

THE MEASUREMENT OF PHYSICAL ACTIVITY SELF-EFFICACY IN INTERVENTIONS  
THAT PROMOTE PHYSICAL ACTIVITY IN ADULTS

By

André Godfrey Bateman

A DISSERTATION

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

Kinesiology—Doctor of Philosophy

2022

## PUBLIC ABSTRACT

### THE MEASUREMENT OF PHYSICAL ACTIVITY SELF-EFFICACY IN INTERVENTIONS THAT PROMOTE PHYSICAL ACTIVITY IN ADULTS

By

André Godfrey Bateman

Physical activity has been associated with positive health outcomes. Insufficient physical activity has therefore become a global pandemic in the general adult population. There is a rich literature on the positive association between various forms of self-efficacy beliefs and physical activity outcomes. Consequently, self-efficacy has been employed as a motivational construct in physical activity promoting interventions. These physical activity promoting interventions typically specify self-efficacy as a mediator of physical activity participation.

It is important that the measurement of self-efficacy associated with physical activity in physical activity promoting interventions is valid (i.e., congruent with theory). For example, self-efficacy theory specifies different forms of self-efficacy (e.g., self-regulatory efficacy and task-related self-efficacy), each of which are conceptually distinct. To ensure validity of scores, conceptual distinctions should be reflected in the measurement of each of these constructs. Issues currently exist with the measurement of self-efficacy associated with physical activity in physical activity promoting interventions. These issues include making the distinction between the measurement of different forms of self-efficacy, along with other concerns (e.g., uncertainty about dimensionality of measures). These issues, if not uncovered and addressed, will have implications for the reliability and validity of the scores produced by self-efficacy measures. The overarching focus of the current project is therefore to improve the reliability and validity of self-efficacy measurement in physical activity promoting interventions. The project is therefore expected to make a necessary contribution to self-efficacy measurement literature.

This project consists of two studies aimed at targeting the issues which exist in the measurement of self-efficacy in physical activity promoting interventions designed for adults. Study 1, a systematic review, is focused on uncovering the issues which exist in the measurement of physical activity self-efficacy in physical activity promoting interventions. Study 2, in response to the findings of the first study focuses on exploring the measurement properties (e.g., dimensionality) of the self-efficacy to regulate physical activity scale as utilized in a recent physical activity promoting intervention targeting adults with obesity (Fun For Wellness; FFW). Both studies were guided by established recommendations for the measurement of psychological constructs, the measurement of self-efficacy in general, and the measurement of self-efficacy in the domain of human physical performance (i.e., physical activity). The findings of both studies are expected to clarify the issues that exist and contribute to an improvement in the measurement of self-efficacy for physical activity in physical activity promoting interventions.

## ABSTRACT

### THE MEASUREMENT OF PHYSICAL ACTIVITY SELF-EFFICACY IN INTERVENTIONS THAT PROMOTE PHYSICAL ACTIVITY IN ADULTS

By

André Godfrey Bateman

This dissertation comprises two studies focused on the measurement of self-efficacy associated with physical activity-promoting interventions in adults. Recent research indicates that most adults do not achieve sufficient daily physical activity for health. The research also shows that adults with obesity are even less likely to engage in sufficient physical activity for health. Physical inactivity is associated with negative health outcomes such as cardiovascular disease and is therefore a major public health concern. There is however evidence that certain motivational constructs, such as self-efficacy are associated with increased physical activity in adults. As a result, behavioral interventions utilizing these constructs as modifiable mediators of physical activity behavior have been employed to increase physical activity in different populations.

Study 1 is a systematic review focused on examining the theoretical and measurement quality of physical activity self-efficacy scales in physical activity-promoting interventions for adults. The search strategy was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. One hundred sixteen studies were reviewed, from which the physical activity self-efficacy scales were identified and extracted. Of the scales identified, 14 were multi-item and five were single item scales. The systematic review uncovered that the identified scales had varying conceptual and measurement related properties despite having good administrative quality in general. The major issues identified with self-efficacy measurement were: (a) a lack of concordance between self-efficacy and physical activity measurement, (b) a

lack of specified physical activity levels to which the self-efficacy measurements refer, (c) self-efficacy scales described with theoretically imprecise construct labels, (d) a lack of emphasis on essential conceptual properties of self-efficacy scales, (e) a lack of specification of the dimensionality of self-efficacy scales and (f) the use of single-item measures of self-efficacy. Essential conceptual and measurement related recommendations were made in response to these issues to improve the measurement of physical activity self-efficacy in physical activity-promoting interventions.

Study 2 employed a latent variable approach to explore the dimensionality, temporal invariance, and external validity of responses to the self-efficacy to regulate physical activity scale (SERPA). The SERPA is a modified version of the barriers self-efficacy scale. This study analyzed data from the Well-Being and Physical Activity Study (WBPA; ClinicalTrials.gov, identifier: NCT03194854). The WBPA consisted of 461 participants at baseline which decreased to 427 participants at 30 days post baseline. The WBPA deployed the Fun For Wellness (FFW) intervention. One objective of the FFW intervention was to promote physical activity in adults with obesity. A two-dimensional factor structure explained responses to the SERPA at baseline. Factor 1 was conceptualized as self-efficacy to regulate barriers to physical activity participation based on social considerations. Factor 2 was conceptualized as self-efficacy to regulate internally perceived barriers to physical activity participation. There was strong evidence for the effectiveness of the FFW intervention to exert a direct effect on the proposed two-dimensional structure of latent self-efficacy to regulate physical activity in adults with obesity at 30 days post-baseline.

Copyright by  
ANDRÉ GODFREY BATEMAN  
2022

This dissertation is dedicated to my mother, the strongest woman I know, the one who I have always tried to emulate.

## ACKNOWLEDGEMENTS

First, I would like to acknowledge Dr. Nicholas Myers, my advisor for bearing with me through this process. I know working with me could not have been easy. Dr. Myers, your mentorship has been excellent. What I appreciate the most is that you challenged me develop in the areas that were deficient and made sure I improved. Your thoughtful, strategic approach to mentorship has also not gone unnoticed and though there are many ways I can still improve I am confident that I am now well-positioned to become an outstanding scholar.

Second, I would like to thank my close family and friends for their support throughout this journey. Thank you to my wife, father, and brother for their unwavering support. Thanks, Rochette, for doing all the things a best friend does, including tolerating those long phone calls where I vented all my frustrations. Thank you, Sara, for always being available to review my work for language accuracy. Thanks, auntie Christine for your timely phone calls to check in, and always making your home in New York available for me when I needed a break from Michigan. Thanks also to Damion Myers, Vanburn Phillips, Andrew McIntosh, Marvin Powell, Dorian Hayden, Lucas Capalbo, Karim Blake, Chelsi Ricketts, Morgan Anderson, Emily Werner, Allie Tracey, and James, and Linda Pivarnik for their support in various ways.

Third, a special thanks goes out to the supportive administrative staff of the Kinesiology department. Christina, Michelle, Marlene, Mary-Anne thank you for being so helpful and unselfish, it was greatly appreciated. Thanks also to Dr. Al Smith who became a second advisor, mentor, source of support through difficult times, and friend. Finally, I thank God for blessing me with the intellect and perseverance to be successful during this journey and for all the people who supported me along the way.



## PREFACE

Study 1 was financially supported, in part, with a Summer Research Fellowship awarded in 2021 by the College of Education at Michigan State University. Study 1 from this dissertation is published in a peer reviewed journal, *Measurement in Physical Education and Exercise Science*. Permission has been received from the journal to use the published manuscript as part of this dissertation. Study 2 was financially supported, in part, with a Dissertation Completion Fellowship awarded in 2022 by the College of Education at Michigan State University.

### **Study 1 Citation**

Bateman, A., Myers, N. D., Chen, S., & Lee, S. (2021). Measurement of Physical Activity Self-Efficacy in Physical Activity-Promoting Interventions in Adults: A Systematic Review. *Measurement in Physical Education and Exercise Science*. Advance online publication. <https://doi.org/10.1080/1091367X.2021.1962324>.

## TABLE OF CONTENTS

|  |     |
|--|-----|
| LIST OF TABLES .....   | x   |
| LIST OF FIGURES .....  | xi  |
| KEY TO ABBREVIATIONS.....  | xii |
| CHAPTER I: GENERAL INTRODUCTION .....  | 1   |
| REFERENCES .....   | 6   |
| CHAPTER II: STUDY 1 .....  | 11  |
| ABSTRACT.....  | 11  |
| INTRODUCTION .....   | 11  |
| METHODS .....  | 17  |
| Search Strategy .....  | 17  |
| Eligibility Criteria .....   | 17  |
| Data Extraction .....  | 18  |
| Conceptual and Measurement-related Examination of Physical Activity Self-<br>efficacy..... | 19  |
| Examination of Administrative Criteria.....  | 20  |
| RESULTS .....  | 21  |
| Descriptive Results .....  | 22  |
| Conceptual and Measurement-Related Aspects of Physical Activity Self-<br>efficacy.....     | 24  |
| Examination of Administrative Criteria.....  | 29  |
| DISCUSSION .....   | 31  |
| Limitations .....  | 36  |
| Conclusion and Recommendations.....  | 37  |
| REFERENCES .....   | 39  |
| CHAPTER III: STUDY 2.....  | 46  |
| ABSTRACT.....  | 46  |
| INTRODUCTION .....   | 46  |
| The Well-being and Physical Activity (WBPA) Study .....                                    | 47  |
| The Intervention.....  | 47  |
| Self-efficacy Theory as the Basis for the Intervention and Measurement.....                | 48  |
| Measurement of Self-efficacy to Regulate Physical Activity.....                            | 50  |
| Results Under a Traditional Observed Score Approach .....                                  | 50  |
| Objective and Exploratory Research Questions .....   | 52  |
| GENERAL METHODS.....   | 53  |
| Study Design.....  | 53  |
| Participants.....  | 54  |
| Procedures.....  | 54  |

|  |    |
|--|----|
| Measures .....   | 55 |
| Development of the SERPA .....   | 55 |
| Data Collection, Demographics, and Descriptive Statistics .....  | 57 |
| Data Analysis .....  | 58 |
| STUDY 2A METHODS .....   | 60 |
| STUDY 2A RESULTS .....   | 60 |
| Interpreting the Two-Factor Solution .....   | 62 |
| STUDY 2A CONCLUSION.....   | 63 |
| STUDY 2B METHODS .....   | 64 |
| STUDY 2B RESULTS .....   | 64 |
| STUDY 2B CONCLUSION.....   | 66 |
| STUDY 2C METHODS .....   | 67 |
| STUDY 2C RESULTS .....   | 68 |
| STUDY 2C CONCLUSION.....   | 70 |
| BRIEF GENERAL DISCUSSION.....  | 70 |
| APPENDICES .....   | 75 |
| APPENDIX A: Search Strategy: Embase.....   | 76 |
| APPENDIX B: Reviewer Rating Scale .....  | 77 |
| APPENDIX C: Self-Efficacy to Regulate Physical Activity (SERPA) Scale.....   | 81 |
| APPENDIX D: Unstandardized Direct Effects ( $\gamma$ ) in the Path Model for Factor 1 and<br>Factor 2 of the demographic covariates at Time 2 (N = 424) Regressed on FFW ..... | 83 |
| REFERENCES .....   | 84 |
| <br>   |    |
| CHAPTER IV: SUMMARY/CONCLUSIONS .....  | 91 |
| STRENGTHS AND LIMITATIONS .....  | 93 |
| CONCLUSION.....  | 95 |
| FUTURE RESEARCH DIRECTIONS .....   | 95 |
| REFERENCES .....   | 97 |

## LIST OF TABLES

|  |    |
|--|----|
| <i>Table 1. Multi-item Scales Identified and Extracted with Construct Descriptions and Other Characteristics</i> .....   | 23 |
| <i>Table 2. Reviewer Consensus Scores for the Conceptual-related Assessment the of Extracted Scales as Described in their Cited Studies</i> .....  | 25 |
| <i>Table 3. Measurement-Related and Administrative Properties of Extracted Scales</i> .....  | 26 |
| <i>Table 4. Administrative Properties of Extracted Scales</i> .....  | 30 |
| <i>Table 5. Distribution of Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) items at Time 1</i> .....   | 59 |
| <i>Table 6. Distribution of Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) items at Time 2</i> .....   | 59 |
| <i>Table 7. Number of Factors Warranted to Explain Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 1 (N = 461)</i> .....  | 61 |
| <i>Table 8. The Accepted Geomin-Rotated (<math>\epsilon = .1</math>) Pattern Coefficients (<math>\Lambda^*</math>), Inter-Factor Correlation (<math>\psi</math>), and Coefficient H, for the Self-efficacy to Regulate Physical Activity Scale (SERPA) Factors at Time 1 (N = 461)</i> .....                           | 63 |
| <i>Table 9. Longitudinal Measurement Invariance for Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 1 (N = 461) and Time 2 (N = 424)</i> .....  | 66 |
| <i>Table 10. Model-Data Fit, Percentage of Latent Variable Variance Accounted for (<math>R^2</math>), and Unstandardized Direct Effects (<math>\gamma</math>) in the Path Model for Factor 1 and Factor 2 of the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 2 (N = 424) Regressed on FFW</i> .. | 69 |
| <i>Table 11. Unstandardized Direct Effects (<math>\gamma</math>) in the Path Model for Factor 1 and Factor 2 of the demographic covariates at Time 2 (N = 424) Regressed on FFW</i> .....  | 83 |

## LIST OF FIGURES

*Figure 1. Schematic review of study and scale selection based on the PRISMA guidance.....22*

*Figure 2. Focal Parameters (i.e.,  $\gamma_1$  and  $\gamma_2$ ) from the Path Model for the Self-Efficacy to Regulate Physical Activity Scale (SERPA) at T1 (Baseline) and T2 (30 Days Post-Baseline) .....68*

