of this review was to examine how the home physical environment relates to adult and child PA and SD behaviors. Articles were searched during May 2014 using Medline, PsycINFO, PubMed, Scopus, and SPORTDiscus databases which yielded 3265 potential studies. Papers were considered eligible if they investigated the presence of PA (eg. exercise equipment, exergaming devices) or SD (eg. television, videogames) equipment and PA or SD behavior. After, screening and manual cross-referencing, 49 studies (20 experimental and 29 observational designs) were found to meet the eligibility criteria. Interventions that reduced sedentary time by using TV limiting devices were shown to be effective for children but the results were limited for adults. Overall, large exercise equipment (eg. treadmills), and prominent exergaming materials (exergaming bike, dance mats) were found to be more effective than smaller devices. Observational studies revealed that location and quantity of televisions correlated with SD behavior with the latter having a greater effect on girls. This was similarly found for the quantity of PA equipment which also correlated with behavior in females. Given the large market for exercise equipment, videos and exergaming, the limited work performed on its effectiveness in homes is alarming. Future research should focus on developing stronger RCTs, investigate the location of PA equipment, and examine mediators of the gender discrepancy found in contemporary studies.
Context

Regular physical activity has been associated with the prevention of at least 25 chronic diseases (D. E. R. Warburton et al., 2007). Despite these findings, physical activity (PA) remains low and consequently obesity and the comorbidities associated with low PA levels have increased (Shields et al., 2010; Tremblay et al., 2010). Researchers have identified several correlates of PA behavior which can be broadly defined into categories of: demographic, biological, intra-individual/psychological, behavioral, social/inter-individual, environmental and policy (Bauman et al., 2012; Ferreira et al., 2007; Trost, Owen, Bauman, Sallis, & Brown, 2002). More recently, understanding the correlates of sedentary (SD) behaviors has become an important and emerging issue. Sedentary behavior is defined as energy expenditure at \( \leq 1.5 \) METs (Pate, O’Neill, & Lobelo, 2008) (Metabolic Equivalent of Task). Despite meeting PA guidelines, excessive sedentary lifestyle can deteriorate health over time (Owen, Healy, Matthews, & Dunstan, 2010; Proper, Singh, Van Mechelen, & Chinapaw, 2011). These findings have prompted the creation of sedentary behavior guidelines for Canadian and Australian children (AGDH, 2013; CSEP, 2012).

The importance of the environment on PA and SD behaviors is reflected in social ecological models (Sallis, Bauman, & Pratt, 1998; Spence & Lee, 2003; Wachs, 1992). The physical environment can represent a discriminate stimulus (Skinner, 1954) which can prompt predictable human behavior (Spence & Lee, 2003). Ecological models also posit that individuals adapt or alter their behaviors in response to the resources in the extra-individual environment. Research on how the neighbourhood environment predicts PA behavior has seen noticeable growth in the past decade; it has been shown to represent approximately 30% of all published research in PA (Rhodes & Nasuti, 2011). While this evidence clearly demonstrates the
importance of the neighbourhood environment on PA, some researchers have suggested that understanding the effects of the home, or micro environment also deserves attention (Sirard, Laska, Patnode, Farbakhsh, & Lytle, 2010a). Individuals are likely to receive higher exposure to stimuli in their homes compared to their neighbourhood environment. For instance, the home environment has been shown to be a determining factor in understanding nutritional choices (K. J. Campbell et al., 2007; Hendrie, Sohonpal, Lange, & Golley, 2013; MacFarlane, Cleland, Crawford, Campbell, & Timperio, 2009) and smoking behavior (Hiemstra et al., 2014; Rushton, 2004). With regard to active lifestyle, the convenience and advancements of technology (e.g., laptops, video game consoles, tablets, etc.) are likely factors that prompt sitting. An average American spends 8 hours per day being sedentary (Matthews et al., 2008) and children spend on average 7.5 hours/day using various entertainment media (i.e., computers, televisions, cell phones) (Prevention, 2014). The home provides personal comfort which makes it an ideal environment to engage in common sedentary activities. However, exercise equipment such as treadmills and exergaming can also provide a convenient method for staying active at home. Both of these types of equipment would seem essential to consider for children and adults. Despite this rationale, only one review has been conducted on the home environment (Maitland, Stratton, Foster, Braham, & Rosenberg, 2013). The results generally supported the premise that the home environment is reliably linked to PA and SD, but it was limited to children, and thematic analyses were constrained to broad classifications (e.g., equipment vs. no equipment, placement of equipment).

Thus, the purpose of this paper was to create a systematic review (Moher, Liberati, Tetzlaff, & Altman, 2010) which would complement the prior review (Maitland et al., 2013) by including adults and updating the contemporary literature on how the home physical
environment relates to adult and child PA and SD. It was hypothesized that the physical components of PA and SD equipment (e.g. quantity and location) and variables within these components (type, and individual factors) would correlate with SD time and PA time respectively.

**Methods**

**Evidence Acquisition**

**Eligibility Criteria**

Studies that were published in English peer-reviewed journals were considered for this review. The journal articles were considered eligible if they investigated: i) the presence of PA (e.g. treadmill) or SD (e.g. television) equipment, and ii) an outcome of PA or SD behavior.

**Exclusion Criteria**

Studies were excluded if the setting was other than a family home (e.g. nursing homes, schools, and recreation facilities); the authors wanted to examine a setting in which the individuals have the autonomy to manipulate their surroundings.

**Search Strategy**

Articles were searched during May 2014 using Medline, PsycINFO, PubMed, Scopus, and SPORTDiscus databases. A combination of the following terms was used to search in title, abstract, and key-terms which included: *home, home environment, physical activity, sedentary, television, screen-time, obesogenic, obesogenic environment, home-based intervention, exercise video, exercise program, exercise dvd, treadmill, bike, exercise bike, cues, stimulus control, exergaming, exercise gaming, eyetoy, dancemat, playstation.*

**Screening**
The screening process of articles by title, abstract and full article was performed based on the eligibility criteria. Figure 1 displays the screening process of the articles.

**Quality Assessment**

Two separate instruments were used to evaluate experimental research and observational studies, respectively. The Cochrane Collaboration Risk of Bias Tool was used for the experimental studies which appeared to be suitable for this review (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012). The observational studies were evaluated by using a modified version of the Downs and Black’s 22-item assessment tool (the modified version excludes items which assess RCTs) (Kirk & Rhodes, 2011).

**Data Abstraction and Analysis**

Items for the data abstraction included: primary author and year, participants, inclusion criteria, length/design, intervention (if applicable), relevant primary instruments, correlations/beta and significance and quality rating. Information on participants provided data such as adults/children, sample size, clinical sample (if applicable). Themes were created on physical factors that could predict behavior: quantity, location, equipment, materials, backyard, and environment interaction. However, themes were established if the findings were present in three or more studies (Sallis et al., 2000). From these themes, common findings which emerged such as gender, ethnicity, type of exergaming, etc. were further grouped within each of the themes when possible. It was later found that there was not sufficient homogeneity of data to conduct a meta-analysis (Hunter & Schmidt, 2004).
Results

Evidence Synthesis

The primary author and a research assistant performed two independent literature searches. The initial search yielded 3,265 potentially relevant articles. After duplicates were removed (1,187 articles), the remaining articles were filtered from title and abstract by removing studies that did not fit the inclusion criteria such as non-home environment, or intervention which did not include correlates of physical home environment. This shortened the list to 96 articles to further evaluate. The next stage involved thoroughly reading the studies which resulted in finding 42 applicable studies for this review. These studies were manually cross-referenced which yielded 7 new articles for a total of 49 studies with independent data-sets for the systematic review. Figure 1 illustrates the search process.

Study Characteristics

Of the 49 studies reviewed, 20 were experimental and 29 were observational studies. Characteristics of experimental and observational studies can be found in Tables 1A, 1B and Tables 2A, 2B respectively. The length of interventions ranged from 6 weeks to 16 months and length of passive prospective designs ranged from 1 to 5 years. In the 20 experimental studies, three focused on reducing sedentary behavior and 17 focused on increasing PA. Quality evaluation of experimental studies resulted in eight strong, five moderate, and seven weak interventions (see tables 3A-3C). In the observational studies, 15 focused on SD behavior and 14 focused on PA behavior. Quality evaluation of the studies revealed that half were high (score: 12-15) and the other half were medium (score: 9-11). Tables 4A-4C display the study
characteristics of all observational studies. Appendix 1A and 1B provide a summary of the items from the two instruments.

*Decreasing Sedentary Behavior*

Studies which manipulated the physical environment to reduce sedentary behavior were placed in this category. Three studies were identified, and all used the same intervention approach which involved the implementation of a TV limiting device (French, Gerlach, Mitchell, Hannan, & Welsh, 2011; French, Mitchell, & Hannan, 2012; Ni Mhurchu et al., 2009). The first two studies (French et al., 2011; French et al., 2012) which were evaluated as high quality interventions (see Table 3) spanned for over a year. Two of the three studies found TV limiting devices to be an effective method for reducing television time among children with medium effect sizes (French et al., 2012; Ni Mhurchu et al., 2009) and the only study that measured adults also found it to be a successful tactic for reducing TV viewing time (French et al., 2011).

*Increasing Physical Activity - Implementation of Exercise Equipment*

A total of seventeen intervention studies focused on facilitating physical activity in homes (Baranowski et al., 2012; Canning, Allen, Dean, Goh, & Fung, 2012; Graves, Ridgers, Atkinson, & Stratton, 2010; Jakicic, Winters, Lang, & Wing, 1999; Khalil et al., 2012; Maddison et al., 2011; Madsen, Yen, Wlasiuk, Newman, & Lustig, 2007; Maloney et al., 2008; Mark & Rhodes, 2013; Moore et al., 2009; Ni Mhurchu et al., 2008; Oka, DeMarco, & Haskell, 2005; Owens, Garner, Loftin, van Blerk, & Ermin, 2011; Paez, Maloney, Kelsey, Wiesen, & Rosenberg, 2009; Plotnikoff, Eves, et al., 2010; Rhodes, Warburton, & Bredin, 2009; Vestergaard, Kronborg, & Puggaard, 2008). The method of environment manipulation used was categorized as either: i) the implementation of exercise equipment (Canning et al., 2012; Jakicic
et al., 1999; Khalil et al., 2012; Moore et al., 2009; Oka et al., 2005; Plotnikoff, Eves, et al., 2010; Vestergaard et al., 2008), or ii) the modification of sedentary equipment (Baranowski et al., 2012; Graves et al., 2010; Maddison et al., 2011; Madsen et al., 2007; Maloney et al., 2008; Mark & Rhodes, 2013; Ni Mhurchu et al., 2008; Owens et al., 2011; Paez et al., 2009). In the first category, the equipment could be subdivided into exercise machines (Canning et al., 2012; Jakicic et al., 1999; Oka et al., 2005; Plotnikoff, Eves, et al., 2010) and exercise videos (Khalil et al., 2012; Moore et al., 2009; Vestergaard et al., 2008). Among exercise machines, three of these interventions implemented a treadmill in houses (Canning et al., 2012; Jakicic et al., 1999; Oka et al., 2005) and one implemented a multi-functional exercise machine (Plotnikoff, Eves, et al., 2010); all studies involved adult samples. One study was a true randomized controlled trial (Jakicic et al., 1999) and the other three were quasi-experimental designs (Canning et al., 2012; Oka et al., 2005; Plotnikoff, Eves, et al., 2010). In the RCT, the intervention group which received a treadmill demonstrated greater adherence compared to the control group with a medium sized difference (70.1%; h=.43). In Canning et al., (Canning et al., 2012) the treadmill group appeared to demonstrate better exercise adherence, (2.6 sessions/week) compared to the control group (1.25 session/week) over 6 weeks; however the effect size is unknown as standard deviations were not provided. Finally Plotnikoff et al., (Plotnikoff, Eves, et al., 2010) found the mean adherence of their program was 71±22%; although this may appear adequate, the lack of a control group limits interpreting the effectiveness of their intervention.