## KEY TO ABBREVIATIONS

|           |  |
|-----------|--|
| AERA      | American Educational Research Association                              |
| APA       | American Psychological Association                                     |
| BARSE     | Barriers specific self-efficacy scale                                  |
| BET I CAN | Behaviors Emotions Thoughts Interactions Contexts Awareness Next steps |
| CFI       | Comparative Fit Index  |
| C-SES     | Craig self-efficacy scale  |
| ESE       | Exercise self-efficacy scale   |
| ESEM      | Exploratory Structural Equation Modeling                               |
| EXSE      | Exercise self-efficacy scale   |
| FFW       | Fun For Wellness   |
| HBS       | Health behavior scale  |
| ML        | Maximum-likelihood   |
| NCME      | National Council on Measurement in Education                           |
| PAAI      | Physical activity assessment inventory                                 |
| PASE      | Physical activity self-efficacy  |
| PRISMA    | Preferred Reporting Items for Systematic Reviews and Meta-Analyses     |
| P-SES     | Plotnikoff self-efficacy scale   |
| RMSEA     | Root Mean Square Error of Approximation                                |
| SEB       | Self-efficacy for exercise behaviors scale                             |
| SEE       | Self-efficacy for exercise scale                                       |
| SEI       | Self-efficacy inventory  |

|        |   |
|--------|---|
| SEPA   | Self-efficacy for physical activity scale                   |
| SE-PBC | Self-efficacy related to perceived behavioral control scale |
| SEQ    | Self-efficacy for exercise questionnaire                    |
| SERPA  | Self-efficacy to regulate physical activity                 |
| SRMR   | Standardized Root Mean Squared Residual                     |
| TLI    | Tucker-Lewis Index  |
| UC     | Usual care  |
| USDHHS | United States Department of Health and Human Services       |
| WBPA   | Well-Being and Physical Activity                            |
| WHO    | World Health Organization                                   |
| W-SES  | Wooldridge self-efficacy scale                              |

## CHAPTER I: GENERAL INTRODUCTION

Self-efficacy is a well-established construct in sport and exercise psychology literature. Consequently, many reviews have considered the general history, theoretical structure, measurement, determinants, and/or consequences of self-efficacy perceptions (e.g., Bandura, 1997, 2006; Beauchamp et al., 2012; Feltz et al., 2008; Jackson et al., 2020). Self-efficacy beliefs, broadly defined, refer to situational or domain-specific self-confidence. Self-efficacy judgements have been positively associated with performance in sport and exercise contexts (e.g., Bauman et al., 2012; Moritz et al., 2000). Consequently, self-efficacy theory has been increasingly applied in interventions designed to promote physical activity participation or performance in different populations (e.g., Anderson et al., 2007; Mailey et al., 2010; Myers et al., 2020). Typically, these interventions conceptualize a person's self-efficacy for (or associated with) physical activity as a malleable belief system, which under the right circumstances (e.g., when exposed to sources of self-efficacy information) can be increased. This is expected to result in increased participation in measured physical activity. It is therefore important that steps be taken to ensure the measurement of self-efficacy associated with physical activity in such interventions is valid (i.e., congruent with theory).

Physical activity participation has been associated with positive health and wellness outcomes such as decreased obesity and cardiovascular disease (United States Department of Health and Human Services, 2018). Despite this, many adults globally continue to not meet recommended amounts of physical activity for health (Guthold et al., 2016). Obesity has also become increasingly prevalent worldwide (Haththotuwa et al., 2020). These coinciding problems of insufficient physical activity and obesity are particularly prevalent in the United States (Ogden et al., 2020). Obesity and insufficient physical activity may also be associated with each other



(Donnelly et al., 2009). Higher levels of insufficient physical activity have led to an increase in the recommendations for programs and interventions aimed at promoting physical activity participation in different sub-populations (DiPietro et al., 2020). Psychosocial approaches have subsequently been incorporated into physical activity interventions as modifiable mediators of physical activity behavior (Curry et al., 2018). Self-efficacy is one such motivational construct, which has been employed in psychosocial interventions as a mediator of physical activity in physical activity interventions (Bauman et al., 2012). For example, Myers et al. (2020) applied two forms of self-efficacy, physical activity self-efficacy (PASE) and self-efficacy to regulate physical activity (SERPA) in an online physical activity promoting intervention designed for adults with obesity.

Bandura (1997) formally defines self-efficacy as one's beliefs in their capabilities to organize and execute the behavior required to produce desired outcomes in a specific domain. One's perceptions about their ability to perform a behavior appears to play a fundamental role in promoting behavior change (Jackson et al., 2020). Consequently, many models of health behavior change (e.g., the health belief model; Rosenstock et al., 1988) now incorporate aspects of self-efficacy theory (Bandura, 1997). Self-efficacy is a component of the broader social cognitive theory (Bandura, 2001), which theorizes that human beings are capable of determining their own thoughts, feelings, and actions as they interact with the environment. Bandura (1997) outlined that self-efficacy beliefs may vary according to their level (e.g., straightforward, or burdensome), strength (from low to high), and generality (degree of transferability from one context to another). Four antecedents or sources of self-efficacy beliefs have also been identified: mastery experiences, verbal persuasion, vicarious influences (or modeling), and awareness of physiological/psychological states. Consequently, a core postulate of self-efficacy theory is that a

strong sense of self-efficacy and/or an environment where sources of self-efficacy information are abundant will motivate behavioral or performance outcomes. For example, Anderson-Bill et al. (2011) found evidence for the positive association of self-efficacy beliefs with physical activity outcomes. Bandura (1997) also referred to two different forms of self-efficacy beliefs, task-related self-efficacy, and self-regulatory efficacy. Task-related self-efficacy beliefs refer to one's belief in their ability to accomplish a specific task, while self-regulatory efficacy refers to one's belief in their ability to manage their behavior over time (e.g., despite barriers). Evidence has been found to support the relevance of both forms of self-efficacy beliefs in physical activity performance contexts (e.g., McAuley 1992, 1993; Myers et al., 2020).

Bandura (1997, 2006) provides detailed recommendations for the measurement of the self-efficacy construct. More generally, the measurement of psychological (or psychosocial) constructs is based on the premise that these underlying constructs are inferred from a set of similar observable indicators (e.g., responses to self-report items) conceptually believed to be indicative of the construct (Raykov & Marcoulides, 2011). The items comprising self-report self-efficacy measures should therefore be substantively representative of the construct as outlined by self-efficacy theory. A review of the self-efficacy literature (e.g., Bandura, 1997, 2006; Feltz & Chase, 1998; Feltz et al., 2008; Myers & Feltz, 2007; Myers et al., 2005; Myers et al., 2008) reveals the major principles of the valid measurement of the self-efficacy construct.

Bauman et al. (2012) emphasizes that an understanding of the correlates and determinants of physical activity is essential in designing interventions aimed at promoting physical activity. In the review, Bauman et al. (2012) identified self-efficacy as a correlate of physical activity in adults. Targeting self-efficacy in physical activity interventions could therefore be important in improving the effectiveness of physical activity interventions. Despite the potential for self-

efficacy as a potential correlate or determinant of physical activity in adults, there appears to be some issues in the conceptual and measurement-related application of this construct in physical activity-promoting interventions.

Though self-efficacy has been incorporated in many physical activity promoting interventions, the construct has been measured using different scales in different studies. For example, Robertson et al. (2020) used a single-item scale, while Anderson-Bill et al. (2011) used a two-dimensional, 23-item scale to measure self-efficacy associated with physical activity. Additionally, scales used to measure self-efficacy associated with physical activity seem to have varying dimensionality. For example, the self-efficacy for physical activity scale (SEPA; Marcus et al., 1992) appears to be treated as unidimensional, while another commonly used scale, the self-efficacy for exercise behaviors scale (SEB; Sallis et al., 1988) appears to be two-dimensional. Another concern for the measurement of self-efficacy in physical activity interventions in general is whether each scale used subscribes to the conceptual (e.g., focused on current capability judgment) and measurement-related (e.g., optimally categorized rating scales) recommendations according to the literature.

The current project focuses on addressing the issues that exist in the measurement of self-efficacy beliefs in physical activity promoting interventions. The aforementioned issues may have implications for the validity of scores produced by self-efficacy measurement. Validity, the overall evaluation of the extent to which evidence and theory support the interpretation of scores obtained from a measure (Messick, 1995) is a fundamental principle in the measurement of psychological constructs. Importantly, the validation of constructs is an ongoing process of collecting evidence that the underlying construct accurately represents the concept being measured (Raykov & Marcoulides, 2011). Furthermore, according to the American Educational

Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME; 2014), validity evidence should be provided for psychological measures, and these measures should be revised if there is evidence for improved validity. The overarching purpose of this study is to improve the measurement of self-efficacy associated with physical activity.

The current study has two major aims. Each major aim will be addressed in separate studies, that is, Study 1 and Study 2. The aim of Study 1 is to systematically review the measurement of self-efficacy in physical activity promoting interventions for adults. The aim of Study 2 is to examine the measurement of self-efficacy to regulate physical activity as used in the Fun For Wellness (FFW) intervention (e.g., Myers et al., 2020). Achieving these aims is expected to result in the advancement of the measurement of self-efficacy associated with physical activity by producing two major outcomes. First, to uncover the current state of the art in the measurement of self-efficacy in physical activity interventions for adults. Second, to provide validity evidence for a scale used to measure self-efficacy to regulate physical activity in a recently published physical activity promoting intervention.

## REFERENCES

## REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderson-Bill, E. S., Winett, R. A., Wojcik, J. R., & Winett, S. G. (2011). Web-based guide to health: relationship of theoretical variables to change in physical activity, nutrition and weight at 16-months. *Journal of medical Internet research*, *13*(1), e27.  
<https://doi.org/10.2196/jmir.1614>
- Anderson, E. S., Winett, R. A., & Wojcik, J. R. (2007). Self-regulation, self-efficacy, outcome expectations, and social support: Social cognitive theory and nutrition behavior. *Annals of Behavioral Medicine*, *34*(3), 304-312.  
<https://psycnet.apa.org/doi/10.1007/BF02874555>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1-26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.
- Bauman, A. E., Reis R. S., Sallis, J. F., Wells, J. C., Loos, R. J. F., & Martin, B. W. (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet*, *380*, 258-271. [https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)
- Beauchamp, M. R., Jackson, B., & Morton, K. (2012). Efficacy beliefs and human performance: From independent action to interpersonal functioning. In S. Murphy (Ed.), *The Oxford handbook of sport and performance psychology* (pp. 273–293). New York, NY: Oxford University Press.
- Curry, S. J., Krist, A. H., Owens, D. K., Barry, M. J., Caughey, A. B., Davidson, K. W., ... & Kubik, M. (2018). Behavioral weight loss interventions to prevent obesity-related morbidity and mortality in adults: US Preventive Services Task Force recommendation statement. *Jama*, *320*(11), 1163-1171.  
<http://jamanetwork.com/journals/jama/fullarticle/10.1001/jama.2018.13022>
- DiPietro, L., Al-Ansari, S. S., Biddle, S. J., Borodulin, K., Bull, F. C., Buman, M. P., ... & Willumsen, J. F. (2020). Advancing the global physical activity agenda: recommendations for future research by the 2020 WHO physical activity and sedentary

- behavior guidelines development group. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 1-11. <https://doi.org/10.1186/s12966-020-01042-2>
- Donnelly, J. E., Blair, S. N., Jakicic, J. M., Manore, M. M., Rankin, J. W., & Smith, B. K. (2009). Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Medicine & Science in Sports & Exercise*, 41(2), 459-471. <https://doi.org/10.1249/mss.0b013e3181949333>
- Feltz, D. L., & Chase, M. A. (1998). The measurement of self-efficacy and confidence in sport. In J. L. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 63- 78). Morgantown, WV: Fitness Information Technology.
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport: Research and strategies for working with athletes, teams, and coaches*. Champaign, IL: Human Kinetics.
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport: Research and strategies for working with athletes, teams, and coaches*. Champaign, IL: Human Kinetics.
- Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1· 9 million participants. *The lancet global health*, 6(10), e1077-e1086. [https://doi.org/10.1016/S2214-109X\(18\)30357-7](https://doi.org/10.1016/S2214-109X(18)30357-7)
- Haththotuwa, R. N., Wijeyaratne, C. N., & Senarath, U. (2020). Worldwide epidemic of obesity. In *Obesity and obstetrics* (pp. 3-8). Elsevier. <https://doi.org/10.1016/B978-0-12-817921-5.00001-1>
- Mailey, E. L., Wójcicki, T. R., Motl, R. W., Hu, L., Strauser, D. R., Collins, K. D., & McAuley, E. (2010). Internet-delivered physical activity intervention for college students with mental health disorders: a randomized pilot trial. *Psychology, health & medicine*, 15(6), 646-659. <https://doi.org/10.1080/13548506.2010.498894>
- Marcus, B. H., Selby, V. C., Niaura, R. S., & Rossi, J. S. (1992). Self-efficacy and the stages of exercise behavior change. *Research Quarterly for Exercise and Sport*, 63(1), 60-66. <https://doi.org/10.1080/02701367.1992.10607557>
- McAuley, E. (1992). The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. *Journal of Behavioral Medicine*, 15(1), 65-88. <https://psycnet.apa.org/doi/10.1007/BF00848378>
- McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of behavioral medicine*, 16(1), 103-113. <https://doi.org/10.1007/BF00844757>

- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational measurement: Issues and practice*, 14(4), 5-8. <https://doi.org/10.1111/j.1745-3992.1995.tb00881.x>
- Moritz, S. E., Feltz, D. L., Fahrback, K. R., & Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport*, 71, 280–294. <https://doi.org/10.1080/02701367.2000.10608908>
- Myers, N. D., & Feltz, D. L. (2007). From self-efficacy to collective efficacy in sport: Transitional methodological issues. In G. Tenenbaum & R.C. Eklund (Eds.), *The handbook of sport psychology* (pp. 799–819). Wiley.
- Myers, N. D., Feltz, D. L., & Wolfe, E. W. (2008). A confirmatory study of rating scale category effectiveness for the coaching efficacy scale. *Research Quarterly for Exercise and Sport*, 79(3), 300-311. <https://doi.org/10.1080/02701367.2008.10599493>
- Myers, N. D., McMahon, A., Prilleltensky, I., Lee, S., Dietz, S., Prilleltensky, O., ... & Brincks, A. M. (2020). Effectiveness of the fun for wellness web-based behavioral intervention to promote physical activity in adults with obesity (or overweight): Randomized controlled trial. *JMIR formative research*, 4(2), e15919. <https://doi.org/10.2196/15919>.
- Myers, N. D., Wolfe, E. W., & Feltz, D. L. (2005). An evaluation of the psychometric properties of the coaching efficacy scale for coaches from the united states of america. *Measurement in Physical Education and Exercise Science*, 9(3), 135-160. [https://doi.org/10.1207/s15327841mpee0903\\_1](https://doi.org/10.1207/s15327841mpee0903_1)
- Ogden, C. L., Fryar, C. D., Martin, C. B., Freedman, D. S., Carroll, M. D., Gu, Q., & Hales, C. M. (2020). Trends in obesity prevalence by race and hispanic origin—1999-2000 to 2017-2018. *Jama*, 324(12), 1208-1210. <https://doi.org/10.1001/jama.2020.14590>
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Taylor & Francis.
- Robertson, M. C., Green, C. E., Liao, Y., Durand, C. P., & Basen-Engquist, K. M. (2020). Self-efficacy and physical activity in overweight and obese adults participating in a worksite weight loss intervention: multistate modeling of wearable device data. *Cancer Epidemiology and Prevention Biomarkers*, 29(4), 769-776. <https://doi.org/10.1158/1055-9965.EPI-19-0907>
- Rosenstock, I. M., Strecher, V. J., & Becker, M. H. (1988). Social learning theory and the health belief model. *Health education quarterly*, 15(2), 175-183. <https://doi.org/10.1177/109019818801500203>
- Sallis, J. F., Pinski, R. B., Grossman, R. M., Patterson, T. L., & Nader, P. R. (1988). The development of self-efficacy scales for health-related diet and exercise behaviors. *Health Education Research*, 3(3), 283-292. <https://doi.org/10.1093/her/3.3.283>



United States Department of Health and Human Services: 2018 Physical activity guidelines advisory committee. (2018). *2018 Physical activity guidelines advisory committee scientific report*. <https://health.gov/paguidelines/second-edition/report/>

## CHAPTER II: STUDY 1

### ABSTRACT

Self-efficacy is a psychosocial determinant of physical activity in adults. Different scales have been used to measure physical activity self-efficacy. This review examines the theoretical and measurement quality of scales measuring physical activity self-efficacy in physical activity-promoting interventions. The search strategy was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Studies were included if they measured physical activity self-efficacy in adults aged 18 to 65. One hundred sixteen studies were reviewed. Fourteen multi-item and five single-item scales were identified. The properties of the scales varied. The following issues were identified: (a) a lack of concordance between self-efficacy and physical activity measurement, (b) not specifying physical activity levels, (c) theoretically imprecise construct labels, (d) not emphasizing essential conceptual properties, (e) not reporting dimensionality and (f) the use of single-item measures. The scales showed good administrative properties. Recommendations are made to improve the measurement of physical activity self-efficacy.

### INTRODUCTION

There is a rich literature on the potential of increasing self-efficacy associated with physical activity as a mechanism for promoting physical activity in adults (Bauman et al., 2012). Physical activity has been associated with positive health outcomes such as the reduction of obesity and cardiovascular disease in adults (i.e., people within the 18-64 age-range; United States Department of Health and Human Services [USDHHS], 2018). Insufficient physical activity however, has become a global pandemic in the general adult population (Kohl et al., 2012; Sallis et al., 2016). Addressing this pandemic through physical activity-promoting

interventions is likely to have a positive impact on public health leading to disease prevention and health promotion (USDHHS, 2018).

Self-efficacy theory is a component of the broader social cognitive theory. Self-efficacy judgements are domain-specific beliefs held by an individual regarding their capability to perform a specific behavior or task given certain situational demands (Bandura, 1997). Scales measuring self-efficacy within a specified domain should therefore measure a respondent's confidence in their capability to perform specific behaviors within the specified domain (Beauchamp, 2016). Consequently, self-efficacy associated with physical activity generally describes the degree to which an individual believes that they have the capability to engage in physical activity behavior. More specifically, task-related self-efficacy beliefs (e.g., physical activity self-efficacy) within a specified domain refer to an individual's beliefs in their ability to accomplish levels of a task (e.g., engage weekly in at least 150 minutes of moderate intensity physical activity). Self-regulatory efficacy (i.e., self-efficacy to regulate a behavior) refers to an individual's beliefs in their ability to overcome challenging situations or possible barriers to accomplishing a task that he or she already knows how to do (e.g., engage in physical activity even if under personal stress; Bandura, 1997; Jackson et al. 2020). Self-efficacy judgements are subsequently inextricably linked to performance within the specified domain. As a result, the validity of scores produced by self-efficacy measures is dependent on the extent to which the content of the scale represents the construct and how well the construct predicts performance within the specified domain (Bandura, 1997).

Self-efficacy has been found to be a predictor of physical activity in different adult sub-populations (e.g., Fjeldsoe et al., 2020; Mailey et al., 2010; Rovniak et al., 2002; Young et al., 2016). There have been suggestions however that the measurement of self-efficacy beliefs may

not always be consistent or appropriate within the context of human physical performance (Feltz et al., 2008). Bandura (1997, 2006) describes self-efficacy as a theoretical construct with an appropriate measurement model. Scales measuring self-efficacy within a specified domain should therefore be grounded in self-efficacy theory, designed based on rigorous psychometric principles and administered appropriately to maximize the validity of the scores produced. Well-constructed scales that are valid for their intended purposes and used as intended, are likely to provide substantial benefits to its users (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014).

Validity is a fundamental consideration in scale development and refers to the degree to which empirical evidence and theory support test scores and their use (AERA, APA, NCME, 2014). Valid measurement of psychological constructs is considered essential for the advancement of legitimate science in health psychology and related disciplines (Williams & Rhodes, 2016). As a result, using different scales with varying theoretical and measurement properties to measure physical activity self-efficacy may result in issues integrating knowledge obtained from research studies. In addition to using scales with high quality theoretical and measurement properties, studies should use scales as intended to maximize the validity of the scores produced.

This review of the measurement of physical activity self-efficacy beliefs is based on both conceptual and measurement-related recommendations for the development of theoretically grounded self-efficacy scales. The conceptual recommendations are focused on the extent to which the measurement of self-efficacy beliefs (e.g., scale content) corresponds with self-efficacy theory and includes the instructions and items comprising the scale. The measurement-

related recommendations refer to evidence for both reliability and validity of the measurement of the self-efficacy construct. These conceptual and measurement-related recommendations are based on Bandura's work (e.g., Bandura, 1997) along with extensions of his work as applied to the physical activity domain (e.g., Feltz & Chase, 1998). The alignment of conceptual and measurement recommendations is essential for improved measurement of psychosocial constructs such as self-efficacy (AERA, APA, NCME, 2014; Bandura, 1997, 2006).

Based on conceptual recommendations, self-efficacy beliefs should be specific to a domain of functioning (e.g., physical activity) and self-efficacy scales should be tailored to this specified domain (Feltz & Chase, 1998; Myers & Feltz, 2007). Concordance should therefore be established between the aspects (e.g., intensity levels) of self-efficacy beliefs and the domain of interest (e.g., physical activity). For example, if the measure of physical activity assesses intensity of physical activity the self-efficacy scale should also assess capability beliefs regarding different levels of physical activity intensity. Increased concordance improves the predictive validity of the scores produced by scales measuring self-efficacy beliefs (Feltz & Chase, 1998; Feltz et al., 2008; Myers & Feltz, 2007). Self-efficacy scales should measure both strength (i.e., degree of confidence) and levels (i.e., degree of situational demands, which could range from straightforward to burdensome) of self-efficacy beliefs (Bandura, 1997). Self-efficacy scale items should also be guided by an expert a priori conceptual analysis of the skills required for successful performance. The scale and scale label should reflect the content of the activity domain (Feltz & Chase, 1998; Myers & Feltz, 2007). The items comprising the scale should measure strength in one's belief in their capability to accomplish a task and therefore should be phrased in terms of capability (e.g., "can do") and not intention (e.g., "will do"). The items should only represent beliefs about personal abilities to produce specified levels of performance.

Preliminary instruction should also establish appropriate judgment based on one's current state (i.e., current capability) and not predictions about a future state (Bandura, 1997, 2006).

Based on measurement-related recommendations, empirical evidence should be provided for the dimensionality of self-efficacy scales. The empirical evidence for dimensionality should also have conceptual support. Not identifying dimensionality is likely to have implications for the validity and internal consistency of the scores produced by the measures (Bandura, 1997, 2006; Feltz & Chase, 1997; Feltz et al., 2008; Myers & Feltz, 2007). The internal consistency of self-efficacy scales should also be determined and reported (Bandura, 2006; Feltz & Chase, 1997). The validity of the scores produced by self-efficacy measurement should be determined by the consequences of self-efficacy judgements, that is, the extent to which levels of self-efficacy predict performance in a specified domain (Bandura, 1997, 2006; Feltz & Chase, 1997; Feltz et al., 2008; Myers & Feltz, 2007). The response scales of self-efficacy measures should also be effectively categorized to produce optimal psychometric characteristics. For example, measures of coaching efficacy using a 10-point Likert scale were found to be somewhat problematic (Myers et al., 2005). Subsequently, Myers et al. (2008) found evidence for improved validity of the scores produced by shorter rating scales (i.e., five or four categories) for measures of self-efficacy, as opposed to a minimum of 10 categories as suggested by Bandura (1997). Additionally, scholars caution against using single-item measures of self-efficacy beliefs within a specified domain (e.g., Myers & Feltz, 2007). In addition to problems with reliability and validity, single-item measures of self-efficacy beliefs may produce a restricted range of scores and reduce the predictive power of self-efficacy on performance outcomes (Bandura, 1997; Feltz & Chase, 1997).

Despite the consistent evidence supporting physical activity self-efficacy as a determinant of physical activity in physical activity-promoting interventions, there appears to be no current review of physical activity self-efficacy scales in adults. Systematic reviews provide evidence to stakeholders about the risk, harms, and benefits of interventions (Moher et al., 2009). Interventions typically involve experimental research designs such as randomized controlled trials which are generally considered to produce a higher quality of evidence than other designs (e.g., observational studies; Djulbegovic & Guyatt, 2017). Systematic reviews are therefore essential for advancing evidence-based practice (Alper & Haynes, 2016). Considering the potential issues in the measurement of physical activity self-efficacy in physical activity-promoting interventions, a systematic review of these scales as used in recent interventions (i.e., within the past 10 years) focused on quality measurement and theoretical grounding is important in improving the measurement of the construct thus potentially enhancing the benefits of these interventions.

The present systematic review intends to update and extend the current knowledge of the measurement of physical activity self-efficacy to inform interventions measuring physical activity self-efficacy. The present systematic review has three aims. First, to identify scales that have been used to measure physical activity self-efficacy in physical activity-promoting interventions. Second, to review conceptual, measurement-related, and administrative issues in the measurement of physical activity self-efficacy in physical activity-promoting interventions. Third, to make recommendations for the development and use of physical activity self-efficacy scales with improved quality.

## **METHODS**

### **Search Strategy**

The search strategy was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009). A systematic search of three electronic databases, Embase, EBSCOhost and PsycInfo was conducted in December 2020 in accordance with some recent recommendations (e.g., Bramer et al., 2017). Journal articles published in English from January 2010 to December 2020 that measured physical activity self-efficacy in physical activity-promoting interventions (e.g., longitudinal) designs were identified. A 10-year period was chosen as it corresponds with the time interval in which some key updated physical activity guidelines are based and published (e.g., USDHHS, 2018). Furthermore, this timeframe keeps the results relevant while based on the premise that increasing the timeframe is likely to increase the number of eligible studies without necessarily impacting the results in a meaningful way. The search terms used in all the database searches were chosen based on their relevance to the measurement of physical activity self-efficacy in physical activity-promoting interventions. The terms used in title (i.e., within the title of the manuscript) or topic (i.e., within the abstract, keywords etc.) searches were: “physical activity”, “self-efficacy”, “trial” and “intervention” (see Appendix A for the full electronic search strategy used for the Embase database). After duplicates were removed one reviewer screened the title and abstract of each retrieved article. The full text was reviewed, as necessary. Studies that met the eligibility criteria were selected.

### **Eligibility Criteria**

Peer reviewed journal articles reporting physical activity-promoting interventions measuring physical activity self-efficacy as a mechanism for increasing physical activity behavior were included. Once the authors reported the measurement of any form of physical



activity (e.g., exercise) in the study it was deemed eligible based on this criterion. These longitudinal studies included randomized controlled trials, effectiveness trials, pretest-posttest designs, feasibility trials and pilot studies. Included studies have samples of adults between the ages of 18 and 65 years of age (adult or middle-aged). Studies were selected and extracted that reported the measurement of physical activity self-efficacy by a self-report scale administered in English. Scales that were translated between English and other languages were not included. Scales designed for specific, well-defined sub-populations in clinical settings (e.g., traumatic brain injury patients) were also not included.

### **Data Extraction**

Data extraction and scale reviews were overseen by a senior reviewer with expertise in the measurement of self-efficacy associated with physical activity in physical activity-promoting interventions. One reviewer conducted the initial data extraction, and a second reviewer reviewed the initial data extraction results. The second reviewer randomly sampled 25% of the studies selected as meeting the eligibility criteria and disagreements were resolved by discussion with the third expert reviewer when necessary. Inter-rater reliability was then estimated, using proportion of agreement, based on the initial scores of the first two raters. Proportion of agreement was calculated by dividing the consistent classifications of both reviewers by the total number of classifications. Each physical activity self-efficacy scale that was identified and met the eligibility criteria was extracted. The original source (as cited in the extracted study) of the development of each physical activity self-efficacy scale was then found. The reviewers then also attempted to obtain the complete scales used to measure physical activity self-efficacy and physical activity from the original sources. If the required information was not available in the cited original study the authors of the studies were contacted and the additional necessary

information requested. Once a complete survey battery was received from the authors of a study it was used in the review. If the battery was not received the reviewers extracted the necessary information from the published study. The initial source of the scale was the focal point of the review process based on the premise that all subsequent versions of the scale in question would be directly or indirectly affected by the initial study presenting the development of the scale. Each extracted study was assessed for conceptual, measurement-related, and administrative quality.