Three studies were found that used exercise DVDs as home PA interventions among adults (Moore et al., 2009) (Khalil et al., 2012; Vestergaard et al., 2008). These experiments were difficult to compare as the quality of each exercise DVD could mediate motivation. Moreover, the interventions were rated weak with only one earning a moderate score (see Table 3A). In the
three studies that provided exercise videos, the length of intervention and adherence rates found were: 6 weeks, 69% (33/48) (Moore et al., 2009); eight weeks, 53% (Khalil et al., 2012); and five months, 89.2% (Vestergaard et al., 2008). It’s important to note that all studies used different clinical samples, which could attribute to variability in barriers to continue performing. Overall, providing exercise DVDs could be an economical procedure to start PA; however, further research with stronger methodology would lead to more conclusive findings.

*Increasing Physical Activity - Modification of Sedentary Equipment*

Studies that modified a SD device with the objective of increasing PA were categorized under this heading. The most common studies were those that modified a standard video game console into an exergaming system (Baranowski et al., 2012; Graves et al., 2010; Maddison et al., 2011; Madsen et al., 2007; Maloney et al., 2008; Mark & Rhodes, 2013; Ni Mhurchu et al., 2008; Owens et al., 2011; Paez et al., 2009). Three of the studies earned a strong quality rating (Baranowski et al., 2012; Maddison et al., 2011; Mark & Rhodes, 2013), and the remainder were moderate. Nine studies targeted children, and two included both adults and children (Mark & Rhodes, 2013; Owens et al., 2011). The length of interventions ranged from 6-28 weeks with adherence rates spanning from 69-100%. The majority of the experiments (seven studies) compared the effectiveness of the intervention to a control group (Baranowski et al., 2012; Graves et al., 2010; Maddison et al., 2011; Maloney et al., 2008; Mark & Rhodes, 2013; Ni Mhurchu et al., 2008; Paez et al., 2009).

The types of devices and effect sizes are as follows: exercise bikes $d=.85$ ($p<0.01$) (Mark & Rhodes, 2013); no significance for Nintendo Wii (Baranowski et al., 2012; Owens et al., 2011); significance for Dance-Dance Revolution (DDR) with the latter study showing OR of
participating increase from 1.2 to 3.485 (p<0.05) if other video games were absent (Maloney et al., 2008; Paez et al., 2009); and no significance for peripheral devices (Sony Eyetoy and Wii jOG) (Graves et al., 2010; Maddison et al., 2011; Ni Mhurchu et al., 2008). However, the use of inactive games significantly decreased in the group that received the Sony Eyetoy (Maddison et al., 2011; Ni Mhurchu et al., 2008). Overall, it appears that devices which provide cardiovascular type of exercises such as exercise bikes or DDR were effective in facilitating PA.

Observational Studies

Sedentary Equipment

Quantity of SD Equipment and SD time

Ten studies investigated the correlation between the quantity of media resources and total TV viewing/sedentary time (Adachi-Mejia et al., 2007; Barr-Anderson, Van Berg, Neumark-Sztainer, & Story, 2008; Bauer, Neumark-Sztainer, Fulkerson, Hannan, & Story, 2011; Dennison, Erb, & Jenkins, 2002; Gorin, Phelan, Raynor, & Wing, 2011; Hoyos Cillero & Jago, 2011; James N. Roemmich, Leonard H. Epstein, Samina Raja, & Li Yin, 2007; Salmon et al., 2013; Sirard, Laska, Patnode, Farbakhsh, & Lytle, 2010b; Van Zutphen, Bell, Kremer, & Swinburn, 2007). Of these, two conducted analysis with children (Bauer et al., 2011; Van Zutphen et al., 2007) and six conducted gender analysis (Adachi-Mejia et al., 2007; Barr-Anderson et al., 2008; Dennison et al., 2002; Hoyos Cillero & Jago, 2011; J. N. Roemmich, L. H. Epstein, S. Raja, & L. Yin, 2007; Sirard et al., 2010a). All studies were cross-sectional designs. Four studies scored high quality ratings (Bauer et al., 2011; James N. Roemmich et al., 2007; Salmon et al., 2013; Sirard et al., 2010b) and two scored a moderate rating (Hoyos Cillero & Jago, 2011; Van Zutphen et al., 2007).
Some studies indicated mixed results on the relationship between the number of television sets and viewing time (Bauer et al., 2011; Gorin et al., 2011; Van Zutphen et al., 2007); however, when the researchers conducted separate gender analysis, adolescent girls were associated with a greater number of media equipment in homes and higher viewing time $r = 0.12-0.29$ (Barr-Anderson et al., 2008; Dennison et al., 2002; Hoyos Cillero & Jago, 2011; James N. Roemmich et al., 2007; Sirard et al., 2010b). These results suggest that quantity of media equipment in homes is more likely to have a behavioral effect (increase sedentary time) for girls compared to boys.

**Location of Sedentary Equipment**

Eight studies examined the relationship between the location of the equipment and behavior (Adachi-Mejia et al., 2007; Barr-Anderson et al., 2008; Dennison et al., 2002; Sirard et al., 2010a; Van Zutphen et al., 2007); all of these studies investigated the behavior effects of children having a television in their bedrooms. Five studies were high quality (Adachi-Mejia et al., 2007; Dennison et al., 2002; Salmon et al., 2013; Sirard et al., 2010b; Wethington, Pan, & Sherry) and three were moderate (Barr-Anderson et al., 2008; Hoyos Cillero & Jago, 2011; Van Zutphen et al., 2007). Two studies were prospective designs (Barr-Anderson et al., 2008; Salmon et al., 2013), while the remaining six were cross-sectional (Adachi-Mejia et al., 2007; Dennison et al., 2002; Hoyos Cillero & Jago, 2011; Sirard et al., 2010b; Van Zutphen et al., 2007; Wethington et al.).

Overall, seven studies found that televisions in bedrooms were related to SD and PA behavior (Barr-Anderson et al., 2008; Dennison et al., 2002; Sirard et al., 2010b; Van Zutphen et al., 2007). Of these studies, six found a significant difference of SD time between children who
had TVs in their bedrooms compared to those who did not with very large effect sizes (d=7.2-11.16) (Barr-Anderson et al., 2008; Dennison et al., 2002; Salmon et al., 2013; Sirard et al., 2010a; Van Zutphen et al., 2007; Wethington et al.) indicating a strong positive correlation between television in bedrooms and SD time. Four studies further revealed a negative association between television in bedrooms and PA time (Barr-Anderson et al., 2008; Bauer et al., 2011; Sirard et al., 2010b). In Barr-Anderson et al., (Barr-Anderson et al., 2008) televisions in bedrooms was significantly different in "moderate/vigorous activity" (d=3.27) and "vigorous activity" (d=3.77). However, households with boys were a predictor of having television sets in their bedrooms and greater viewing time (Adachi-Mejia et al., 2007; Barr-Anderson et al., 2008; Dennison et al., 2002). All seven studies found households with children to be correlated with quantity and television sets in bedrooms. These findings suggest that television sets located in a bedrooms is correlated with unhealthy behavior- both, an increase in SD and decrease in PA time.

Finally, three studies were found that investigated the moderating role of ethnicity and television in bedrooms with behavior (Adachi-Mejia et al., 2007; Barr-Anderson et al., 2008; Dennison et al., 2002). Dennison et al., (Dennison et al., 2002) showed that more Black children (51%; h=.66) and more Hispanic children (50%; h=.64) had a TV set in their bedroom than White children (20%) or other race children (31%; h=.25), and these results were parallel to Adachi-Mejia et al., (2007) (Adachi-Mejia et al., 2007). Interestingly, in Barr-Anderson et al., (Barr-Anderson et al., 2008) the researchers found TVs in bedrooms in Asian children (39%) to be lower than White (60.2%; h=.42), Black (81.5%; h=.90), Hispanic (66.3%; h=.55) and other racial groups (78.8%; h=.83). Overall, these results suggest that children from white
families have a lower prevalence of televisions in their bedrooms than most other racial groups, with the potential exception of Asians.

**Physical Activity Equipment and Materials**

*Quantity of Physical Activity Equipment and Usage*

Fourteen studies investigated the association between the quantity of home PA equipment and behavior (Atkinson, Sallis, Saelens, Cain, & Black, 2005; Bauer et al., 2011; Dunton, Jamner, & Cooper, 2003; Gorin et al., 2011; Kerr, Norman, Sallis, & Patrick, 2008; Maddison et al., 2009; Patnode et al., 2010; Reed & Phillips, 2005; Ries, Dunsiger, & Marcus, 2009; Sirard et al., 2010b; Spurrier, Magarey, Golley, Curnow, & Sawyer, 2008; Stuckyropp & Dilorenzo, 1993; Van Dyck et al.; David M. Williams et al., 2008). Six of these studies examined adult samples (Atkinson et al., 2005; Gorin et al., 2011; Reed & Phillips, 2005; Ries et al., 2009; Van Dyck et al.; David M. Williams et al., 2008) and the remaining studied children/adolescents (Bauer et al., 2011; Dunton et al., 2003; Kerr et al., 2008; Maddison et al., 2009; Patnode et al., 2010; Sirard et al., 2010b; Spurrier et al., 2008; Stuckyropp & Dilorenzo, 1993). The results in this category are divided into: exercise equipment and physical activity materials. We define exercise equipment as objects designed for repetitive exercise behavior such as a treadmill, exercise bike, weights, or other exercise machines. PA materials were identified as either mobile or require a partner for use (e.g., Frisbee, tennis racket).

*Exercise Equipment*

Seven studies assessed behavior based on the presence of exercise equipment in the home (e.g., treadmill, bicycle, trampoline and weights) (Atkinson et al., 2005; Dunton et al., 2003;
Gorin et al., 2011; Kerr et al., 2008; Reed & Phillips, 2005; Van Dyck et al.; David M. Williams et al., 2008). Six studies administered the Perceived Environment Related to Physical Activity questionnaire (Sallis, Johnson, Calfas, Caparosa, & Nichols, 1997) or a modified version (Atkinson et al., 2005; Dunton et al., 2003; Kerr et al., 2008; Reed & Phillips, 2005; Van Dyck et al.; David M. Williams et al., 2008). Dunton et al., (Dunton et al., 2003) sampled adolescent girls and found that use frequency correlated with availability (r=.22) and variety (r=.67) of equipment. Reeds and Phillips (Reed & Phillips, 2005) assessed the quantity of exercise equipment in the home with adults and separated their findings into specific components of PA behavior. Significance was found for all types of PA equipment as predictors of behavior, but only for adult females. Overall these results suggest that the presence of exercise equipment at home appears to be more likely used by adolescent and adult females.