### **Conceptual and Measurement-related Examination of Physical Activity Self-efficacy**

A reviewer rating scale (see Appendix B) was developed (because the authors were unaware of a relevant existing instrument with validity evidence) based on a review of the literature on self-efficacy theory and measurement as outlined in the introduction. This new approach was an initial attempt to provide a practical and objective descriptive tool to rate the conceptual and measurement quality of the scales. The items comprising the reviewer rating scale were developed, discussed, iteratively modified, and agreed upon by three reviewers. The reviewer rating scale was used to assess the extent to which each extracted self-efficacy scale and their cited studies are in alignment with self-efficacy theory and measurement. The reviewer rating scale consisted of 17 items (e.g., “the self-efficacy instrument measures beliefs specific to the physical activity domain”). The items were rated on a 3-point Likert scale. A score of 2 (i.e., yes) for an item indicates that the scale ideally satisfied the criterion, a score of 1 (i.e., somewhat) indicates that the scale did not completely satisfy the ideal criterion and a score of 0 (i.e., no) indicates that the criterion was not satisfied.

Three reviewers conducted the rating of the conceptual and measurement-related aspects of the physical activity self-efficacy scale using the reviewer rating scale. Each reviewer

independently reviewed the cited studies in which the scales were described including the content (e.g., instructions, items, and rating scale) of the corresponding physical activity self-efficacy scales. The reviewers then independently assessed each study using the reviewer rating scale. Finally, the reviewers as a group, justified each of their item ratings and a final score for each item in the reviewer rating scale was determined for each study by discussion and consensus. All three reviewers were required to agree to the final score on each item in the reviewer rating scale. A total consensus score was not calculated for the rating of each self-efficacy scale because the purpose of the study was to describe (not compare scale scores) the theoretical and measurement quality of scales measuring physical activity self-efficacy in physical activity-promoting interventions with our newly developed reviewer rating scale.

### **Examination of Administrative Criteria**

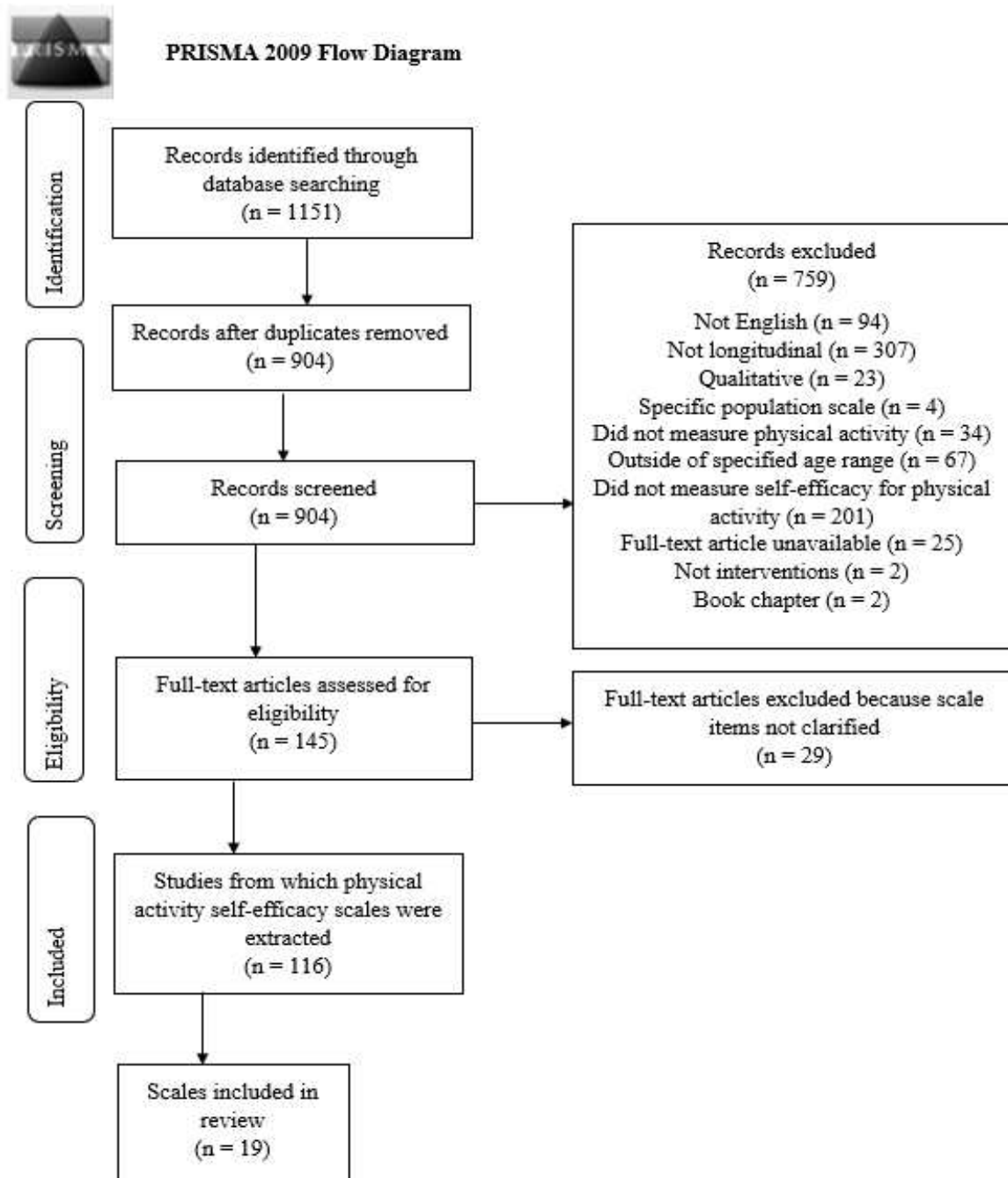
The following administrative criteria were assessed for each identified scale: (a) time to administer, (b) ease of scoring and (c) readability and comprehension. The rating of these administrative criteria was assessed independently by two reviewers who completed each scale on their own. Inter-rater reliability was then estimated, using proportion of agreement, based on the scores of both raters. The reviewers discussed each of their ratings and a final score was determined for each scale by discussion and consensus. Each scale achieved a rating within the range of 0 to 3. A score of 3 indicates that satisfaction of the criterion was ideal, a score of 2 indicates that the scale fell short of the ideal criterion, a score of 1 indicates that the scale was significantly below the acceptable criterion and a score of 0 indicates that the criterion was not satisfied. The criteria used to assess administrative quality, though not typically used as a requirement for the assessment of self-reported motivational constructs were deemed a useful guide in optimizing the quality of measurement.

Time to administer refers to the time needed to complete the measure (Bot et al., 2004). For a positive rating of time to administer, participants were able to complete the scale in less than or equal to 10 minutes (Terwee et al., 2007). Ease of scoring refers to the extent the scale can be scored by a trained investigator (Bot et al., 2004). For a positive rating, the scale had to be generated by summing items, or the formula used to compute total score had to be a simple one, such as the reversal of specific items (Terwee et al., 2007). Readability and comprehension refer to the extent to which the wording and language of the scale is understandable for all participants (Terwee et al., 2007). For a positive rating, the scale had to be comprised of content that is easily understood by most adults, that is, at approximately the grade six level.

## **RESULTS**

The results of the database search, screening and study inclusion are included in Figure 1. Fourteen multi-item and five single-item distinct physical activity self-efficacy scales were extracted and reviewed from the 116 manuscripts that met the inclusion criteria. Inter-rater agreement for data extraction was found to be substantial, where the observed proportion of agreement,  $P = .87$ . The psychometric information for the exercise self-efficacy scale (ESE; Bandura, 2006) and the self-efficacy inventory (SEI) were recovered from Everett et al. (2009) and Lipschitz et al. (2015) respectively. Extracted scales that were described as a modified version of another original scale were included as part of the assessment of the original scale since the initial study from which the original scale was developed is likely to have a notable influence on the modified versions of the scale.

Figure 1. Schematic review of study and scale selection based on the PRISMA guidance.



## Descriptive Results

Fourteen multi-item scales were identified and extracted as measuring self-efficacy associated with physical activity by self-report (see Table 1). The most frequently used (28

instances) scale was the self-efficacy for physical activity (SEPA; Marcus et al., 1992). Eleven of the fourteen multi-item scales were clearly presented as measuring a construct described as self-efficacy related to engaging in physical activity despite barriers or specific situations (e.g., bad weather). These scales (e.g., the ESE) were collectively classified as measuring self-efficacy to regulate physical activity. Three scales: 1) the barriers specific self-efficacy scale (BARSE) combined with the exercise self-efficacy scale (EXSE), 2) the self-efficacy related to perceived behavioral control scale (SE-PBC), and 3) the Wooldridge self-efficacy scale (W-SES) did not clearly appear to measure self-efficacy to regulate physical activity. For example, the items of the EXSE subscale of the BARSE combined with the EXSE scale appear to measure efficacy for varying levels of exercise (i.e., task-related physical activity self-efficacy). The number of items in the identified scales ranged from three to 23. Two scales, the health behavior scale (HBS) and self-efficacy for exercise behaviors (SEB) were reported as being multidimensional, both consisting of two subscales. The extracted single-item physical activity self-efficacy scales generally required respondents to indicate their level of confidence in their ability to engage in physical activity or exercise over a certain specified timeframe. For example, the single-item measure of physical activity self-efficacy in Mosher et al. (2013) was worded as follows, “How sure are you that you could exercise at least 30 min a day at least 5 days a week?”.

Table 1. *Multi-item Scales Identified and Extracted with Construct Descriptions and Other Characteristics*

| Study | Scale | Construct Description as Reported by Author                          | Sub-scales | No. of Items | No. of Studies |
|-------|-------|--|------------|--------------|----------------|
| 1     | ESE   | SE to regulate exercise  | 1          | 18           | 11             |
| 2     | SEB   | SE for exercise behavior adoption and maintenance                    | 2          | 12           | 21             |
| 3     | SEE   | SE for exercise  | 1          | 9            | 10             |
| 4     | SEPA  | Confidence in ability to persist with exercise in various situations | 1          | 5            | 28             |

Table 1 (cont'd)

|    |            |   |   |    |    |
|----|------------|---|---|----|----|
| 5  | BARSE      | Perceived capabilities to exercise in the face of barriers to participation   | 1 | 13 | 12 |
| 6  | BARSE+EXSE | Efficacy with respect to continued exercise participation                     | 1 | 8  | 12 |
| 7  | SEQ        | Confidence in one's ability to exercise when faced with potential barriers    | 1 | 16 | 7  |
| 8  | HBS        | SE to face barriers   | 2 | 23 | 3  |
| 9  | P-SES      | SE  | 1 | 8  | 2  |
| 10 | SEI        | Confidence to engage in regular PA across a variety of challenging situations | 1 | 6  | 1  |
| 11 | SE-PBC     | SE component of PBC   | 1 | 4  | 2  |
| 12 | PAAI       | SE for PA   | 1 | 13 | 2  |
| 13 | C-SES      | Confidence to be physically active  | 1 | 3  | 1  |
| 14 | W-SES      | SE for PA   | 1 | 4  | 1  |

*Note.* ESE = Exercise Self-efficacy, SEB = Self-efficacy for Exercise Behaviors, SEE = Self-efficacy for Exercise Scale, SEPA = Self-efficacy for Physical Activity, BARSE = Barriers Specific Self-efficacy Scale, EXSE = Exercise Self-efficacy Scale, SEQ = Self-efficacy for Exercise Questionnaire, HBS = Health Behavior Scale, P-SES = Plotnikoff Self-efficacy Scale, SEI = Self-efficacy Inventory, SE-PBC = Self-efficacy Related to Perceived Behavioral Control, PAAI = Physical Activity Assessment Inventory, C-SES = Craig Self-efficacy Scale, W-SES = Wooldridge Self-efficacy Scale, N/A = Not Applicable, 0 = No, 1 = Somewhat, 2 = Yes.

### **Conceptual and Measurement-Related Aspects of Physical Activity Self-efficacy**

The results of the conceptual-related assessment of each study in which each of the extracted 14, multi-item physical activity self-efficacy instruments are cited are presented in Table 2. Items 15, 16 and 17 of the reviewer rating scale correspond to the measurement-related recommendations while the other items refer to the conceptual recommendations as described previously. The results of the measurement-related assessment are summarized in Table 3. The following paragraphs summarize the results of the assessment of the physical activity self-efficacy instruments using the reviewer rating scale (see Appendix B for more details of score justifications).

Table 2. Reviewer Consensus Scores for the Conceptual-related Assessment the of Extracted Scales as Described in their Cited Studies

|      | ESE<br>Study<br>1 | SEB<br>Study<br>2 | SEE<br>Study<br>3 | SEPA<br>Study<br>4 | BARSE<br>Study<br>5 | BARSE<br>+ EXSE<br>Study 6 | SEQ<br>Study<br>7 | HBS<br>Study<br>8 | P-SES<br>Study<br>9 | SEI<br>Study<br>10 |
|------|-------------------|-------------------|-------------------|--------------------|---------------------|----------------------------|-------------------|-------------------|---------------------|--------------------|
| Item |                   |                   |                   |                    |                     |                            |                   |                   |                     |                    |
| 1    | 2                 | 1                 | 2                 | 2                  | 2                   | 2                          | 2                 | 2                 | 2                   | 1                  |
| 2    | 1                 | 1                 | 2                 | 1                  | 1                   | 1                          | 1                 | 1                 | 1                   | 1                  |
| 3    | 0                 | 2                 | 0                 | 0                  | 2                   | 2                          | 2                 | 2                 | 0                   | 0                  |
| 4    | 2                 | 2                 | 2                 | 2                  | 2                   | 2                          | 2                 | 2                 | 2                   | 2                  |
| 5    | 2                 | 2                 | 2                 | 1                  | 2                   | 2                          | 1                 | 2                 | 2                   | 1                  |
| 6    | 2                 | 1                 | 2                 | 1                  | 2                   | 2                          | 1                 | 2                 | 2                   | 1                  |
| 7    | 0                 | 1                 | 0                 | 0                  | 0                   | 1                          | 1                 | 1                 | 2                   | 1                  |
| 8    | 0                 | 1                 | 2                 | 0                  | 0                   | 1                          | 1                 | 0                 | 2                   | 1                  |
| 9    | 2                 | 2                 | 2                 | 2                  | 2                   | 2                          | 2                 | 2                 | 2                   | 2                  |
| 10   | 2                 | 2                 | 2                 | 2                  | 2                   | 1                          | 2                 | 2                 | 2                   | 2                  |
| 11   | 0                 | 2                 | 0                 | 2                  | 2                   | 2                          | 0                 | 0                 | 0                   | 0                  |
| 12   | 0                 | 1                 | 1                 | 1                  | 2                   | 1                          | 1                 | 1                 | 2                   | 1                  |
| 13   | 0                 | 1                 | 0                 | 2                  | 1                   | 1                          | 1                 | 1                 | 2                   | 1                  |
| 14   | 0                 | 0                 | 2                 | 1                  | 1                   | 1                          | 1                 | 1                 | 2                   | 1                  |

|      | SE-PBC<br>Study<br>11 | PAAI<br>Study<br>12 | C-SES<br>Study<br>13 | W-SES<br>Study<br>14 |
|------|-----------------------|---------------------|----------------------|----------------------|
| Item |                       |                     |                      |                      |
| 1    | 1                     | 2                   | 1                    | 1                    |
| 2    | 1                     | 1                   | 1                    | 1                    |
| 3    | 2                     | 0                   | 0                    | 2                    |
| 4    | 2                     | 2                   | 2                    | 1                    |
| 5    | 2                     | 1                   | 0                    | 1                    |
| 6    | 2                     | 1                   | 0                    | 1                    |
| 7    | 2                     | 0                   | 0                    | 0                    |
| 8    | 2                     | 0                   | 0                    | 0                    |
| 9    | 2                     | 2                   | 2                    | 2                    |
| 10   | 1                     | 2                   | 2                    | 0                    |
| 11   | 0                     | 2                   | 0                    | 2                    |
| 12   | N/A                   | 1                   | 0                    | 0                    |
| 13   | N/A                   | 2                   | 2                    | 0                    |
| 14   | N/A                   | 2                   | 0                    | 0                    |



Table 2 (cont'd)

*Note.* ESE = Exercise Self-efficacy, SEB = Self-efficacy for Exercise Behaviors, SEE = Self-efficacy for Exercise Scale, SEPA = Self-efficacy for Physical Activity, BARSE = Barriers Specific Self-efficacy Scale, EXSE = Exercise Self-efficacy Scale, SEQ = Self-efficacy for Exercise Questionnaire, HBS = Health Behavior Scale, P-SES = Plotnikoff Self-efficacy Scale, SEI = Self-efficacy Inventory, SE-PBC = Self-efficacy Related to Perceived Behavioral Control, PAAI = Physical Activity Assessment Inventory, C-SES = Craig Self-efficacy Scale, W-SES = Wooldridge Self-efficacy Scale, N/A = Not Applicable, 0 = No, 1 = Somewhat, 2 = Yes.

Table 3. *Measurement-Related Properties of Extracted Scales*

| Scale        | Administrative Properties     |                         |                                      |
|--------------|-------------------------------|-------------------------|--------------------------------------|
|              | Internal Consistency Evidence | Dimensionality Evidence | Response Scale Optimally Categorized |
| ESE          | +                             | +                       | -                                    |
| SEB          | +                             | +                       | +                                    |
| SEE          | +                             | -                       | -                                    |
| SEPA         | +                             | +/-                     | -                                    |
| BARSE        | +                             | -                       | -                                    |
| BARSE + EXSE | +                             | +/-                     | -                                    |
| SEQ          | +                             | -                       | -                                    |
| HBS          | +                             | +                       | -                                    |
| P-SES        | +                             | -                       | +                                    |
| SEI          | -                             | -                       | +                                    |
| SE-PBC       | -                             | +                       | -                                    |
| PAAI         | +                             | +                       | -                                    |
| C-SES        | +                             | -                       | -                                    |
| W-SES        | +                             | -                       | -                                    |

*Note.* ESE = Exercise Self-efficacy, SEB = Self-efficacy for Exercise Behaviors, SEE = Self-efficacy for Exercise Scale, SEPA = Self-efficacy for Physical Activity, BARSE = Barriers Specific Self-efficacy Scale, EXSE = Exercise Self-efficacy Scale, SEQ = Self-efficacy for Exercise Questionnaire, HBS = Health Behavior Scale, P-SES = Plotnikoff Self-efficacy Scale, SEI = Self-efficacy Inventory, SE-PBC = Self-efficacy Related to Perceived Behavioral Control, PAAI = Physical Activity Assessment Inventory, C-SES = Craig Self-efficacy Scale, W-SES = Wooldridge Self-efficacy Scale, Somewhat = 1, + = Yes = 2.

Most of the self-efficacy scales (n = 8) clearly referred to at least one specific level of physical activity. For example, the BARSE (McAuley, 1992) clearly specifies frequency (i.e., "3 times per week") of exercise as a level of physical activity performance to which the self-efficacy beliefs refer. Frequency (e.g., at least 3 times per week) was the most common (n = 7),

clearly referred to level of physical activity. Duration ( $n = 3$ ) and intensity ( $n = 2$ ) were less commonly specified. For example, the ESE (Everett et al., 2009) specifies frequency of physical activity (i.e., most days of the week), the P-SES (Plotnikoff et al., 2001) specifies intensity of physical activity (i.e., vigorous) and the self-efficacy for exercise scale (SEE; Resnick & Jenkins, 2000) specifies duration (i.e., for 20 minutes). Two studies clearly specified the three levels of physical activity to which the self-efficacy scale items refer. For example, Rhodes and Courneya (2003) mentioned that participants were asked to use a definition of regular exercise which specified levels of physical activity (i.e., frequency, intensity, duration). Though specifying the three levels of physical activity, the SE-PBC (Rhodes & Courneya, 2003) was not concordant with a measure of physical activity since physical activity did not appear to be measured as reported in the manuscript. Only in one study, the P-SES (Plotnikoff et al., 2001), was the measure of physical activity self-efficacy completely concordant with the measure of physical activity for each level of physical activity. The P-SES referred to regular exercise, which was clearly defined in terms of frequency (i.e., at least three times per week), intensity (i.e., providing examples of vigorous exercise) and duration (i.e., at least 20 minutes each time). The stage of change measure of physical activity used in the study also referred to the same specific definition of regular exercise and was therefore completely concordant with the measure of self-efficacy for frequency, intensity, and duration.

Most studies ( $n = 11$ ) clearly explained that the primary focus of the self-efficacy scale was to measure personal capability to overcome barriers to physical activity performance. For example, the items and instructions of the Craig self-efficacy scale (Craig et al., 2015) require respondents to indicate their confidence to be physically active despite three barriers. These barriers were: (a) “no matter how busy your day is”, (b) “on a day when you don’t really

feel like doing it” and (c) “and still spend time with your family”. Six studies clearly reported that a conceptual analysis (including modification of existing scales) was conducted (within the same study) to determine the items comprising the self-efficacy scale. For example, the physical activity assessment inventory (PAAI) as reported in Haas & Northam (2010) was developed according to guidelines from Bandura (1997). Additionally, the items comprising the PAAI were derived from the literature and revised based on the advice of expert reviewers who were consulted to optimize content validity (Haas & Northam, 2010).

All the studies specified a form of physical activity (e.g., exercise) as the performance domain to which the self-efficacy scale items refer. For example, the instructions of the SEPA scale (Marcus et al., 1992) clearly specified that the items referred to participation in regular exercise. Most of the self-efficacy scales (n = 9) clearly emphasized the measurement of appropriate capability judgment (i.e., “can do”). For example, the instructions of the HBS (Anderson et al., 2007) clearly indicates appropriate capability judgement using the statement “can do” to precede the actions required or barriers to overcome in order to engage in physical activity. Most of the studies (n = 13) clearly emphasized the measurement of strength of personal capability beliefs. For example, the instructions of the self-efficacy questionnaire (Garcia & King, 1991) clearly refer to degree (i.e., strength) of confidence using the question “How confident are you that you could exercise under each of the following conditions over the next 6 months?”.

Some scales (n = 7) clearly emphasized a future time for which the self-efficacy beliefs were being measured. For example, McAuley (1993) which presents the BARSE combined with the EXSE included the statements "next 3 months" and "next 1-12 weeks" in their self-efficacy scale instructions which indicates that the items refer to a future time. Most studies (n = 13) used

present tense language to word the instructions and/or items which may imply an evaluation of current capability. For example, the SEQ used the wording, “How confident are you that you could exercise under each of the following conditions over the next 6 months?” which is worded in the present tense while referring to a future timeframe. Only one study however, clearly emphasized that the self-efficacy scale items referred specifically to a current (i.e., right now) evaluation of capability. The instructions and items of the SEE scale (Resnick & Jenkins, 2000) indicated the measurement of current judgement by including the statement “right now” in clarifying the time-period to which the items comprising the scale refer.

From a measurement-related perspective, most studies (n = 12) reported evidence of internal consistency. For example, the W-SES (Wooldridge et al., 2019) reported internal consistency, Cronbach’s alpha of 0.83 and 0.89 at two measurement occasions. Five studies reported complete evidence for the dimensionality of the physical activity self-efficacy scale. For example, Sallis et al. (1988) reported factor analysis evidence which supported a two-factor structure of the SEB, where one factor was labeled “resisting relapse” and the other labeled “making time for exercise”. Finally, only three studies were assessed as having response scales that were optimally categorized (i.e., consisting of five or less categories). For example, the items comprising the SEI (Lipschitz et al., 2015) consisted of a response scale with 5 categories. The categories were labeled from 1 to 5, where 1 = not at all confident, 2 = somewhat confident, 3 = moderately confident, 4 = very confident, and 5 = completely confident.

### **Examination of Administrative Criteria**

The proportion of agreement, P for time to administer, ease of scoring and readability and comprehension were .80, 1.0, and .93, respectively. The self-efficacy scales showed good administrative properties in general (see Table 4). All scales utilized a Likert scale from which

responses were either summed or a mean score was calculated. All scales received maximum scores for ease of scoring. Only one scale SE-PBC (Rhodes & Courneya, 2003) did not receive a maximum score for readability and comprehension, because the items comprising this scale consist of varying stem patterns. Four scales received a score of 2 for time to administer. The SEB (Sallis et al., 1988), the ESE (Bandura, 2006), the BARSE combined with the EXSE (McAuley, 1993) and the HBS (Anderson et al., 2007) are likely to require respondents more than 10 minutes to read, interpret and respond to all items.

Table 4. *Administrative Properties of Extracted Scales*

| Scale        | Administrative Properties |                 |                               |
|--------------|---------------------------|-----------------|-------------------------------|
|              | Time to Administer        | Ease of Scoring | Readability and Comprehension |
| ESE          | 2                         | 3               | 3                             |
| SEB          | 2                         | 3               | 3                             |
| SEE          | 3                         | 3               | 3                             |
| SEPA         | 3                         | 3               | 3                             |
| BARSE        | 3                         | 3               | 3                             |
| BARSE + EXSE | 2                         | 3               | 3                             |
| SEQ          | 3                         | 3               | 3                             |
| HBS          | 2                         | 3               | 3                             |
| P-SES        | 3                         | 3               | 3                             |
| SEI          | 3                         | 3               | 3                             |
| SE-PBC       | 3                         | 3               | 2                             |
| PAAI         | 3                         | 3               | 3                             |
| C-SES        | 3                         | 3               | 3                             |
| W-SES        | 3                         | 3               | 3                             |

*Note.* ESE = Exercise Self-efficacy, SEB = Self-efficacy for Exercise Behaviors, SEE = Self-efficacy for Exercise Scale, SEPA = Self-efficacy for Physical Activity, BARSE = Barriers Specific Self-efficacy Scale, EXSE = Exercise Self-efficacy Scale, SEQ = Self-efficacy for Exercise Questionnaire, HBS = Health Behavior Scale, P-SES = Plotnikoff Self-efficacy Scale, SEI = Self-efficacy Inventory, SE-PBC = Self-efficacy Related to Perceived Behavioral Control, PAAI = Physical Activity Assessment Inventory, C-SES = Craig Self-efficacy Scale, W-SES = Wooldridge Self-efficacy Scale, Somewhat = 1, + = Yes = 2.

## DISCUSSION

The overarching aim of the present systematic review is to update and extend the current knowledge of the measurement of physical activity self-efficacy in physical activity-promoting interventions using the PRISMA guidelines. The review uncovered several outcomes, however, six major findings will be emphasized. First, from a conceptual perspective, comprehensively emphasizing the levels of physical activity to which the self-efficacy scale is referring to and ensuring concordance of those levels with the measure of physical activity appears to be a major limitation of physical activity self-efficacy measurement. Second, the construct label given to scales measuring self-efficacy associated with physical activity may be misleading and theoretically imprecise. Third, most studies were not clear in emphasizing the measurement of current capability judgement. Fourth, from a measurement perspective, determining the dimensions of the physical activity self-efficacy scale appears to be a major limitation. Fifth, from a measurement perspective, the response scales of most physical activity self-efficacy scales may not be optimally categorized, and the use of single-item physical activity scales is concerning. Sixth, the administrative properties of physical activity self-efficacy scales are generally acceptable. The following six paragraphs discuss each of these six major findings.

Most scales in this review reflected strength of capability beliefs but did not clearly establish levels of physical activity performance to which the scale refers. Most scales therefore had a lack of concordance with the physical activity measure in terms of physical activity levels. Bandura (1997, 2006) emphasizes that a necessary condition for valid measurement of self-efficacy is concordance between the domain-specific self-efficacy beliefs and the proposed outcome of interest. Measures of self-efficacy beliefs within a specified domain of interest should therefore be concordant with the external variable (e.g., physical activity) that represents

the performance domain (Feltz & Chase, 1998; Myers & Feltz, 2007). Feltz and Chase (1997) further assert that the ability of self-efficacy measures to predict performance decreases as the concordance between the efficacy beliefs and the performance domain decreases.

Most scales in this review purporting to measure self-efficacy associated with physical activity appear to be primarily measuring self-efficacy to regulate physical activity. Bandura (1997) distinguishes between task-related self-efficacy beliefs (e.g., self-efficacy for varying levels of physical activity) and self-regulatory efficacy (e.g., engaging in physical activity despite barriers). These distinct forms of self-efficacy judgments are associated with the physical activity performance domain therefore each emphasize different aspects of personal capability beliefs, both of which are relevant in exercise settings (Anderson et al., 2007; Brawley et al., 2011; Martin Ginis et al., 2011; McAuley, 1992, 1993; Megakli et al., 2017). Consequently, since the constructs are conceptually distinct from each other the content (e.g., items and instructions) and labels for scales measuring each should accurately indicate whether the scale is measuring task-related physical activity self-efficacy or self-efficacy to regulate physical activity behavior. Making this distinction also prevents the jingle-jangle fallacy, that is, using scales with similar names that might measure different constructs (jingle fallacy) and/or using scales with dissimilar labels that might measure similar constructs (jangle fallacy; Marsh 1994).