Physical Activity Materials

Six studies were found that investigated the presence of PA materials (Bauer et al., 2011; Maddison et al., 2009; Patnode et al., 2010; Ries et al., 2009; Sirard et al., 2010b; Stuckyropp & Dilorenzo, 1993). All studies used adolescents with a cross-sectional design, with the exception of Ries et al., (Ries et al., 2009) who used adult samples with a prospective follow-up after one year. Two studies investigated the presence of PA materials by administering the Physical Activity and Media Inventory (PAMI) (Patnode et al., 2010; Sirard et al., 2010b) and other measurements include Physical Activity Interview (Stuckyropp & Dilorenzo, 1993), Perceived Environment Related to Physical Activity (Ries et al., 2009), Home Equipment Scale (Rosenberg et al., 2010) and custom scales (Maddison et al., 2009).
The majority of studies found correlations between exercise behavior and the presence of PA materials at home; however, the results were often moderated by gender. Patnode et al., (Patnode et al., 2010) found that availability and accessibility of PA materials significantly predicted MVPA among boys but not girls. Sirard et al., (Sirard et al., 2010b) found the availability of equipment to correlate with the PA of both boys and girls with small effects. Analysis by categories revealed that sport materials also correlated with MVPA accelerometer minutes for boys and girls, while outdoor materials as found to only correlate with girls but not for boys. The researchers also found that the strongest association with MVPA was predicted by total PA material density (total number of items divided by the total number of rooms/locations). Maddison et al., (Maddison et al., 2009) found that ownership of PA materials predicted PA for undergraduate students, and the effect size was similar to Stuckyropp and Dilorenzo (Stuckyropp & Dilorenzo, 1993) who identified the quantity of these items predicted PA only for girls. Similar to exercise equipment, the measurement of PA materials in various environmental factors (accessibility, density, quantity, etc.) predominantly predicted PA in girls.

*Backyards as Predictor of Behavior*

A total of eight studies were found that investigated the relationship between PA and a backyard for children (Bauer et al., 2011; Dunton et al., 2003; Liao, Intille, & Dunton, 2014; Patnode et al., 2010; Ries et al., 2009; Spurrier et al., 2008; Trang, Hong, Dibley, & Sibbritt, 2009; Veitch, Salmon, & Ball, 2010) Two studies reported that 30-33% of physical activity took place in the yard (Liao et al., 2014; Veitch et al., 2010). The presence of a backyard alone decreased SD time, d=.38 (Trang et al., 2009) and yard size also demonstrated a small positive correlation with PA (r =.17) (Spurrier et al., 2008). Of the five studies that included backyard
equipment in their measures (Bauer et al., 2011; Dunton et al., 2003; Patnode et al., 2010; Ries et al., 2009; Spurrier et al., 2008), only one study conducted a separate analysis on yard equipment and found it to correlate with PA (r=.20) (Spurrier et al., 2008). The presence of a backyard and PA materials appear to provide an opportunity for children to engage in PA or reduce their SD behavior.

**Micro and Macro Environment Interaction**

The interaction between the home physical (micro) and neighbourhood (macro) environment for predicting behavior was found in four studies (Kerr et al., 2008; Salmon et al., 2013; Veitch et al., 2010; Wong et al., 2010). All four studies used cross-sectional designs and only one was an adult sample (Kerr et al., 2008). Evaluation of these studies revealed only one to have strong score (Salmon et al., 2013), and the other three were medium quality (Kerr et al., 2008; Veitch et al., 2010; Wong et al., 2010).

Three studies found neighbourhood safety to moderate PA at home; in particular, safety concerns correlated with greater PA at home rather than away from home (Kerr et al., 2008; Salmon et al., 2013; Veitch et al., 2010). For instance, Kerr et al., (Kerr et al., 2008) found a significant interaction between perceived safety and equipment usage, where home equipment use was not related to PA in safe neighborhoods (OR=1.07) but was strongly related to PA in dangerous neighborhoods (OR=4.40). Veitch et al., (Veitch et al., 2010) echoed similar findings with the prediction of playing in the home yard, when there was a “stranger danger” concern (OR=2.32) or road safety concerns (Salmon et al., 2013). Finally, the availability of sport facilities in neighborhood was significantly associated of PA only if video game consoles were absent in home (OR_{boys}=1.26; OR_{girls}=1.34) (Wong et al., 2010). Overall, home equipment use was predictive of PA more if there were neighbourhood safety concerns.
Discussion

The home environment can provide ease of access to a variety of equipment which could prompt both PA and SD activities. More importantly, we have the autonomy to modify the equipment in our homes and can essentially shape our own PA and SD behavior. Hence, understanding how the characteristics of physical equipment in our homes predict our behaviors could provide insight on conducting pragmatic interventions. The purpose of the present review was to summarize and appraise the literature on how the home physical environment affects both PA and SD behavior. The collated results helped identify some important findings and highlight methodological limitations to consider for future research.

Experimental Designs

Currently, the majority of research has focused on facilitating PA by implementing exergaming consoles over providing exercise equipment. The present review found exercise games that required cardiovascular exercises such as exercise bikes, and DDR were among the most effective in promoting PA. However, the effectiveness of these games was moderated by other environmental factors such as if inactive video games were present (Maloney et al., 2008; Paez et al., 2009) and if adolescents received new active video games to maintain novelty/stimulus strength (Maddison et al., 2011). The present review found only two studies that investigated the effectiveness of exergaming on adults (Mark & Rhodes, 2013; Owens et al., 2011). The adult population deserves more attention considering that: i) the average video game player is 37 years old with 53% belonging to 18-49 age group (EESA, 2013) and ii) children are likely to model adult behavior (Rhodes & Quinlan, 2014).
Providing exercise equipment, particularly substantive exercise equipment (treadmills, portable exercise systems) showed promise in exercise participation. However, only one intervention was a true RCT and the remaining were quasi-experimental designs with partial analysis. Moreover, in the majority of interventions, it was unclear whether it was equipment or usage prompts which facilitated behavior (Canning et al., 2012; Plotnikoff, Eves, et al., 2010). The best RCT on the matter showed a medium to large effective size between treadmills in homes and behavior (Jakicic et al., 1999), however, more research employing RCT methodology is needed. Finally, a small set of studies examined exercise DVDs for home use. Although effective and economical, the combination of lack of control groups, variability of types of exercise videos, participant variability in health conditions, and self-reported measures compromise the fidelity of these studies.

The Sport and Fitness Industry Association found that selling exercise equipment is a $4.49 billion business with treadmill sales accounting for 25.5% of entire category (SFIA, 2012). In addition, the sales for exergaming consoles are projected to reach $40 billion by 2015 (WebMD, 2013). The North American population is making substantial investments on equipment that can facilitate their PA in their homes yet the sparse research on effectiveness of these products is alarming.

In terms of SD behavior, another set of experimental studies examined the effectiveness of TV limiting devices. These devices were generally effective but findings were limited for adults. Television viewing is a common sedentary activity (Rhodes & Dean, 2009) which could become habitual. Maintaining the requirements of a TV limiting device (e.g. frequent deposit of coins) complicates the behavior and likely reduces some reward, both which have been shown to
be necessary antecedents for habit formation (P. Lally & B. Gardner, 2011). Thus, these devices may show efficacy for adults who wish to break SD habits but it is unlikely to thwart SD motivation.

Observational Evidence

Overall, 29 observational studies were found that investigated the physical components of the micro environment; this is a limited and disproportionate number compared to research on the macro environment (Rhodes & Nasuti, 2011). Although the physical components of the external or macro environment are important, as evidenced by a substantial amount of research in neighborhood studies (see review of reviews) (Gebel, Ding, & Bauman, 2014), external physical components are relatively non-mutable (Kaushal & Rhodes, 2014). One of the most interesting findings from the observational studies was a notable gender difference. Although boys and minority children were found to have more TVs in their bedrooms, the total quantity of media equipment in the home was correlated with SD behaviors only for girls. Thus, interventions that inform parents on limiting media equipment, particularly in bedrooms, would be an important preventive measure for healthy homes.

The use of PA equipment showed similar findings to SD equipment; exercise equipment and PA materials both correlated with behavior for females but not males. It has been found that most women do not prefer traditional exercise environments due to gender differences in exercise context (Kruisselbrink, Dodge, Swanburg, & MacLeod, 2004). Hence, women may feel more comfortable exercising in their own homes. With the combined findings, we suggest that PA equipment designed towards women could be effective in facilitating their exercise behavior at home.
Another interesting theme that emerged was the interaction between the physical home and neighborhood. Overall, a perceived lack of neighbourhood safety predicted greater use of equipment (PA or SD) at home. However, this is a complex relationship, as low socioeconomic status families usually reside in neighbourhoods with safety concerns (Weir, Etelson, & Brand, 2006). It has been previously suggested that indoor screen-based entertainment is a convenient method for parents to keep children entertained and safe (Tandon et al., 2012). Moreover, it might be unlikely that low SES homes would consist of physical features that correlate with PA such as treadmills or a backyard.

**Limitations and Future Research**

A noticeable limitation found in these studies involved methodological issues which could compromise the quality of findings, particularly, study design and measurement validity. Out of the 29 observational studies, six were prospective designs, and only 11/20 experiments were true RCTs. Some studies did not use validated scales and some that did, failed to use the potential of the subscales in the measure (i.e., types of equipment, location, etc). Finally, the variability of populations could also be a limitation such as clinical populations, ethnic backgrounds and SES.

The present review also consists of some limitations which are also important to address. First, only English peer-reviewed published articles were considered for this study. Therefore, potential studies which could have been relevant (e.g. Thesis or Non-English) were not included. Second, the search criterion was limited to the terms and databases described in the method section.
In conclusion, the (micro) home environment represents a potentially important context for PA and SD behavior intervention based on theoretical and pragmatic grounds. Our review identified 49 studies in this context. Interventions that reduced sedentary time by using TV limiting devices were shown to be effective, but results were limited particularly for adults. Overall, prominent exercise and exergaming equipment were found to be more effective than smaller devices. Although exercise DVDs were shown to be effective, future studies should incorporate controlled trial methodology and also consider other modes of exercise video such as Netflix or streaming devices. Observational studies revealed that the location and quantity of televisions correlated with SD behavior with the latter having a greater influence on females. This was similarly found for the quantity of PA equipment which also correlated with behavior in girls. Given the large market for exercise equipment, videos and exergaming, the limited and relatively low-quality work performed on its effectiveness in homes is alarming. Future research should focus on developing stronger RCTs, investigate the location of PA equipment, and examine mediators of the gender discrepancy found in contemporary studies.
Table 1A: Data Extraction of Experimental Studies: Participant and Study Characteristics

<table>
<thead>
<tr>
<th>Primary Author and Year</th>
<th>Participants</th>
<th>Inclusion Criteria</th>
<th>Length/Design</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baranowski 2012</td>
<td>78 children</td>
<td>Participants were children 9 to 12 years of age, with a BMI &gt;50th percentile, but &lt; 99th</td>
<td>RCT 13 weeks (PA measured for 5 weeks)</td>
<td>Wii console, no prescription</td>
</tr>
<tr>
<td>Canning 2012</td>
<td>20 participants</td>
<td>Inclusion: patients with Parkinson's disease</td>
<td>RCT 6 weeks (semi-supervised)</td>
<td>Home treadmill walking</td>
</tr>
<tr>
<td>French 2011</td>
<td>90 Households (HH)</td>
<td>(i) at least one child ages ≥5 years two HH members ages ≥12 years (ii) residence in a private house or apartment within 20 miles of the university (iii) HH TV viewing weekly average of ≥10 h per person (iv) no HH members with dietary, medical, psychological, or physical limitations that would prevent their participation</td>
<td>RCT 1 year</td>
<td>TV limiting device</td>
</tr>
</tbody>
</table>
in intervention activities; and
(v) willingness to be randomized to
active intervention or control group.