Furthermore, at least one recent intervention study, Myers et al. (2020) has distinguished between and measured both physical activity self-efficacy and self-efficacy to regulate physical activity and found evidence that improving self-efficacy to regulate physical activity may be effective in increasing physical activity behavior in adults with obesity. Additionally, some scholars have identified other forms of self-efficacy beliefs which may also be relevant in physical activity settings such as coping self-efficacy and scheduling self-efficacy (e.g.,

DuCharme & Brawley, 1995; Rodgers et al., 2008). The content and labels of scales measuring these, and other similar constructs should also be distinct from each other while still aligning with self-efficacy theory.

While most studies in this review stipulated the measurement of appropriate judgment of personal capabilities specific to the physical activity domain, most studies were not clear about the timeframe to which these beliefs refer. This lack of emphasis on current capabilities while referring to a future timeframe may have the unintended effect of influencing participants to base their responses on expected future capabilities. Self-efficacy measurement should however establish appropriate judgment based on one's current state and not one's potential or expected future capabilities (Bandura, 1997, 2006). This specification distinguishes self-efficacy (can do) from intention (will do; Bandura, 1997, 2006) and other related but different constructs, such as: motivation, locus of control, outcome expectancies, self-esteem, or self-worth (Bandura, 2006; Beauchamp, 2016; Gist & Mitchell, 1992; Williams & Rhodes, 2016).

Most of the studies reviewed did not provide complete evidence for the dimensionality of the physical activity self-efficacy scale used. Determining the dimensionality of self-efficacy scales is important to establish the factor structure of the items and empirically establish construct validity (Bandura, 1997, 2006; Feltz & Chase, 1997; Feltz et al., 2008; Myers & Feltz, 2007). Empirical tests of dimensionality can also be used to confirm expected a priori conceptual factor structures. Not confirming the factor structure of the physical activity self-efficacy scale is likely to reduce the performance predicting effect of the scale. Consequently, the scale will have less validity in accurately monitoring changes in physical activity self-efficacy and in predicting physical activity behavior during physical activity-promoting interventions.



The response scales of most of the multi-item self-efficacy scales in this review were not assessed as being optimally categorized. Though Bandura (1997) suggests a minimum of 10 response categories, analysis of the response categories of self-efficacy scales in some domains have indicated that more (e.g., greater than five) response categories may be associated with lowered validity (e.g., Myers et al., 2008; Zhu & Kang, 1998). Further analysis of the optimal response categories in physical activity self-efficacy rating scales may therefore be necessary. Additionally, five single-item scales purported to measure self-efficacy associated with physical activity were identified in this review. Generally, the measurement of psychological constructs using single-item scales is presumed to have unacceptably low or unknown reliability and is typically discouraged (Wanous et al., 1997). Self-efficacy theorists also warn against measuring self-efficacy beliefs using a single item. Single-item self-efficacy scales have been found to reduce the predictive power of self-efficacy beliefs on performance and fail to differentiate between differing levels of personal efficacy (Bandura, 1997). Additionally, single-item self-efficacy scales have been reported to have problems with reliability and validity (Feltz & Chase, 1997; Myers & Feltz, 2007). Furthermore, the most used index of reliability, the internal consistency coefficient, Cronbach's alpha (Cronbach, 1951) cannot be calculated for a single-item scale. Consequently, no evidence of scale reliability was provided in any reviewed study using a single-item measure of physical activity self-efficacy.

The administrative properties of the physical activity self-efficacy scales were consistently good and suggest good practical utility of the scales. It is notable however, that evidence of an examination of administrative properties was not reported in any of the studies cited for any of the scales examined. Good administrative properties are likely to reduce participant burden when completing these self-report measures while increasing participants'

understanding of the items and scale instructions. Participant responses will therefore likely be more reflective of their internal processes and result in increased interpretability and validity of scores.

The present review highlighted five other noteworthy findings. First, each of the multi-item scales identified consists of varying numbers of items and other characteristics. This use of several different scales purported to measure a similar construct is likely to result in inconsistencies in the literature as suggested by Feltz et al. (2008). Second, some intervention studies (e.g., Murru & Ginis, 2010) report using modified versions of existing scales and may not report thorough evidence of the modified scale's reliability and validity in the new study. Voskuil et al. (2017) emphasize that adapting physical activity self-efficacy scales without conducting psychometric analyses on the modified scales may affect the measurement properties of the scale. Third, though most studies reported the internal consistency of the physical activity self-efficacy scales, there are limitations with the use of Cronbach's alpha as an indicator of reliability using a traditional observed score approach (Raykov & Marcoulides, 2019). A useful alternative may be to model physical activity self-efficacy as a latent variable within a structural equation modeling framework. Fourth, most intervention studies appear to cite the original source of the development of the physical activity self-efficacy scale being used. These sources, however, may not adequately report the measurement properties of the scale and may also involve a sample with characteristics that are different from the intervention study. As a result, there may be limited evidence of the validity of the self-efficacy measurement in the study to provide a sound scientific basis for proposed score interpretations and subsequent decision-making (AERA, APA, NCME, 2014). Fifth, most of the self-efficacy scales included in this review were assessed as being related to the physical activity domain, however most studies

presenting the development of the scales did not provide evidence of an a priori conceptual analysis that determined the items comprising the self-efficacy scale. Failure to determine and incorporate the skills hypothesized to be necessary for successful performance into the self-efficacy measurement is likely to limit the predictive validity of the self-efficacy measure on performance (Bandura, 1997; Feltz, et al., 2008).

### **Limitations**

The present review focused on examining the initial publication of each physical activity self-efficacy scale extracted from review studies meeting the search criteria. This approach however, may have the effect of not completely accounting for modifications and adaptations that have been subsequently made to these instruments. Future studies reviewing the quality of modified versions of specific scales in specific contexts may be useful. The present review focused on the aspects of physical activity self-efficacy measurement that are mostly dependent on the scale itself such as the item content and response scales. The validity of a measurement however, is also dependent on other factors such as characteristics of the subjects being measured. Future studies reviewing a representative sample of all the intervention studies using specific physical activity self-efficacy scales would provide further evidence for the validity of those scales. The measurement focus of the present review did not allow for the analysis and synthesis of outcome results (e.g., effect sizes) in the included intervention studies. Future studies (e.g., meta-analyses) reviewing quality measurement of physical activity self-efficacy and its effect on the outcome of interest (i.e., physical activity behavior) would be helpful in clarifying the predictive validity of measured self-efficacy judgements on the physical activity behavior of participants.

## **Conclusion and Recommendations**

Alignment of the conceptual and measurement-related characteristics of a psychosocial construct is essential for ensuring that test scores and their application are valid (AERA, APA, NCME, 2014). Consequently, appropriate attention should be given to the conceptual and measurement-related aspects of physical activity self-efficacy measurement in research aimed at using increasing physical activity self-efficacy to promote physical activity behavior. It appears however that a few issues exist with the measurement of self-efficacy associated with physical activity in physical activity-promoting interventions.

To improve the measurement of self-efficacy associated with physical activity, researchers should specify whether they are measuring physical activity self-efficacy (task-related) or the self-regulatory aspects of self-efficacy (e.g., overcoming barriers such as bad weather). Researchers may also benefit from measuring both the task and self-regulatory aspects of self-efficacy and determine if both should be combined into a single scale with a composite score as a comprehensive measure of self-efficacy associated with physical activity. Researchers should also ensure concordance between the self-efficacy measure and the outcome measure of physical activity by ensuring that the levels of physical activity (e.g., intensity) being measured are represented in the self-efficacy scale. Reporting evidence of reliability and dimensionality of scores is also recommended, especially when modifications are made to an original scale and the modified scale is used in a target population different from the population for which it was originally designed. Researchers should avoid using single-item scales to measure self-efficacy associated with physical activity due to issues of unknown reliability and compromised validity. Finally, scales with the most comprehensive combination of good theoretical, psychometric,

administrative, and other desirable properties should be given priority for use in intervention studies.

## REFERENCES

## REFERENCES

- Alper, B. S., & Haynes, R. B. (2016). EBHC pyramid 5.0 for accessing preappraised evidence and guidance. *BMJ evidence-based medicine*, *21*(4), 123-125.  
<http://dx.doi.org/10.1136/ebmed-2016-110447>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderson, E. S., Winett, R. A., & Wojcik, J. R. (2007). Self-regulation, self-efficacy, outcome expectations, and social support: Social cognitive theory and nutrition behavior. *Annals of Behavioral Medicine*, *34*(3), 304-312.  
<https://psycnet.apa.org/doi/10.1007/BF02874555>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.
- Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J., Martin, B. W., & Lancet Physical Activity Series Working Group. (2012). Correlates of physical activity: why are some people physically active and others not?. *The lancet*, *380*(9838), 258-271.  
[https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)
- Beauchamp, M. R. (2016). Disentangling motivation from self-efficacy: Implications for measurement, theory-development, and intervention. *Health Psychology Review*, *10*(2), 129-132. <https://doi.org/10.1080/17437199.2016.1162666>
- Bot, S. D. M., Terwee, C. B., van der Windt, D A W M, Bouter, L. M., Dekker, J., & de Vet, H C W. (2004). Clinimetric evaluation of shoulder disability questionnaires: A systematic review of the literature. *Annals of the Rheumatic Diseases*, *63*(4), 335-341.  
<http://dx.doi.org/10.1136/ard.2003.007724>
- Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic reviews*, *6*(1), 1-12. <https://doi.org/10.1186/s13643-017-0644-y>
- Brawley, L., Rejeski, W. J., Gaukstern, J. E., & Ambrosius, W. T. (2012). Social cognitive changes following weight loss and physical activity interventions in obese, older adults in poor cardiovascular health. *Annals of Behavioral Medicine*, *44*(3), 353-364.  
<https://doi.org/10.1007/s12160-012-9390-5>

- Craig, C. L., Bauman, A., Latimer-Cheung, A., Rhodes, R. E., Faulkner, G., Berry, T. R., ... & Spence, J. C. (2015). An evaluation of the My ParticipACTION campaign to increase self-efficacy for being more physically active. *Journal of health communication, 20*(9), 995-1003. doi: 10.1080/10810730.2015.1012240
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Djulbegovic, B., & Guyatt, G. H. (2017). Progress in evidence-based medicine: a quarter century on. *The Lancet, 390*(10092), 415-423. [https://doi.org/10.1016/S0140-6736\(16\)31592-6](https://doi.org/10.1016/S0140-6736(16)31592-6)
- Ducharme, K. A., & Brawley, L. R. (1995). Predicting the intentions and behavior of exercise initiates using two forms of self-efficacy. *Journal of Behavioral Medicine, 18*(5), 479-497. <https://doi.org/10.1007/BF01904775>
- Everett, B., Salamonson, Y., & Davidson, P. M. (2009). Bandura's exercise self-efficacy scale: Validation in an Australian cardiac rehabilitation setting. *International Journal of Nursing Studies, 46*(6), 824-829. [https://doi.org/10.1016/S0140-6736\(16\)31592-6](https://doi.org/10.1016/S0140-6736(16)31592-6)
- Feltz, D. L., & Chase, M. A. (1998). The measurement of self-efficacy and confidence in sport. In J. L. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 63- 78). Morgantown, WV: Fitness Information Technology.
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport: Research and strategies for working with athletes, teams, and coaches*. Champaign, IL: Human Kinetics.
- Fjeldsoe, B. S., Miller, Y. D., Prosser, S. J., & Marshall, A. L. (2020). How does MobileMums work? Mediators of a physical activity intervention. *Psychology & health, 35*(8), 968-983. <https://doi.org/10.1080/08870446.2019.1687698>
- Garcia, A. W., & King, A. C. (1991). Predicting long-term adherence to aerobic exercise: A comparison of two models. *Journal of Sport & Exercise Psychology, 13*(4), 394-410. <https://doi.org/10.1123/jsep.13.4.394>
- Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review, 17*(2), 183-211. <https://doi.org/10.5465/amr.1992.4279530>
- Haas, B. K., & Northam, S. (2010). Measuring self-efficacy: development of the Physical Activity Assessment Inventory. *Southern Online Journal of Nursing Research, 10*(4), 35-51.
- Jackson, B., Beauchamp, M. R., & Dimmock, J. A. (2020). Efficacy beliefs in physical activity settings. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (pp. 57-80). John Wiley & Sons. <https://doi.org/10.1002/9781119568124.ch4>



- Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., . . . Lancet Physical Activity Series Working Group. (2012). The pandemic of physical inactivity: Global action for public health. *The Lancet (British Edition)*, *380*(9838), 294-305. [https://doi.org/10.1016/S0140-6736\(12\)60898-8](https://doi.org/10.1016/S0140-6736(12)60898-8)
- Lipschitz, J. M., Yusufov, M., Paiva, A., Redding, C. A., Rossi, J. S., Johnson, S., Blissmer, B., Simay Gokbayrak, N., Velicer, W.F., & Prochaska, J. O. (2015). Transtheoretical principles and processes for adopting physical activity: A longitudinal 24-month comparison of maintainers, relapsers, and nonchangers. *Journal of Sport & Exercise Psychology*, *37*(6), 592-606. <https://doi.org/10.1123/jsep.2014-0329>
- Mailey, E. L., Wójcicki, T. R., Motl, R. W., Hu, L., Strauser, D. R., Collins, K. D., & McAuley, E. (2010). Internet-delivered physical activity intervention for college students with mental health disorders: a randomized pilot trial. *Psychology, health & medicine*, *15*(6), 646-659. <https://doi.org/10.1080/13548506.2010.498894>
- Marcus, B. H., Selby, V. C., Niaura, R. S., & Rossi, J. S. (1992). Self-efficacy and the stages of exercise behavior change. *Research Quarterly for Exercise and Sport*, *63*(1), 60-66. <https://doi.org/10.1080/02701367.1992.10607557>
- Martin Ginis, K. A., Latimer, A. E., Arbour-Nicitopoulos, K. P., Bassett, R. L., Wolfe, D. L., & Hanna, S. E. (2011). Determinants of physical activity among people with spinal cord injury: A test of social cognitive theory. *Annals of Behavioral Medicine*, *42*(1), 127-133. <https://doi.org/10.1007/s12160-011-9278-9>
- Marsh, H. W. (1994). Sport motivation orientations: Beware of jingle-jangle fallacies. *Journal of Sport and Exercise Psychology*, *16*(4), 365-380. <https://doi.org/10.1123/jsep.16.4.365>
- McAuley, E. (1992). The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. *Journal of Behavioral Medicine*, *15*(1), 65-88. <https://psycnet.apa.org/doi/10.1007/BF00848378>
- McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of behavioral medicine*, *16*(1), 103-113. <https://doi.org/10.1007/BF00844757>
- Megakli, T., Vlachopoulos, S. P., Thøgersen-Ntoumani, C., & Theodorakis, Y. (2017). Impact of aerobic and resistance exercise combination on physical self-perceptions and self-esteem in women with obesity with one-year follow-up. *International Journal of Sport and Exercise Psychology*, *15*(3), 236-257. <https://doi.org/10.1080/1612197X.2015.1094115>
- Moher, D., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264. <https://doi.org/10.1371/journal.pmed.1000097>

- Mosher, C. E., Lipkus, I., Sloane, R., Snyder, D. C., Lobach, D. F., & Demark-Wahnefried, W. (2013). Long-term outcomes of the FRESH START trial: exploring the role of self-efficacy in cancer survivors' maintenance of dietary practices and physical activity. *Psycho-oncology*, 22(4), 876-885. <https://doi.org/10.1002/pon.3089>
- Murru, E. C., & Ginis, K. A. M. (2010). Imagining the possibilities: The effects of a possible selves intervention on self-regulatory efficacy and exercise behavior. *Journal of Sport and Exercise Psychology*, 32(4), 537-554. <https://doi.org/10.1123/jsep.32.4.537>
- Myers, N. D., & Feltz, D. L. (2007). From self-efficacy to collective efficacy in sport: Transitional methodological issues. In G. Tenenbaum & R.C. Eklund (Eds.), *The handbook of sport psychology* (pp. 799–819). Wiley.
- Myers, N. D., Feltz, D. L., & Wolfe, E. W. (2008). A confirmatory study of rating scale category effectiveness for the coaching efficacy scale. *Research Quarterly for Exercise and Sport*, 79(3), 300-311. <https://doi.org/10.1080/02701367.2008.10599493>
- Myers, N. D., McMahon, A., Prilleltensky, I., Lee, S., Dietz, S., Prilleltensky, O., ... & Brincks, A. M. (2020). Effectiveness of the fun for wellness web-based behavioral intervention to promote physical activity in adults with obesity (or overweight): Randomized controlled trial. *JMIR formative research*, 4(2), e15919. <https://doi.org/10.2196/15919>.
- Myers, N. D., Wolfe, E. W., & Feltz, D. L. (2005). An evaluation of the psychometric properties of the coaching efficacy scale for coaches from the united states of america. *Measurement in Physical Education and Exercise Science*, 9(3), 135-160. [https://doi.org/10.1207/s15327841mpee0903\\_1](https://doi.org/10.1207/s15327841mpee0903_1)
- Plotnikoff, R. C., Blanchard, C., Hotz, S. B., & Rhodes, R. (2001). Validation of the decisional balance scales in the exercise domain from the transtheoretical model: A longitudinal test. *Measurement in Physical Education and Exercise Science*, 5(4), 191-206. [https://doi.org/10.1207/S15327841MPEE0504\\_01](https://doi.org/10.1207/S15327841MPEE0504_01)
- Raykov T, & Marcoulides G. A. (2019) Thanks coefficient alpha, we still need you. *Educational and Psychological Measurement*, 79, 200-10. <https://doi.org/10.1177/0013164417725127>
- Resnick, B., & Jenkins, L. S. (2000). Testing the reliability and validity of the self-efficacy for exercise scale. *Nursing Research (New York)*, 49(3), 154-159. <https://doi.org/10.1097/00006199-200005000-00007>
- Rhodes, R. E., & Courneya, K. S. (2003). Self-efficacy, controllability and intention in the theory of planned behavior: Measurement redundancy or causal independence? *Psychology & Health*, 18(1), 79-91. <https://doi.org/10.1080/0887044031000080665>

- Rodgers, W. M., Wilson, P. M., Hall, C. R., Fraser, S. N., & Murray, T. C. (2008). Evidence for a multidimensional self-efficacy for exercise scale. *Research Quarterly for Exercise and Sport*, 79(2), 222-234. <https://doi.org/10.1080/02701367.2008.10599485>
- Rovniak, L. S., Anderson, E. S., Winett, R. A., & Stephens, R. S. (2002). Social cognitive determinants of physical activity in young adults: A prospective structural equation analysis. *Annals of Behavioral Medicine*, 24(2), 149-156. [https://doi.org/10.1207/S15324796ABM2402\\_12](https://doi.org/10.1207/S15324796ABM2402_12)
- Sallis, J. F., Bull, F., Guthold, R., Heath, G. W., Inoue, S., Kelly, P., . . . Lancet Physical Activity Series 2 Executive Committee. (2016). Progress in physical activity over the olympic quadrennium. *The Lancet (British Edition)*, 388(10051), 1325-1336. [https://doi.org/10.1016/S0140-6736\(16\)30581-5](https://doi.org/10.1016/S0140-6736(16)30581-5)
- Sallis, J. F., Pinski, R. B., Grossman, R. M., Patterson, T. L., & Nader, P. R. (1988). The development of self-efficacy scales for healthrelated diet and exercise behaviors. *Health Education Research*, 3(3), 283-292. <https://doi.org/10.1093/her/3.3.283>
- Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, Daniëlle A.W.M, Knol, D. L., Dekker, J., Bouter, L.M., de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
- United States Department of Health and Human Services: 2018 Physical activity guidelines advisory committee. (2018). *2018 Physical activity guidelines advisory committee scientific report*. <https://health.gov/paguidelines/second-edition/report/>
- Voskuil, V. R., Pierce, S. J., & Robbins, L. B. (2017). Comparing the psychometric properties of two physical activity self-efficacy instruments in urban, adolescent girls: Validity, measurement invariance, and reliability. *Frontiers in Psychology*, 8, 1301. <https://doi.org/10.3389/fpsyg.2017.01301>
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology*, 82(2), 247-252. <https://psycnet.apa.org/doi/10.1037/0021-9010.82.2.247>
- Williams, D. M., & Rhodes, R. E. (2016). The confounded self-efficacy construct: Conceptual analysis and recommendations for future research. *Health Psychology Review*, 10(2), 113-128. <https://doi.org/10.1080/17437199.2014.941998>
- Wooldridge, J. S., Ranby, K. W., Roberts, S., & Huebschmann, A. G. (2019). A Couples-Based Approach for Increasing Physical Activity Among Adults With Type 2 Diabetes: A Pilot Feasibility Randomized Controlled Trial. *The Diabetes Educator*, 45(6), 629-641. <https://doi.org/10.1177%2F0145721719881722>

Young, M. D., Plotnikoff, R. C., Collins, C. E., Callister, R., & Morgan, P. J. (2016). A test of social cognitive theory to explain men's physical activity during a gender-tailored weight loss program. *American journal of men's health*, 10(6), NP176-NP187. <https://doi.org/10.1177%2F1557988315600063>

Zhu, W., & Kang, S. J. (1998). Cross-cultural stability of the optimal categorization of a self-efficacy scale: A Rasch analysis. *Measurement in Physical Education and Exercise Science*, 2(4), 225-241. [https://doi.org/10.1207/s15327841mpee0204\\_3](https://doi.org/10.1207/s15327841mpee0204_3)

## CHAPTER III: STUDY 2

### ABSTRACT

The objective of this study was to improve the measurement of self-efficacy to regulate physical activity in adults with obesity. To accomplish this a latent variable approach was used to explore dimensionality, temporal invariance, and external validity of responses to a slightly modified version of the barriers self-efficacy scale: the self-efficacy to regulate physical activity scale (SERPA). Data ( $N_{baseline} = 461$  and  $N_{30\text{ days post-baseline}} = 427$ ) from the Well-Being and Physical Activity Study (ClinicalTrials.gov, identifier: NCT03194854), which deployed the Fun For Wellness (FFW) intervention, were analyzed. A two-dimensional factor structure explained responses to the SERPA at baseline. There was strong evidence for at least partial temporal measurement invariance for this two-dimensional structure. There was strong evidence for the effectiveness of the FFW intervention to exert a direct effect on the proposed two-dimensional structure of latent self-efficacy to regulate physical activity in adults with obesity at 30 days post-baseline.

### INTRODUCTION

Insufficient physical activity in adults has become a global pandemic (Kohl et al., 2012; Sallis et al., 2016) and physical activity may play an important role in weight regulation (Ilacqua et al., 2019). Obesity is characterized by increased weight and therefore an increased body mass index (BMI), which is associated with other noncommunicable diseases such as cardiovascular disease (United States Department of Health and Human Services [USDHHS], 2013). Regular physical activity is positively associated with weight loss and is an essential component of the prevention of obesity-related pathology (Donnelly et al., 2009). Most adults with obesity however, do not meet public health guidelines for physical activity (Tran et al., 2020; Tudor-

Locke et al., 2010). Consequently, motivational constructs have been applied in developing effective health-behavior change (e.g., physical activity promoting) programs to address this global problem of adult obesity (Curry et al., 2018; Gourlan et al., 2011). Self-efficacy theory (Bandura, 1997) was utilized in the Well-Being and Physical Activity (WBPA) study in which researchers investigated the effectiveness of the Fun For Wellness (FFW) intervention (e.g., Myers et al., 2020) to promote physical activity in adults with obesity. The two other major objectives of the WBPA study were to promote subjective well-being and well-being actions; the effectiveness reports of these outcomes are presented in (Myers et al., 2020) and (Lee et al., 2021) respectively.

### **The Well-Being and Physical Activity (WBPA) Study**

The WBPA study targeted two forms of self-efficacy as potentially modifiable mediators of physical activity: (a) self-efficacy for varying levels of physical activity measured by the physical activity self-efficacy (PASE) scales and (b) self-efficacy to regulate physical activity measured by the self-efficacy to regulate physical activity scale (SERPA). Both forms of self-efficacy were measured in the WBPA study since they are theoretically distinct and may each be influenced differently by the FFW intervention. The FFW conceptual model for the promotion of physical activity is based on participants' growth in various forms of targeted self-efficacy judgments, such as self-efficacy to regulate physical activity in response to the intervention components.

#### *The Intervention*

The FFW intervention focused on presenting sources of self-efficacy information (e.g., verbal persuasion; Bandura, 1997) in the form of online, interactive challenges (e.g., vignettes performed by professional actors) to provide participants with capability-enhancing learning

opportunities. The theory-based online capability-enhancing components comprising the FFW intervention are described in detail and justified in Myers et al. (2019). The interactive and scenario-based challenges are organized by the BET I CAN acronym (Myers et al., 2017), where each letter represents a type of challenge. For example, “B” refers to *behaviors* (e.g., setting a goal). Participants in the FFW intervention group were given access to 152 BET I CAN challenges, which gave them the opportunity to be exposed to sources of self-efficacy information (e.g., enactive mastery experiences; Bandura, 1997). Engagement with the intervention was expected to build personal capability and directly promote improvement in the targeted performance domain, which in this case was physical activity participation according to specified guidelines (World Health Organization [WHO], 2018; USDHHS, 2013). The WHO (2018) and the USDHHS (2013) both recommend that individuals engage in at least 150 min of physical activity at a moderate intensity, or 75 min of physical activity at a vigorous intensity weekly.

#### *Self-efficacy Theory as the Basis for the Intervention and Measurement*

Self-efficacy is a component of the broader social cognitive theory (Bandura, 2001). Self-efficacy judgments are theorized to influence an individual’s choice regarding their course of action, the effort they exert and how long they will persevere in the face of barriers and adversity (Bandura, 1997). Self-efficacy theory is a popular contemporary theory of motivation, and self-efficacy perceptions occupy a major role in many health-behavior change models (Jackson et al., 2020). The beliefs that an individual holds about their perceived capabilities (i.e., their self-efficacy beliefs) to be successful in a specific performance domain can play an important motivational role in changing or promoting specific behaviors (Gilson & Feltz, 2012). Self-efficacy beliefs are therefore expected to play a central role in motivating behavior change, such

as the initiation and maintenance of regular physical activity. Consequently, valid, and reliable measurement of self-efficacy beliefs including exploring empirical evidence for dimensionality is important in justifying the utility of self-efficacy measurement (Bandura, 1997, 2006; Feltz & Chase, 1998; Feltz et al., 2008; Myers & Feltz, 2007).

Physical activity self-efficacy generally describes the degree to which an individual believes that they have the capability to engage in varying levels (e.g., frequency, intensity, or duration) of physical activity behavior. Successful engagement in sport and exercise however, also relies on the accomplishment of specific tasks, which are in part dependent on one's ability to regulate and manage their behavior over time (Anderson et al., 2007). An individual's confidence regarding these self-regulatory processes, that is, their self-regulatory efficacy is prominent in research on physical activity participation (Jackson et al., 2020). Self-efficacy to regulate physical activity therefore, more specifically refers to an individual's beliefs in their capability to overcome possible barriers (e.g., bad weather) to physical activity. Self-efficacy to regulate physical activity furthermore appears to be theoretically important in playing a motivating role in successful engagement in goal-directed physical activity (Jackson et al., 2020). Multiple studies have reported a positive association between individuals' confidence in their regulatory capacities relating to physical activity and physical activity participation outcomes. For example, Spink and Nickel (2010) found evidence for self-regulatory efficacy as a partial mediator between attributions and intention for health-related physical activity (i.e., moderate, and mild exercise levels). Additionally, Anderson-Bill et al. (2011) reported the results of an intervention where self-efficacy to face barriers to physical activity as measured by a 23-item scale (Anderson et al., 2007) was positively associated with increased physical activity.