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Inclusion Criteria</th>
<th>Study Design</th>
<th>TV Limiting Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>153 adults</td>
<td>1) at least 1 adult and 2 HH members (including the adult) ages &gt; 12 years; (2) residence in a private house or apartment within 20 miles of the university; (3) HH weekly average TV of &gt; 10 hours per person; (4) no HH members with dietary, medical, psychological, or physical limitations that would prevent their participation in intervention activities; and (5) willingness to be randomized to active intervention or control group.</td>
<td>RCT 1 year</td>
<td>TV limiting device</td>
</tr>
<tr>
<td>Graves</td>
<td>42 children</td>
<td>owned a PS2 or PS3 video game console and self-reported playing these for &gt; 2 hr/wk age range, 8-10 years</td>
<td>RCT 12 weeks</td>
<td>jOG device (add to ps2/ps3) records steps on spot</td>
</tr>
<tr>
<td>Jakicic</td>
<td>148 women</td>
<td>exclusion: medical condition</td>
<td>RCT 18 months</td>
<td>3 intervention groups:</td>
</tr>
<tr>
<td></td>
<td>25-45 years old</td>
<td></td>
<td></td>
<td>1. long-bout exercise group</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. short-bout exercise group</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. short-bout + exercise equipment group</td>
</tr>
<tr>
<td>Khalil</td>
<td>15 participants</td>
<td>Inclusion criteria: patients with Huntington's</td>
<td>RCT</td>
<td>Home DVD exercise program</td>
</tr>
<tr>
<td>Year</td>
<td>Study</td>
<td>Participants</td>
<td>Inclusion</td>
<td>Exclusion</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>2012</td>
<td>Maddison</td>
<td>322 children</td>
<td>10-14 years old; overweight or obese; owned PS2/PS3; no active video games; played &gt; 2 hours/week</td>
<td>RCT</td>
</tr>
<tr>
<td>2011</td>
<td>Madsen</td>
<td>30 children</td>
<td>Children aged 9 to 18 years with a BMI above the 95th percentile who owned videogame consoles</td>
<td>RCT</td>
</tr>
<tr>
<td>2007</td>
<td>Maloney</td>
<td>60 children</td>
<td>Inclusion: 7-8 years old</td>
<td>RCT</td>
</tr>
<tr>
<td>2008</td>
<td>Maloney</td>
<td>60 children</td>
<td>Exclusion criteria: debilitating/chronic health problems, played Dance, Dance Revolution more than twice</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>Moore</td>
<td>20 participants</td>
<td>Inclusion: patients with pulmonary disease</td>
<td>RCT</td>
</tr>
<tr>
<td>2008</td>
<td>Ni Mhurchu</td>
<td>29 children</td>
<td>9-12 years</td>
<td>RCT- 6 weeks</td>
</tr>
<tr>
<td>2008</td>
<td>Ni Mhurchu</td>
<td>20 children</td>
<td>aged between 10 and 14 years; owned a PS 2; English speaking; and able to provide informed assent and parental consent</td>
<td>RCT</td>
</tr>
<tr>
<td>Authors</td>
<td>Number</td>
<td>Age (range)</td>
<td>Inclusion</td>
<td>Exercise</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Oka</td>
<td>60</td>
<td>30–76</td>
<td>Men and women over 30 years of age, well-compensated heart failure at baseline, with a diagnosis of heart failure &gt;3 months duration,</td>
<td>RCT</td>
</tr>
<tr>
<td>Owens</td>
<td>Adults (n=9), 37.8 +/- 4.9</td>
<td>8–13 years of age living at the same residence, no Wii console or Wii Fit currently in the home</td>
<td>RCT</td>
<td>3 months</td>
</tr>
<tr>
<td>Patel</td>
<td>15</td>
<td>14.7 ± 5.3</td>
<td>Inclusion: Pediatric heart transplant recipients</td>
<td>RCT</td>
</tr>
<tr>
<td>Paez</td>
<td>60</td>
<td>7- to 8-year-old children (n=60)</td>
<td>Exclusion criteria included individuals with significant somatic or played DDR, StepMania, more than twice before enrollment.</td>
<td>RCT</td>
</tr>
<tr>
<td>Plotnikoff</td>
<td>48</td>
<td>55 ± 12 Experimental: 55 ± 12 Control: 54 ± 12</td>
<td>Inclusion: Sedentary, obese individuals</td>
<td>RCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhodes</td>
<td>Adults (n=59, m=37.07 ± 6.56)</td>
<td>Inclusion: families with two parents and at least one child</td>
<td>RCT</td>
<td>6 weeks</td>
</tr>
<tr>
<td>Vestergaard</td>
<td>53</td>
<td>Inclusion: women over 75 years of age</td>
<td>RCT</td>
<td>Videotape, booklet and</td>
</tr>
</tbody>
</table>
2007

Experimental: 81 (3.3) n=25
Control: 82.7 (3.8) n =28

5 months
( first session supervised for safety)

Exercise bands
### Table 1B: Data Extraction of Experimental Studies: Instruments and Analysis

<table>
<thead>
<tr>
<th>Primary Author and Year</th>
<th>Relevant Primary Instruments</th>
<th>Correlations/Beta and Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baranowski 2012</td>
<td>i) accelerometer</td>
<td>Overall, intervention did not appear to have any significant changes</td>
</tr>
<tr>
<td></td>
<td>ii) Self-Report Wii use</td>
<td></td>
</tr>
<tr>
<td>Canning 2012</td>
<td>i) Self-Report</td>
<td>treadmill group appeared to demonstrate better exercise adherence, (2.6 session times/week) compared to the control group (1.25 session/week) over 6 weeks</td>
</tr>
<tr>
<td>French 2011</td>
<td>i) Self Report TV and computer use</td>
<td>Intervention effect (final and baseline)</td>
</tr>
<tr>
<td></td>
<td>ii) International Physical Activity Questionnaire</td>
<td><strong>Tv Viewing</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult, $B = -0.55$, $SE = 0.20$, $p = 0.01$</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TV is on</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult, $B = 0.9084$, $SE = 0.17$, $p &lt; 0.0001$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adolescent, $B = 0.9800$, $SE = 0.33$, $p = 0.005$</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TV on during meals</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult $B = -18.04$, $SE = 7.61$, $p = 0.02$</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>MVPA (adult)</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$B = 29.633$, $SE = 12.77$, $p = 0.02$</td>
</tr>
<tr>
<td>French 2012</td>
<td>i) Modified International Physical Activity Questionnaire</td>
<td>This was effective tactic for reducing TV viewing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>time only in adults $\beta = .55^*$</td>
</tr>
</tbody>
</table>
ii) Three-Day Physical Activity Recall

iii) TV Self-Report

Graves 2010
i) PA- self-report
ii) accelerometer

No intervention effects on PA variables at 6 and 12 weeks (p=0.17) and (p=.38)

Jakicic 1999
i). self-report
ii) accelerometer

the intervention group, SBEQ demonstrated greater adherence (87%) with moderate-strong effect size compared to the control SB (70.1%; h=.43). Although the number of exercise sessions/week of SBEQ (6.6+/ - 5.3) was significantly greater (p<0.001) than SB group (5.8 +/- 5.7) the effect size was small d=.14).

Khalil 2012
i) Exercise Diary

Adherence: three times over eight weeks, 53%

Maddison 2011
i) accelerometer
ii) self-report

The change in the average daily time spent playing active video games increased by 10 min (95% CI: 6.26, 13.81 min; p= 0.0001) at end of intervention compared with the Control group. The change in average daily time spent in non-active video games Decreased at 24 weeks in favour of intervention group, but was not significant (-9.39 min; 95% CI:- 19.38, 0.59 H89; P = 0.06).

Madsen 2007
i) Self-Report Diary

Few children in this study used DDR regularly, despite frequent telephone encouragement. Even among those who initially played frequently, playingDDR
<table>
<thead>
<tr>
<th>Study</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maloney</td>
<td>i) accelerometer, ii) SST - sedentary screen-time</td>
<td>There were no statistical differences between the intervention and the control groups in vigorous, moderate, light, or sedentary PA. However, there was a significant increase in vigorous PA in the intervention Group (from 10 ± 7.7 mpw to 16.2 ± 11.8 mpw, p &lt;0.0005). The DDR group had decrease in SST of -1.2 ± 3.7 h per week (hpw), p&lt;0.05, whereas the controls reported an increase of + 3.0 ± 7.7 hpw (non-significant). Difference in SST between the groups was significant, with less SST in the DDR group.</td>
</tr>
<tr>
<td>Moore</td>
<td>i) Self-Report</td>
<td>Adherence to DVD program: 69% at 6 weeks.</td>
</tr>
<tr>
<td>Ni Mhurchu</td>
<td>i) Self-Report TV watching, ii) Pedometer</td>
<td>When baseline TV viewing was controlled, there were no significant differences between groups at 6 weeks, F (1, 26)=.09, p=0.77. However, there was a decrease in total self-reported viewing hours (non-significant) in the intervention group compared to the change in the control group.</td>
</tr>
<tr>
<td>Ni Mhurchu</td>
<td>i) accelerometer, ii) Physical Activity Questionnaire for Children</td>
<td>Average PA time was higher in intervention group compared to the control group (difference at 6 weeks = 194 counts/min, p=0.04, 12 weeks=48 counts/min, p= 0.06)</td>
</tr>
<tr>
<td>Oka</td>
<td>i) Self-Report</td>
<td>Average total aerobic adherence for the 3 months of the exercise program was highest at 110%, upper body adherence was 87% and lower body exercise adherence was lowest at 75% for all 3 months. Twelve week adherence slowly decreased for all 3 components.</td>
</tr>
<tr>
<td>Owens</td>
<td>i) Actigraph GT1M accelerometer</td>
<td>No significant changes in physical activity</td>
</tr>
</tbody>
</table>
were observed in adults (p = 0.051) or children (p = 0.89).
82% reduction in minutes of daily Wii Fit during the second
6 weeks was statistically significant (p<0.01).

Paez 2009  i) accelerometer
       ii) DDR self-report usage
Paez participation was significant only absence of other video games
2009  OR 3.485 SE 1.427 p= 0.015

Plotnikoff 2010  i) Self-Report
Plotnikoff Mean adherence of their program was 71±22%;

Rhodes 2013  i) Home exercise log sheet
Rhodes For children, significant usage of game bikes was found
2013  (t36 = 2.61, P = .01, d = .85). NS for parents.

Vestergaard 2007  i) Dynamometer- hand grip
Vestergaard Adherence: 78 minutes/week for five months, 89.2%
2007  ii) Isobex medical device- Bicep strength
       iii) Health Related Quality of Life
<table>
<thead>
<tr>
<th>Primary Author and year</th>
<th>Participants</th>
<th>Inclusion Criteria</th>
<th>Length/Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adachi-Mejia 2007</td>
<td>2 343 child/parent pairs 9-12 years of age</td>
<td>BMI greater or equal to 95th percentile age range, 9-12</td>
<td>cross-sectional</td>
</tr>
<tr>
<td>Atkinson 2005</td>
<td>102 adults (age = 48.2 +/- 11.6 years)</td>
<td>Age 18-65, absence of medical condition</td>
<td>prospective 4 years</td>
</tr>
<tr>
<td>Barr-Anderson 2007</td>
<td>781 adolescents</td>
<td>Not clear, previous study cannot be found.</td>
<td>prospective 5 years</td>
</tr>
<tr>
<td>Bauer 2011</td>
<td>253 parent/adolescent girls age= 15.7 years, range = 14-20.3</td>
<td>Girls with physical activity levels less than one hour/day.</td>
<td>cross-sectional</td>
</tr>
<tr>
<td>Dennison 2002</td>
<td>2 761 adults with children</td>
<td>Adults with children aged 1 through 5 years participating at a nutritional program</td>
<td>cross-sectional</td>
</tr>
<tr>
<td>Dunton 2003</td>
<td>87 girls, 14-17 years old (age = 15.02 +/- .72 years)</td>
<td>(1) failure to meet the minimum physical activity recommendations by the American College of Sports Medicine 2) performance at or below</td>
<td>cross-sectional</td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Methods</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Gorin, 2011</td>
<td>overweight (n=201) normal weight (n=213) adults</td>
<td>Not specified</td>
<td></td>
</tr>
<tr>
<td>Hoyos Cillero, 2010</td>
<td>247 primary school children 256 secondary school children</td>
<td>Not specified</td>
<td></td>
</tr>
<tr>
<td>Jakicic, 1997</td>
<td>194 adults 98 men, 96 women</td>
<td>University faculty and staff</td>
<td></td>
</tr>
<tr>
<td>Kerr, 2008</td>
<td>853 parent/child dyad 878 adolescents 853 parents</td>
<td>exclusion: health conditions</td>
<td></td>
</tr>
<tr>
<td>Liao, 2014</td>
<td>118 adults (age 27-73)</td>
<td>Exclusion: did not speak English household income greater than $210 000 physical disabilities which limited PA</td>
<td></td>
</tr>
<tr>
<td>Maddison, 2009</td>
<td>110 students 12-17 years of age (age= 14.6 +/- 1.55)</td>
<td>age: 12-17 years</td>
<td></td>
</tr>
<tr>
<td>Patnode, 2010</td>
<td>294 youth/parent pairs adolescents: 10-17 years of age</td>
<td>no specific inclusion criteria mentioned</td>
<td></td>
</tr>
</tbody>
</table>
age = 15.4 +/- 1.7

Reed
2005
411 university students
not specified
cross-sectional

Ries
2009
249 adults
participating in moderate or vigorous physical activity for 90 min or less
prospective- 1 year

Roemmich
2007
88 children
greater than 15 hours/week of TV/computer BMI less than 90th percentile
cross-sectional

Salmon
2013
613 children
47% boys (age 9.4 +/- 2.2 years)
Exclusion if parent was <17 or >46 And children <5 or >13
Prospective- 1 year

Spurrier
2008
280 households
(mean = 4.8 years ± 0.21)
Children less than five and a half years of age
Cross-sectional

Sirard
2010
613 parent/adolescent dyads
Health Partners members, in grades 6th through 11th in the fall of 2007, residing in one of the randomly selected middle or high-school districts included in the sample.
cross-sectional

Stucky-ropp
1993
121 girls
5th and 6th grades
cross-sectional

121 boys
(age = 11.2 +/- 0.7)

Tang
2009
2684 children
Children grades 6-9
Cross-sectional

(age- 11-16)
<table>
<thead>
<tr>
<th>Author</th>
<th>N</th>
<th>Description</th>
<th>Sampling Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Dyck</td>
<td>1200 adults (age 20-65) 48% males</td>
<td>None-stratified random sampling Of 24 neighborhoods</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Zutphen</td>
<td>1926 children 4-12 years of age</td>
<td>ages 4-12 and part of a nutrition program</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wethington</td>
<td>23 145</td>
<td>Data from 2007 National Survey of Children’s Health</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams</td>
<td>205 adults</td>
<td>Less than 90 minutes per week of MVPA</td>
<td>Prospective-1 year</td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wong</td>
<td>29 139 children</td>
<td>Secondary schools</td>
<td>Prospective-1 month</td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wong</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2B: Data Extraction of Observational Studies: Instruments and Analysis

<table>
<thead>
<tr>
<th>Primary Author and year</th>
<th>Relevant Primary Instruments</th>
<th>Correlations/Beta and Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adachi-Mejia 2007</td>
<td>i) Self-Report custom questionnaire</td>
<td>Boys more likely to have televisions in their bedrooms (50.3%) compared to girls (46.2%) (p&lt;0.05)</td>
</tr>
<tr>
<td>Atkinson 2005</td>
<td>i) GLTEQ ii) Accelerometer</td>
<td>Quantity of home exercise equipment was correlated with self-reported total ( r = .34^* ) and vigorous leisure-time physical activity ( r = .27^* )</td>
</tr>
</tbody>
</table>
| Barr-Anderson 2007      | i) Modified Leisure Time Exercise Questionnaire ii) Custom items for Sedentary behavior | Sedentary Time for: i) TV in bedroom and ii) no TV in bedroom  
  1. Girls  
  i) 20.7+/-0.78, ii) 15.2 +/-0.90, p=.001  
  2. Boys  
  i) 22.2+/-0.78, ii) 18.2+/-1.16, P= 0.005  
Vigorous PA for i) TV in bedroom and ii) no TV in bedroom  
  1. Girls  
  i) 1.8+/-0.17, ii) 2.5+/-0.20, p=.004  
  2. Boys  
  i) 3.8 +/- 0.17, ii) 3.8 +/- 0.25, p<0.04 |
<table>
<thead>
<tr>
<th>Author</th>
<th>Study Title</th>
<th>Results/Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bauer</td>
<td>i) Family Physical Activity Environment</td>
<td>Total PA and Home PA resources, $r = 0.51$, $p = .67$</td>
</tr>
<tr>
<td>2011</td>
<td>ii) Family Television Use Environment</td>
<td>MVPA and Home PA Resources, $r = 0.41$, $p = 0.063$</td>
</tr>
<tr>
<td></td>
<td>iii) 3-Day Physical Activity Recall</td>
<td></td>
</tr>
<tr>
<td>Dennison</td>
<td>i) Parent report on TV details</td>
<td>Children with a TV set in their bedroom, compared with those without, spent an additional 4.6 hours per week ($p&lt;0.0001$)</td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dunton</td>
<td>i. Modified Perceived Environments Related to Physical Activity instrument</td>
<td>1. home use availability and home use frequency, $r = 0.216$, $p &lt; 0.05$</td>
</tr>
<tr>
<td>2003</td>
<td>ii. Stanford Usual Physical Activity Scale</td>
<td>2. home use frequency and home use variety, $r = 0.667$, $p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>iii. 2 Day Physical Activity Recall</td>
<td>3. Fitness and home availability, $r = 0.224$, $p &lt; 0.05$</td>
</tr>
<tr>
<td></td>
<td>iv. VO2 max test</td>
<td></td>
</tr>
<tr>
<td>Gorin</td>
<td>i) Exercise Environment Questionnaire</td>
<td>In normal weight group, physical activity was associated with aerobic equipment available ($P=.02$). Positive correlation between number of TVs and viewing ($r=.25$, $P&lt;.001$). TVs in bedroom led to significantly longer viewing times ($p&lt;.001$)</td>
</tr>
<tr>
<td>2011</td>
<td>ii) S-R TV use and access</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) Paffenbarger Physical Activity Questionnaire</td>
<td></td>
</tr>
<tr>
<td>Hoyos Cillero</td>
<td>i) TV and media scale-custom</td>
<td>Older females- bedroom TV ($OR:0.32$, $p&lt;.05$), and bedroom console ($OR: 0.26$, $p&lt;0.05$) correlated with &gt;2hrs/day but not for males. More males in secondary school group than females had a console</td>
</tr>
</tbody>
</table>
Younger females exceeding TV guidelines had 2 TV sets in the home.

**Jakicic**

1997

i) Paffenbarger Questionnaire

ii) custom PA environment scale

For women, there were significant correlations between total activity and both recreational equipment ($r=0.22$) and total amount of exercise equipment ($r=0.25$), individual sports equipment ($r=0.20$) and total exercise equipment ($r=0.030$)

Men- team sport equipment and PA ($r=0.20$)

Women- individual sport ($r=0.24$), home equipment ($r=0.24$), total equipment ($r=0.28$) and recreation equipment ($r=0.27$)

**Kerr**

2008

i) Exercise equipment checklist

ii) International Physical Activity Questionnaire

iii) 7 day Physical Activity Recall

The presence of more home-use exercise equipment was related to physical activity in adolescent girls (OR = 1.27, 95% CI = 1.1–1.5). There was also a significant interaction between perceived safety and equipment ($P < 0.01$)

**Liao**

2014

i. accelerometer

ii. electronic momentary assessment

Home was the most important context for Physical and sedentary activity

Women engaged in PA more when they were at home (pred. prob.=.61, SE=.061) than men (pred. prob.=.24, SE=.077)

When at home, men spent more SD time than women whereas when at work, women spent more SD time than men ($p<0.05$)

**Maddison**

2009

i) Perceived ownership and reported use of equipment

ii) accelerometer

iii) Physical Activity Questionnaire for Adolescents (PAQ-A),

Of the perceived variables, home ownership of recreation equipment (standardized effect = .26) had a direct effect on physical activity.
Patterns of Accelerometer

Patnode 2010 i) accelerometer

ii) International Physical Activity Questionnaire

iii) Physical Activity and Media Inventory

Home PA equipment was positively correlated with MVPA $r = 0.21$ ($p < 0.001$)

Reed 2005 Questionnaire which measured exercise intensity, frequency, duration

Female- Quantity of home exercise equipment and:

i) total physical activity $r = .247$, $p < 0.05$

ii) intensity $r = .332$, $p < 0.05$

iii) frequency $r = .166$, $p < 0.05$

iv) duration $r = .310$, $p < 0.05$

Male- Quantity of home exercise equipment and:

i) total physical activity $r = .039$, $p > 0.05$

ii) intensity $r = .009$, $p > 0.05$

iii) frequency $r = .093$, $p > 0.05$

iv) duration $r = .91$, $p > 0.05$

Ries 2009 i) 7-day physical activity recall

ii) Home Environment Scale

Associations between home equipment availability and minutes of physical activity.

A) SIM $B = 5.07$, SE = 1.83, $p < 0.01$

B) STRIDE $B = 3.52$, SE = 10.2, $p < 0.001$

Roemmich 2007 i) Self-Report

ii) accelerometer

number of televisions in the home was correlated to television watching time ($r = .31$, $p \leq .01$).
Salmon 2013

i) custom items for television and PA equipment

ii) Self-report by parents

TV in child’s bedroom and screen-time, B=1.5 (0.8, 2.3), p<0.005

Home PA equipment and sedentary time, B=-3.4 (-5.9, -0.9), P<.01

Neighbourhood road safety concerns moderated screen-time behavior.

B=.9 p<0.054

Sirard 2010

i) Physical Activity and Media Inventory

PAASS (Physical Activity Availability and Accessibility Summary Score)

MAASS: Media Availability and Accessibility Summary Score

ii) accelerometer

iii) Self-Report screen time

Accelerometer MVPA and

(males)

i) PA equipment density r= 0.13, p<0.05;

ii) PAAS, r= 0.15, p<0.05;

(females)

i) PA equipment density r= 0.16, p<0.05

ii) PAAS, r= 0.19, p<0.05

S-R Screen Time and

males- MAASS, r= 0.07, p> 0.05

females- MAASS, r=0.12, p<0.05

Accelerometer and

i) PA Density, B=1.17, SE=0.42, p<0.01

ii) PAASS B=0.03, SE=0.01, p<0.01

Televisions in girl's bedrooms positively associated with time and the ratio of activity (r=0.17, p<0.05)
<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spurrier 2008</td>
<td>i) The Physical and Nutritional Home Environment Inventory</td>
<td>Backyard size ($p = 0.001$) and outdoor play equipment (p = 0.003) were associated with outdoor play</td>
</tr>
<tr>
<td></td>
<td>ii) The Outdoor Playtime and Small Screen Entertainment Checklist</td>
<td>Presence of playstation ($p&lt;0.02$) was associated with SD time</td>
</tr>
<tr>
<td>Stucky-ropp 1993</td>
<td>i) The Physical Activity Interview</td>
<td>Number of exercise related items at home predicted PA for girls. Not significant for boys, no data reported. B= 0.26, $R^2=0.08$, $F=5.06$, $p = 0.008$</td>
</tr>
<tr>
<td></td>
<td>ii) Children's Physical Activity Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) Parental Physical Activity Questionnaire</td>
<td></td>
</tr>
<tr>
<td>Tang 2009</td>
<td>i) The Adolescent Physical Activity Recall Questionnaire</td>
<td>Backyard predicted less SD time (OR = 0.7, 95% CI = 0.6–0.9)</td>
</tr>
<tr>
<td></td>
<td>ii) Home Environment Questionnaire-custom</td>
<td></td>
</tr>
<tr>
<td>Van Dyck 2011</td>
<td>i) IPAQ</td>
<td>PA equipment in home environment (CI=.057, 0.115) was associated with vigorous leisure-time PA</td>
</tr>
<tr>
<td></td>
<td>ii) NEWS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) accelerometer</td>
<td></td>
</tr>
<tr>
<td>Van Zutphen 2007</td>
<td>Self-Report measure (not validated)</td>
<td>Children who had televisions in their bedrooms significantly watched more TV than their counterparts ($p&lt;0.001$)</td>
</tr>
<tr>
<td>Veitch 2010</td>
<td>i) accelerometer</td>
<td>One third of children play in yard per week</td>
</tr>
<tr>
<td></td>
<td>ii) environment scale-custom</td>
<td>“stranger danger” predicted yard play OR=2.32, $p&lt;0.05$</td>
</tr>
<tr>
<td>Wethington 2013</td>
<td>Custom PA and screen-time scales</td>
<td>TV in bedroom and screen-time, OR 1.7 (1.4 to 2.1)</td>
</tr>
</tbody>
</table>
Williams 2008  i) 7-day Physical Activity Recall  PA equipment at home predicted PA adoption (OR=1.73; 95% CI: 1.05, 2.85), but not PA maintenance (OR=0.88; 95% CI: 0.58, 1.35)

ii) Environment Assessment Scale

Wong 2010  Custom PA and screen-time scales  Access to sport facilities were more likely to be physically Active (ORboys=1.26; ORgirls=1.34), while those who additionally Reported computer/internet use were less likely to be physically active (ORboys=.60; ORgirls=.54)
Table 3A: Evaluation of Experimental Studies

<table>
<thead>
<tr>
<th></th>
<th>Baranowski</th>
<th>Canning</th>
<th>French</th>
<th>French 2012</th>
<th>Graves</th>
<th>Jakicic</th>
<th>Juneau</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 2</td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
<td>II. 5</td>
<td>II. 5</td>
<td>II. 5</td>
</tr>
<tr>
<td><strong>B) 1</strong></td>
<td>B) 5</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
</tr>
<tr>
<td><strong>1- strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>3-weak</strong></td>
<td><strong>1- moderate</strong></td>
<td><strong>3-weak</strong></td>
<td><strong>1- strong</strong></td>
</tr>
<tr>
<td><strong>C) 2</strong></td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 3</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
</tr>
<tr>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>2- moderate</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
</tr>
<tr>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 3</td>
</tr>
<tr>
<td><strong>2-moderate</strong></td>
<td><strong>3-weak</strong></td>
<td><strong>2- moderate</strong></td>
<td><strong>2- moderate</strong></td>
<td><strong>2-moderate</strong></td>
<td><strong>2-moderate</strong></td>
<td><strong>2-moderate</strong></td>
<td><strong>2-moderate</strong></td>
</tr>
<tr>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
<td><strong>E) I. 1</strong></td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
</tr>
<tr>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1- strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
<td><strong>1-strong</strong></td>
</tr>
<tr>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
<td><strong>F) I. 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 1</td>
<td></td>
</tr>
<tr>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>2-moderate</td>
<td>1-strong</td>
<td></td>
</tr>
<tr>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td></td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 1</td>
<td>II. 1</td>
<td></td>
</tr>
<tr>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>1-strong</td>
<td></td>
</tr>
<tr>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td></td>
</tr>
<tr>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td></td>
</tr>
</tbody>
</table>

Note. The Cochrane Collaboration Risk of Bias Tool was used to assess these studies. Please see Table 5A for summary of items.
Table 3B: Evaluation of Experimental Studies

<table>
<thead>
<tr>
<th>Khalil</th>
<th>Maddison</th>
<th>Madsen</th>
<th>Maloney</th>
<th>Moore</th>
<th>Ni Mhurchu</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) I.1</td>
<td>A) I.1</td>
<td>A) I.1</td>
<td>A) I.2</td>
<td>A) I.1</td>
<td>A) I.1</td>
</tr>
<tr>
<td>II. 5</td>
<td>II. 5</td>
<td>II. 1</td>
<td>II. 5</td>
<td>II. 1</td>
<td>II. 5</td>
</tr>
<tr>
<td>1-strong</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>2-moderate</td>
</tr>
<tr>
<td>B) 5</td>
<td>B) 1</td>
<td>B) 3</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
</tr>
<tr>
<td>3-weak</td>
<td>1-strong</td>
<td>1-weak</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1- strong</td>
</tr>
<tr>
<td>C) 1</td>
<td>C) 2</td>
<td>C) 3</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 1</td>
</tr>
<tr>
<td>3-weak</td>
<td>1-strong</td>
<td>1-weak</td>
<td>1-strong</td>
<td>1-strong</td>
<td>3- weak</td>
</tr>
<tr>
<td>D) I.1</td>
<td>D) I.1</td>
<td>D) I.1</td>
<td>D) I.1</td>
<td>D) I.1</td>
<td>D) I.1</td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 3</td>
</tr>
<tr>
<td>2-moderate</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>2- moderate</td>
</tr>
<tr>
<td>E) I.3</td>
<td>E) I.1</td>
<td>E) I.1</td>
<td>E) I.1</td>
<td>E) I.3</td>
<td>E) I.1</td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
<td>II. 1</td>
</tr>
<tr>
<td>3- weak</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>3-weak</td>
<td>1- strong</td>
</tr>
<tr>
<td>F) I.2</td>
<td>F) I.1</td>
<td>F) I.1</td>
<td>F) I.3</td>
<td>F) I.1</td>
<td>F) I.3</td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 1</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>1-strong</th>
<th>2-moderate</th>
<th>2-moderate</th>
<th>2-moderate</th>
<th>2-moderate</th>
<th>1-strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 3</td>
</tr>
<tr>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>2-moderate</th>
<th>2-moderate</th>
<th>1-strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 3</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
</tr>
<tr>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1-strong</th>
<th>1-strong</th>
<th>3-weak</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
<td>Weak</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Note. The Cochrane Collaboration Risk of Bias Tool was used to evaluate these studies. Please see Table 5A for summary of items.
Table 3C: Evaluation of Experimental Studies

<table>
<thead>
<tr>
<th>Ni Mhurchu 2009</th>
<th>Oka</th>
<th>Owens</th>
<th>Paez</th>
<th>Plotnikoff</th>
<th>Rhodes</th>
<th>Vestergaard</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
<td>A) I. 1</td>
</tr>
<tr>
<td>II. 5</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
<td>II. 5</td>
<td>II. 3</td>
</tr>
<tr>
<td>3-weak</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>2-Moderate</td>
<td>1-strong</td>
</tr>
<tr>
<td>B) 1</td>
<td>B) 5</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
<td>B) 1</td>
</tr>
<tr>
<td>1-strong</td>
<td>3-weak</td>
<td>3-weak</td>
<td>1-strong</td>
<td>3-weak</td>
<td>1-strong</td>
<td>1-strong</td>
</tr>
<tr>
<td>C) 1</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
<td>C) 2</td>
</tr>
<tr>
<td>2-moderate</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
</tr>
<tr>
<td>D) I. 3</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
<td>D) I. 1</td>
</tr>
<tr>
<td>II. 3</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
<td>II. 2</td>
<td>II. 3</td>
</tr>
<tr>
<td>2-moderate</td>
<td>3-weak</td>
<td>3-weak</td>
<td>2-moderate</td>
<td>3-weak</td>
<td>2-moderate</td>
<td>2-moderate</td>
</tr>
<tr>
<td>E) I. 1</td>
<td>E) I. 1</td>
<td>E) I. 1</td>
<td>E) I. 1</td>
<td>E) I. 1</td>
<td>E) I. 1</td>
<td>E) I. 1</td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
</tr>
<tr>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>1-strong</td>
<td>1-strong</td>
</tr>
<tr>
<td>F) I. 2</td>
<td>F) I. 2</td>
<td>F) I. 1</td>
<td>F) I. 2</td>
<td>F) I. 1</td>
<td>F) I. 1</td>
<td>F) I. 3</td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 1</td>
<td>II. 3</td>
</tr>
<tr>
<td></td>
<td>2-moderate</td>
<td>3-weak</td>
<td>1-strong</td>
<td>2-moderate</td>
<td>1-strong</td>
<td>1-strong</td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>--------</td>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
<td>G) I. 1</td>
</tr>
<tr>
<td>II. 1</td>
<td>II. 2</td>
<td>II. 3</td>
<td>II. 3</td>
<td>II. 3</td>
<td>II. 3</td>
<td>II. 1</td>
</tr>
<tr>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
<td>III. 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
<td>H) I. 1</td>
</tr>
<tr>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
<td>II. 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
<th>1-strong</th>
</tr>
</thead>
</table>

| Moderate | Weak | Weak | Strong | Weak | Strong | Moderate |

Note. The Cochrane Collaboration Risk of Bias Tool was used to evaluate these studies. Please see Table 5A for summary of items.

Table 4A: Evaluation of Observational Studies
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
<td>0-unable</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
<td>0-unable</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>some</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0-unable</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12</strong></td>
<td><strong>10</strong></td>
<td><strong>11</strong></td>
<td><strong>13</strong></td>
<td><strong>13</strong></td>
<td><strong>13</strong></td>
<td><strong>9</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
<td><strong>11</strong></td>
</tr>
</tbody>
</table>

Note. The modified version of the Downs and Black's 22-item assessment total used to assess these studies. Please see Table 5B for summary of items.
<table>
<thead>
<tr>
<th>Question</th>
<th>Maddison</th>
<th>Patnode</th>
<th>Patnode</th>
<th>Reed</th>
<th>Ries</th>
<th>Roemmich</th>
<th>Salmon</th>
<th>Sirard</th>
<th>Spurrier</th>
<th>Stucky-rop</th>
<th>Tang</th>
<th>Timperio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0-unable</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>almost</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13</strong></td>
<td><strong>13</strong></td>
<td><strong>13</strong></td>
<td><strong>10</strong></td>
<td><strong>10</strong></td>
<td><strong>14</strong></td>
<td><strong>13</strong></td>
<td><strong>15</strong></td>
<td><strong>10</strong></td>
<td><strong>11</strong></td>
<td><strong>12</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>

Note. The modified version of the Downs and Black’s 22-item assessment tool used to assess these studies. Please see Table 5B for summary of items.

Table 4C: Evaluation of Observational Studies
<table>
<thead>
<tr>
<th>Question</th>
<th>Van Zutphen</th>
<th>Veitch</th>
<th>Wethington</th>
<th>Williams</th>
<th>Wong</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0-unable</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>large drop out</td>
<td>0</td>
<td>0-unable</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9</strong></td>
<td><strong>11</strong></td>
<td><strong>13</strong></td>
<td><strong>13</strong></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>

Note. The modified version of the Downs and Black's 22-item assessment tool used to assess these studies. Please see Table 5B for summary of items.

Appendix- Table 1A: Summary items from The Cochrane Collaboration Risk of Bias Tool
A) Selection Bias

I) Sample representative of population

1. Very likely
2. Somewhat likely
3. Not likely
4. Can’t tell

II) Percentage participants agreed to participate

1. 80 - 100% agreement
2. 60 – 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can’t tell

B) Study Design

1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify ____________________________
8. Can’t tell

C) Were there important differences between groups prior to the intervention?
1. Yes
2. No
3. Can’t tell

D) Blinding
I) Assessors were aware

1. Yes
2. No
3. Can’t tell

II) Participants were aware

1. Yes
2. No
3. Can’t tell

E) Data Collection Methods
I) Validity of tools - Were tools valid?

1. Yes
2. No
3. Can’t tell

II) Reliability of tools

1. Yes
2. No
3. Can’t tell
F) Withdrawals and drop-outs

I) Drop-outs Reported
1. Yes
2. No
3. Can't tell

II) Percentage of completion
1. 80 - 100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

G) Intervention Integrity

I) Percentage participant received allocated intervention
1. 80 - 100%
2. 60 - 79%
3. less than 60%
4. Can't tell

II) Measurement of consistency
1. Yes
2. No
3. Can't tell
III) Potential contamination

1. Yes
2. No
3. Can’t tell

H) Analyses

I) Statistical Methods appropriate

1. Yes
2. No
3. Can’t tell

II) Intervention allocation status rather than actual intervention received

1. Yes
2. No
3. Can’t tell

Note. Based on selection rules provided by the instrument directory, each section letter was evaluated as 1) strong, 2) moderate or 3) weak

Appendix- Table 1B: The modified version of the Downs and Black’s 22-item assessment tool
1. Hypothesis and aim/objectives clearly described
2. Definitions of PA constructs that are validated are clearly described
3. Participants described
4. Confounders described
5. Missing/incomplete data described
6. Main findings clearly described
7. Information provided about variability of data
8. Effect size reported
9. Recruitment sample representative of population
10. Participants’ representative of population
11. Appropriate statistical tests used
12. Validation of self-report measure
13. Adjustment for confounding analyses
14. Compliance acceptable

Note. This Table provides a summary of items of the modified version of Downs and Black’s 22-item assessment tool. Items were assessed as 1=yes, 0= no, or unable to determine
Potentially relevant publications identified and screened from electronic databases (Scopus, PsychInfo, SportDiscus, Medline, PubMed): 3265

Total Duplicates Removed: 1137

Potentially relevant citations Screened: 2128

Papers excluded after evaluation of abstract: 2032

1381- other design and methods (ie. qualitative, telephone counselling, absence of correlate between physical environment variables and PA/Sedentary time, mixed methods-ie. home and lab intervention)
332- Other topics/ disciplines
215- Non home (school, hospital, nursing home, etc)
53- abstract/books/thesis
15- Reviews (1 relevant-kept aside)
31- non-academic articles
5- non English

Remaining papers for detailed evaluation: 96
Total Papers Reviewed: 49

Papers excluded after reviewing: 54
27- no correlation between physical environment variables and PA/Sedentary time
23- no measurement of behaviour change/unclear, poorly designed
4- Non home (school, hospital, nursing home, etc)

Papers added after manual cross-referencing: 7
References


Appendix 4: A Brief Overview of the Dual Process Approach in Predicting Physical Activity
Overview

The majority of research used to understand human behavior has been based on conscious regulatory models which propose that behavior is a volitional and reflective process (Sheeran et al., 2013). In the PA literature, some of the most commonly used psychological models: The Protection Motivation Theory (W. Rogers, 1974), Health Belief Model (Rosenstock, 1974), Theory of Reasoned Action (Fishbein, 1975), Self-Determination Theory (E. L. Deci & Ryan, 1985) and Social Cognitive Theory (Bandura, 1986) and Theory of Planned Behavior (Ajzen, 1991) have demonstrated their utility in understanding PA behavior over the past few decades (Linke et al., 2014; Rhodes & Nigg, 2011) and have helped identify key correlates which include intention, affective attitude, and perceived behavioral control, (Rhodes & De Bruijn, 2013).

However, some theorists have proposed that not all behavior is a reflective process as denoted by the Unconscious Thought Theory (UTT) (Dijksterhuis & Nordgren, 2006) and Dual process approach (DPA) (Evans, 2008b; Evans & Stanovich, 2013). Although both of these perspectives propose that the unconscious process carries a significant contribution to our behavioral choices, they differ in the interpretation of the unconscious process. The UTT theorizes that the unconscious provides cognition without our awareness whereas the DPT proposes that unconscious works in a reflexive manner (Bargh, 2011; Dijksterhuis & Nordgren, 2006). The DPA surmises that our behavior is influenced by a combination of both unconscious (rapid, automatic) and conscious (slow, deliberative) processes. Despite the changes in nomenclature of these systems over the past thirty years which include automatic and controlled (Schneider & Shiffrin, 1977), stimulus bound vs. higher order (Toates, 2006), Type 1 and Type 2
systems (Evans & Stanovich, 2013), the agreement on definitions and distinction between two systems have remained relatively consistent (Evans, 2008b). Overall, the preferred and most commonly used labels is Type 1 and Type 2 (Stanovich, 1999) systems, as suggested by Evans (2013). For instance, learning to drive will require Type 2 system for processing new information (e.g. driving coordination, traffic cues, new routes). However, repeatedly performing the behavior under familiar contexts such as driving to work every morning will allow the process to gradually become under control of the Type 1 system, or become automatic. The initiation of the goal “go to work” would be followed by a sequence of automatic processes, this has been defined as goal driven automaticity (B. Verplanken et al., 1997). Type 1 is assumed to yield default responses unless intervened on by distinctive higher order reasoning processes (Type 2). Type 2 provides hypothetical thinking and load heavily on working memory (Evans 2013. Driving home in a familiar route can be controlled by Type 1, however, novel stimuli such as construction site would require following new directions and processing new information in which Type 2 would then be activated. Figure 1 illustrates the basic functionality of the dual process approach.

Figure 1: Dual Process Approach

![Dual Process Approach Diagram](image-url)
Note. Type 1 (unconscious process) and Type 2 (conscious process) both provide input in predicting behavior.

Relationship between the two systems: Synergistic and Antagonistic Processes

Interestingly, type 1 and 2 systems can work either synergistically or in an antagonistic manner on behavior. The synergistic work of these systems can be referred to as goal-derived automaticity (Aarts & Dijksterhuis, 2000b; Bargh & Ferguson, 2000; Wood & Neal, 2007). For instance, when solving a problem that requires prior knowledge, the solution can just “come to mind” depending on its relevance to the context (Evans & Stanovich, 2013). In general, it is hypothesized that the systems work complementary together so that Type 1 can handle low-effort processing/familiar background tasks while the Type 2 mind is dedicated to novel stimuli or situations which demand controlled attention. This can be exemplified in a driving scenario; when an individual encounters a construction site while driving, the attention will be directed to reading new road signs/directions. However, the process of driving itself would remain automatic but its contribution would be significantly reduced (eg. more awareness in steering or braking).

Alternatively, the systems can work against each other when the outcome goal is not in agreement with strong unconscious drives, such as breaking a bad habit (Anshel et al., 2010). In most scenarios the outcome can be harmless, such as when an individual forgets to deviate from one’s habitual driving route from home (despite possessing an initial intention to stop along the way). However, a strong bad habit can have much more serious consequences such as when a compulsive gambling habit continues to destroy an individual’s life (Evans, 2014b). Although
the “folk psychological notion” of irrationality carries a perspective that people simply try and fail to achieve their goals, the conscious mind constructs an illusion of free will which justifies the outcome (Bargh & Ferguson, 2000) such as matching bias.

**Early Evidence of Dual Process Approach**

Irrational behaviors can be seen as unconscious influences with emotions (eg. affect), or habits and intuitions which override the conscious rationality (Evans, 2014b). Some of the earliest evidence for the dual process approach was from a study that examined the interference between type 1 and type 2 systems (Wason, 1960). The experiment involved presenting a sequence of numbers to participants based on a particular rule. The participants were instructed to decipher the rule by providing sets of numbers which they believed to match the rule. After presenting each set of sequences, the researcher would then tell them if their sequence matched the rule or not. The participants were allowed to provide as many sequences as they desired and each was evaluated by the researcher. Once they felt comfortable, the participant would then announce the rule. The study found that only 6/29 participants figured out the rule on their first announcement. Wason concluded that participants were irrational as they failed to falsify their own hypothesis during the process, which has been later found to be confirmation or matching bias (Evans, 2014a). The participants basically searched for evidence to support their hypothesis rather than seeking for evidence that could falsify their belief. Hence, participants only provided scenarios in which they believed to match their belief of the rule rather than testing against it to eliminate their believed rule. Evans (2014a) proposes a dual-process hypothesis in which unconscious type 1 processes (matching bias) determined the choices; the type 2 reasoning
processes then provided rationale for the choice. This reasoning essentially matches the
description Wason (1960) provided to explain the participants behavior.

**Dual Process Approach and Physical Activity**

The application of Dual Process approach to predict PA behavior is very limited to a
handful of papers which may not have clearly identified their approach as Dual process, but
nonetheless have compared unconscious vs. conscious constructs to predict PA behavior. Six
studies were found to analyze conscious motivation (e.g., intention) with habit. Of these studies,
four showed support for dual process approach (Conroy et al., 2010; R. E. Rhodes & G. J. De
Bruijn, 2010; Rhodes et al., 2010; B. Verplanken et al., 1997) and two did not (de Bruijn &
Rhodes, 2011; Rebar et al., 2014). Four of these studies particularly investigated the interaction
of intention and habit to predict behavior (de Bruijn & Rhodes, 2011; R. E. Rhodes & G. J. De
Bruijn, 2010; Rhodes et al., 2010; B. Verplanken et al., 1997). Two of these studies found a
significant habit x intention interaction which indicated that intentions were only significantly
related to behavior when habit was weak (de Bruijn & Rhodes, 2011; B. Verplanken et al.,
1997). Generally these results have demonstrated a synergistic relationship between the two
processes, though some studies have shown nonparallel functionality such as intention
significantly predicting the behavior despite weak habits (de Bruijn et al., 2009; de Bruijn &
Rhodes, 2011). Although PA behavior can be performed without habit, the repetitive conscious
effort becomes an extra burden for delivering the behavior which could deteriorate motivation
over time. For instance Kaushal & Rhodes (2015) found that those with stronger habits displayed
much better exercise adherence than their counterparts. They also found habit and intention to
be almost parallel predictors for exercise adherence in their trajectory analysis (\(\beta=.227\), and \(\beta=.232\) respectively).

**Limitations and Future Directions**

Limited findings suggest that synergistic functionality of intention and habit could be effective in predicting exercise adherence. Although these findings provide novel insight into understanding exercise behavior, randomized-controlled trials (RCTs) that apply dual process approach are non-existent. This is a critical shortcoming as an RCT which is able to strengthen a participants’ intention and habit could be a key indicator of a successful intervention for long-term exercise adherence. Although longitudinal designs can help identify *what* causes the two processes change, a critical question of if we can effectively manipulate the change to facilitate exercise adherence cannot be addressed. Hence, analyzing results with a control group in an RCT can help provide insight as to *how* can we implement an exercise habit.
Appendix 5: The Unconscious Thought Theory (Definition)

The UTT (Dijksterhuis & Nordgren, 2006) was developed from the concept of incubation (i.e., you suddenly have a good idea or know what to do regarding an important decision when your conscious mind is focused on another task). Due to uncontrollable multiple processes at work (independent on deliberate cognitive focus), it was hypothesized that the unconscious system consists of parallel instead of serial processes (Dijksterhuis & Nordgren, 2006). It has been theorized that the unconscious system has higher capacity than conscious regulated processes (Bargh, 2011; Dijksterhuis & Nordgren, 2006). For instance, while an individual is consciously focusing on a particular thought or problem, certain components could be associated with previous cues which would prompt other memories without deliberation of those thoughts.
References


Appendix 6: Study III: Worksheets

Exercise Habit Formation Guide

Goal: To exercise at least 4 bouts per week accumulating 150 minutes or more of moderate-to-vigorous physical activity

Fill out the following sections and keep this guide where it is visible to you (e.g. on the fridge, on your office wall, etc).

Consistency (temporal and behavioral pattern)

My workout time will be at ________ am/pm

The two activities prior to my workout will be __________ ________, workout

Examples- (evening): come home from work, eat oatmeal, workout
                          : pick up kids, prepare dinner, workout
                          (morning): brush teeth, eat breakfast, workout

Ease

Keeping your workout materials prepared and at the same location can help make your preparation simple and efficient.

IV. Cues/ Ritual Associations:

Cues work best when they are activated until the behavior is performed then it is turned off. Ig. traffic lights. Cues can be an alarm on your phone, clock, or visual such as running shoes by the door. Placing a cue out in the morning/before you go to bed will remind you to exercise. When you complete your exercise then put the cue out of sight.

Example- in the morning take out a water bottle from your cupboard and place it on your desk, it will remain on your desk until you take it to your workout. After your evening workout you return it to the cupboard. Repeat steps next morning.

Before you leave work select your favourite gym clothes from your closet and place them on your bed. Wear the gym clothes for your workout then put them back in your closet.
The following behaviors will act as my cues to exercise:

1.

2.

3.
Exercise Variety Guide

Goal: To exercise at least 150 minutes or more per week at a moderate-to-vigorous level

Adding Variety!

Instructions. Exercising at different times can prevent you from falling into a stale routine. Exercise should be fun, not a fixed schedule such as your job. For instance, you can exercise in the mornings on Mondays and Wednesdays and save Tuesdays and Thursdays for evening workouts. Finish with a weekend of exercising in the afternoon then switch the times on the following week! “Prior location” indicates your last location before you leave for the gym i.e., your home, your work, after dropping off kids to school, etc.

Monday

<table>
<thead>
<tr>
<th>Time</th>
<th>Morning</th>
<th>Mid-morning</th>
<th>Afternoon</th>
<th>Evening</th>
<th>Night</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tuesday

<table>
<thead>
<tr>
<th>Time</th>
<th>Morning</th>
<th>Mid-morning</th>
<th>Afternoon</th>
<th>Evening</th>
<th>Night</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Wednesday

<table>
<thead>
<tr>
<th>Time</th>
<th>Morning</th>
<th>Mid-morning</th>
<th>Afternoon</th>
<th>Evening</th>
<th>Night</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day</td>
<td>Morning</td>
<td>Mid-morning</td>
<td>Afternoon</td>
<td>Evening</td>
<td>Night</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>-------------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>Thursday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sunday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 7: Study III Questionnaire

Instructions

In this survey, we are going to ask you a series of questions about your beliefs and attitudes toward physical activity. There are no right or wrong answers and all we ask is that you provide responses that are as honest and accurate as possible. Although some of the questions may seem redundant, it is for the reliability of the study, and we ask that you answer them to the best of your ability. The questionnaire should take about 15 minutes for you to complete. All responses are completely confidential and will never be used in any way that could link them to you. If you have any questions, please ask one of the research assistants.

If you have any questions about completing the questionnaire, please contact Navin Kaushal (kaushal@uvic.ca)

Participant Number

For the purpose of matching your baseline and follow-up survey, we would ask you to please create your participant number based on the instructions below: Your participant number will consist of your height, day of your birthday and age. Example, Sarah is 5 feet 4 inches tall, her birthday is June 7th and she is 31 years old. Sarah’s participant number will be 54731.

Please enter your participant number _________________

Exercise Behavior

We would like you to recall your average weekly leisure time physical activity over the past month. Specifically, on average, how many times per week did you walk over the past month and what was the duration of these activities? When answering these questions please:

Only count physical activity that was done during free time (i.e., not occupation, school or housework).

Note that the main difference between the three categories is the intensity of the activity.

Write the average frequency on the first line and the average duration on the second line.
<table>
<thead>
<tr>
<th>Activity</th>
<th>Times Per Week</th>
<th>Average Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. STRENUOUS physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(HEART BEATS RAPIDLY, SWEATING)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e.g., running, jogging, hockey, soccer, squash, cross country skiing, vigorous swimming, vigorous aerobic dance classes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. MODERATE physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(NOT EXHAUSTING, LIGHT PERSPIRATION)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e.g., similar to above but at moderate intensity)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. MILD physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Minimal effort, no perspiration)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following questions ask you to rate how you feel about regular physical activity. We define regular physical activity as accumulating at least 150 minutes of per week during your free time that is in the moderate intensity or higher range. Pay careful attention to the words at each end of the scales and mark that best represents how you feel about regular physical activity.

**Decisional Intention**

1. My goal is to engage in moderate or vigorous exercise for at least 150 min per week over the next month
   ________agree or ________disagree

2. I intend to engage in regular leisure-time physical activity for ________ minutes per week on average over the next month.
Consistency- Adapted from Danner et al, 2008 Stability of context: time (item 1) location (item 2); Tappe and Glanz, 2013: consistency of exercise routine (item 3). Items 1 and 2 also from Ji and Wood 2007

1. How consistently do you exercise at the same time each day? (e.g., exercising every morning at 7 am, or exercising daily after supper).

1                      2                      3                      4                      5
Strongly           Disagree        Neutral        Agree        Strongly
Disagree

Environmental Cues: Danner et al., 2008 (stability of circumstances)

1. I use cues at home to remind me to exercise (ie. Placing a water bottle on my desk or gym clothes on the bed).

1                      2                      3                      4                      5
Strongly           Disagree        Neutral        Agree        Strongly
Disagree

Habit

The following questions ask how you feel about the preparing to exercise. The exercise preparatory phase includes all of the activities you would regularly perform before each exercise session (ie. packing gym bag, equipping mp3 player, having a pre-workout meal, etc).

1. When I prepare to exercise, I do it automatically

1                      2                      3                      4                      5
Strongly           Disagree        Neutral        Agree        Strongly
Disagree

2. I prepare to exercise without having to consciously remember

1                      2                      3                      4                      5
3. I prepare to exercise without thinking

1  2  3  4  5
Strongly  Disagree  Neutral  Agree  Strongly
Disagree  Agree

4. I start preparing to exercise before I realize I am doing it

1  2  3  4  5
Strongly  Disagree  Neutral  Agree  Strongly
Disagree  Agree

**General Health Questions**

The next questions are about your health & physical activity behaviors. Please check one:

1. Do you currently smoke cigarettes?
   - Yes □  How many cigarettes do you usually smoke a day? ______
   - No □  Have you ever smoked cigarettes?  Yes □  No □

We would like to know a little more about your medical and health background...

2. Has a close blood relative (e.g., a parent, brother or sister) ever had heart disease (e.g., heart attack, stroke, and/or angina) before the age of 60?
   - Yes □  No □

3. Has a doctor or nurse ever told you that you have had the following: (please check all that apply)
   a. Angina  Yes □  No □
   b. Heart Attack  Yes □  No □
   c. Stroke  Yes □  No □
   d. High blood cholesterol  Yes □  No □
   e. Cancer  Yes □  No □
   f. High blood pressure  Yes □  No □
   g. Diabetes  Yes □  No □
   If yes, which type?  Type 1 □  Type 2 □
   Gestional □
4. In general, compared to other persons your age, how would you rate your health?

    Poor □  Fair □  Good □  Very Good □  Excellent □

**Demographics**

The following questions are needed to help us understand the characteristics of the people participating in the study. For this reason it is very important information. All information is held in strict confidence and its presentation to the public will be in the form of group data only.

1. Age: ______
2. Gender: Male □  Female □
3. Ethnicity/Race: ____________
4. What is your primary method of transportation?
   - walking  bicycle  bus/city transit  car or other motorized vehicle

5. What is the highest level of education that you completed? Please check only one.
   - □ 8th grade or less
   - □ Vocational school or some college
   - □ Some high school
   - □ College degree
   - □ High school diploma
   - □ Professional or graduate degree

7. What is your current marital status? Please check only one.
   - □ Never married
   - □ Married/common law
   - □ Separated/divorced/widowed

8. What is your job situation? Please check the one that describes you best.
   - □ Homemaker
   - □ Retired
   - □ Paid full-time employment
   - □ Paid part-time employment
   - □ Temporarily unemployed
9. What is your annual household income (after taxes)? Please check only one.

- $35,000 or less
- $50,001 to $75,000
- $100,001 to $150,000
- $35,001 to $50,000
- $75,001 to $100,000
- $150,001 - $200,000
- More than $200,000