### *Measurement of Self-efficacy to Regulate Physical Activity*

The SERPA scale, a modified version of the barriers self-efficacy scale (BARSE; McAuley, 1992) was developed to fit the FFW context (Myers et al., 2020). The SERPA was developed to be congruent with recommendations for the measurement of self-efficacy judgments (e.g., Bandura, 1997, 2006), particularly as it relates to the human performance domain (e.g., Bateman et al., 2021; Feltz & Chase, 1998; Feltz et al., 2008). The SERPA reflects a conceptual analysis of the skills or capabilities required for successful physical activity performance in a specified domain. In the case of the SERPA, the skills or capabilities involve overcoming stipulated barriers to physical activity. Additionally, the SERPA has the following properties: 1) it measures current capability to overcome barriers to physical activity for health, 2) it measures strength of self-efficacy beliefs, 3) it is concordant with the physical activity measure used, and 4) it has a response scale that is optimally categorized as suggested by Myers et al. (2008). The conceptual and measurement properties of the SERPA coincide the with the recommendations made in Bateman et al. (2021) for the measurement of self-efficacy beliefs associated with physical activity performance in physical activity interventions. Consequently, it was expected in the WBPA study that higher levels of self-efficacy to regulate physical activity, as measured by the SERPA, would be associated with higher levels of measured physical activity.

### *Results Under a Traditional Observed Score Approach*

Self-efficacy scale item responses are typically summed and converted to a unidimensional self-efficacy score under a traditional observed score approach (Bandura, 1997, 2006; Feltz & Chase, 1988; Myers & Feltz, 2007). The WBPA study (Myers et al., 2020) reported the measurement of self-efficacy to regulate physical activity by the SERPA under a

traditional observed score approach. There was evidence for the FFW intervention having a positive direct effect on self-efficacy to regulate physical activity ( $\beta = .16, p = .01, d = .25$ ) measured at T2. There was also evidence of scale reliability of the SERPA ( $\alpha = .90$ ), under the assumption that the construct is unidimensional.

There are some disadvantages of using a traditional observed score approach. First, an observed score approach does not account for measurement error. Second, an observed score approach often assumes temporal measurement invariance of the construct (i.e., longitudinal measurement invariance). Third, a traditional observed score approach does not fully account for the dimensionality of the construct because unidimensionality is typically only assumed and not explicitly tested. Bandura (2006) however, recommends testing the factor structure (i.e., dimensionality) of self-efficacy scales instead of assuming a unidimensional structure. Failure to account for dimensionality may have implications for the reliability and validity of the scores produced by a self-efficacy scale (Feltz et al., 2008). The limitations of a traditional observed score analysis have substantial conceptual implications because the validity of self-efficacy measures is largely dependent on the extent to which the content of the scale represents the construct and how well the construct is aligned with self-efficacy theory (e.g., growth in strength of capability beliefs in response to exposure to sources of self-efficacy information; Bandura, 1997). Myers et al. (2021) used a latent variable approach in response to the recommendations from Bandura (2006), and the limitations of not accounting for dimensionality (e.g., Feltz et al., 2008) to explore validity evidence for the PASE scales used in the WBPA study (Myers et al., 2020). Evidence was found for a two-dimensional (as compared to an assumed unidimensional) factor structure of the PASE scales (Myers et al., 2021).

According to the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME; 2014), validity evidence should be provided when a measurement instrument is developed or revised. Furthermore, Millsap (2012) recommends that evidence should be provided for temporal measurement invariance to justify that observed changes in the measured scores of a construct are not due to a change in the construct (i.e., the construct remains valid) itself across measurement occasions. The current study is necessary because the factor structure, longitudinal measurement invariance, and external validity of the SERPA has yet to be determined within a latent variable framework.

### **Objective and Exploratory Research Questions**

The purpose of the current study was to explore validity evidence for the scores produced by the SERPA as used in the WBPA study. The specific objectives of the study were: 1) to explore a measurement model that fit the data collected from the WBPA study, 2) to explore evidence for temporal measurement invariance, and 3) to test the external validity of the SERPA according to self-efficacy theory within a latent variable framework. The extent to which the FFW intervention is effective at promoting latent self-efficacy to regulate physical activity demonstrates the external validity of the measures produced by the SERPA. Three exploratory research questions were posed. First, how many factors were empirically justified to explain responses to the SERPA? Second, was there evidence of temporal measurement invariance of the SERPA? Third, was there evidence for the effectiveness of the FFW intervention to promote latent self-efficacy to regulate physical activity?

To investigate these research questions a multiple study design approach was used. This approach was chosen because results to an earlier research question (e.g., examining

dimensionality, Research question 1) had implications for the investigation of later research questions (e.g., examining temporal invariance, Research question 2). A general method section is first provided followed by study-specific method, results, and conclusion sections prior to a brief general discussion section as is consistent with a multiple study design approach.

## **GENERAL METHODS**

### **Study Design**

The data described and analyzed in this paper were collected as part of the WBPA study (ClinicalTrials.gov, identifier: NCT03194854). Consequently, some of the text used in this manuscript is similar to that used to describe the methods in Myers et al. (2020). This approach of using similar text was taken so that readers would not have to consult a previously published work in order to fully understand the methods used in the WBPA study that are important in understanding the present study (American Psychological Association, 2020). In summary, the WBPA study was an online, large-scale, prospective, double-blind, parallel-group, randomized controlled trial. Data collection was conducted from August 2018 through November 2018. Data collection occurred at three measurement occasions (i.e., baseline [T1], 30 days [T2] and 60 days [T3] after baseline). Detailed descriptions of the methodology of the WBPA study are described in the study protocol, Myers et al. (2019), and in Myers et al. (2020). Myers et al. (2020) also includes a populated Consolidated Standards of Reporting Trials-EHEALTH checklist and information regarding compliance with ethical standards in the WBPA study.

Myers et al. (2020) used a traditional observed score approach to analyze the SERPA data. The current study however, uses a latent variable, structural equation modeling framework to reanalyze the data in examining the validity and reliability of the responses to the SERPA items. The present study is based on the FFW conceptual model for the promotion of physical

activity (Myers et al., 2020). This paper emphasizes the part of the conceptual model which proposes that the FFW intervention will have a positive, direct effect on self-efficacy to regulate physical activity at T2 (Myers et al., 2020).

### **Participants**

Participants were recruited through a panel recruitment company. To be eligible for the study, participants had to live in the USA, be able to access the online intervention, be within the age range of 18 to 64 years, and be overweight or obese (i.e., have a BMI greater than 25.00 kg/m<sup>2</sup>). Participants could not be simultaneously enrolled in another similar (i.e., promoting physical activity or wellbeing) intervention program.

### **Procedures**

After giving informed consent, participants were randomly assigned to the intervention (FFW) or control (usual care; UC) group. Randomization was accomplished using software code that was written to accomplish equal allocations to the intervention and control groups. Each participant in each group was given a unique login credential to access the secure FFW website. Participants assigned to the FFW group were given immediate access to the online intervention program. Participants in the FFW group were asked to engage with the FFW intervention. FFW participants had the opportunity to earn a total of \$45 worth of Amazon electronic gift cards based on their engagement with the intervention components (e.g., \$10 for completing both the survey battery at 30 days post-baseline and at least 15 BET I CAN post-introductory challenges). The intervention consists of BET I CAN challenges (e.g., “making a plan”) which provide capability-enhancing opportunities (e.g., “next steps”) to the participants (Myers et al., 2020). The FFW engagement scoring system is explained in detail in Myers et al. (2019). In short, participants were required to earn at least 21 participation points (i.e., high impact for completing

a non-introductory BET I CAN challenge) to be classified as engaged with the FFW intervention. Participants in the UC group were put on a waitlist to access the intervention and asked to conduct their lives as usual. Both groups were given access to the survey battery which was completed at the appropriate measurement occasions (i.e., from T1 to T3). UC participants had the opportunity to earn up to \$30 worth of Amazon electronic gift cards. Details of the remuneration plan for both the FFW and UC groups is presented in Myers et al. (2019).

## **Measures**

The present study focused on analyzing participant data from the SERPA (see Appendix C). The complete survey battery consisted of self-report instruments designed to measure demographic information, physical activity, and various forms of self-efficacy. Self-report demographic data were collected at baseline and included participant gender, race/ethnicity, highest level of education completed, marital status, employment status, age, and household annual income. These demographic variables (i.e., demographic covariates) have been proposed as covariates of physical activity (Bauman et al., 2012; Rubenstein et al., 2016). A detailed description of all the measures used in the study is given in Myers et al. (2019). The present study analyzed participant data from the SERPA. The fitted statistical model exploring external validity also included group allocation (i.e., UC or FFW) along with the demographic covariates.

### *Development of the SERPA*

The SERPA was developed prior to the WBPA study by making modifications to the BARSE scale. The SERPA was developed to measure self-efficacy to regulate engagement in the recommended amount of weekly physical activity for health in adults with obesity. The BARSE scale however, was developed to measure an individual's perceived capabilities to exercise three times per week for 40 min over the next 2 months in the face of commonly identified barriers to

exercise participation. The BARSE scale, which was not used in this study, consists of 13 items on an 11-point Likert scale. Each item consists of a potential barrier to exercise (e.g., “it was not fun or enjoyable”). These barriers were determined through an attributional analysis of reasons why individuals drop out of exercise (McAuley et al., 1990 as cited in McAuley 1992). For each item of the BARSE, participants indicate their confidence to execute exercise behavior on a 10-point scale comprised of 10-point increments, ranging from 0% (not at all confident) to 100% (highly confident). The total measure of self-efficacy for each individual is calculated by finding the sum of the confidence ratings and dividing it by the total number of items in the scale (i.e., a traditional observed score approach).

The SERPA consists of 13 items on a 5-point Likert scale. Each item consists of a potential barrier to physical activity (e.g., “the weather is very bad”). For each item of the SERPA participants indicate their confidence to engage in a recommended amount of weekly physical activity on 5-point increments, ranging from 0 (no confidence) to 4 (complete confidence). Response points 1, 2 and 3 represent low, moderate, and high confidence, respectively. A 5-point rating scale was used even though Bandura (1997) recommended a minimum of 10 response categories because Myers et al. (2008) has subsequently found evidence for improved validity of the scores produced by measures of self-efficacy consisting of shorter rating scales (i.e., five or four categories). The total measure of self-efficacy for each individual is calculated by finding the sum of the confidence ratings and dividing it by the total number of items in the scale as explained in Myers et al. (2020; i.e., a traditional observed score approach).

The SERPA was operationally defined as a self-regulatory efficacy construct according to Bandura (1997). As a result, each item refers to a specific situation that may interfere with an

individual's physical activity effort. Each item comprising the SERPA corresponds to the same 13 barriers in the same order as the items comprising the BARSE, however the SERPA items are framed within the context of physical activity participation and not exercise exclusively. The instructions for the SERPA emphasize an evaluation of current (not future) capability, and strength of self-efficacy beliefs (i.e., how confident) to overcome possible barriers to engagement in a recommended amount of weekly physical activity for health. Physical activity for health is defined in the scale instruction in terms of frequency (i.e., weekly), intensity (i.e., moderate, or vigorous) and duration (i.e., minutes). Moderate and vigorous intensity physical activity are also defined in terms of physical effort and associated breathing changes. Common examples of moderate and vigorous activities are also given (e.g., bicycling at a moderate intensity). The scale items of the SERPA were also designed to be concordant with the physical activity measure used in the study (i.e., the International Physical Activity Questionnaire; Ainsworth et al., 2000; Craig et al., 2003). in terms of recommended amount of weekly physical activity for health based on frequency, intensity, and duration.

### **Data Collection, Demographics, and Descriptive Statistics**

Myers et al. (2020) presented a figure depicting participant flow over the three measurement occasions for self-efficacy to regulate physical activity data. In summary, 461 ( $n_{FFW} = 219$ ,  $n_{UC} = 242$ ) and 424 participants ( $n_{FFW} = 195$ ,  $n_{UC} = 229$ ) provided SERPA data at baseline and T2, respectively. For the present study, beyond attrition there were no missing data on variables used in the data analysis for Research questions 1 and 2. For Research question 3, four cases were missing data on one or more demographic variables. Most participants identified as female (65.4%), White, non-Hispanic (74.2%), having completed at least a 4-year college degree (42.5%), married (68.2%), and as a full-time employee (67.2%). The average age of the



participants was 41.8 years, and the average annual income was approximately \$77,500. There were no statistically significant differences in the observed proportions of demographic characteristics or the mean self-efficacy to regulate physical activity scores at baseline by randomization group. Most (91.7%) of the participants who were assigned to the FFW group were classified as engaged with the FFW intervention. See Scarpa et al. (2021) for full engagement results.

### **Data Analysis**

Statistical models were fit in Mplus 8 with maximum-likelihood (ML; Muthén & Muthén, 2017) estimation with robust standard errors ( $\Delta\chi^2_R$ ). Type I error was set as equal to .05. Missing data were handled with full information ML estimation using the observed information matrix, the default approach in Mplus, under the assumption of missing at random (Little & Rubin, 1987; Schafer & Graham, 2002). Finney and DiStefano (2006) additionally support using ML estimation for data consisting of five or more response categories (i.e., approximately continuous) that is approximately normally distributed (i.e., univariate skewness  $< 2$ , univariate kurtosis  $< 7$ ) – even though technically discrete. The distribution of 5-point Likert responses to the SERPA items were therefore treated as normally distributed and continuous at T1 (see Table 5) and T2 (see Table 6). Exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) provided the data analytic approach that was used to explore factor structure, and invariance testing. Geomin was used as the oblique rotation criteria because it appears to perform well when there is little knowledge about the true loading structure (e.g., Myers et al., 2015). In deciding which models to accept, the conceptual interpretability of the estimated rotated pattern matrix was also considered. The indices of model-data fit used were  $\chi^2$ , RMSEA, CFI, TLI and SRMR. Model-data fit heuristic classifications were consistent with Hu and Bentler (1999).

Coefficient  $H$  was the measure of construct reliability by using structure coefficients (Graham et al., 2003) and  $H \geq .80$  was considered desirable (Hancock & Mueller, 2001).

Table 5. *Distribution of Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) items at Time 1*

| Time 1 |                            |      |      |      |      |      |       |       |       |
|--------|----------------------------|------|------|------|------|------|-------|-------|-------|
| Item   | Percent Observed Responses |      |      |      |      | M    | SD    | Skew  | Kurt  |
|        | 0                          | 1    | 2    | 3    | 4    |      |       |       |       |
| 1      | 10.2                       | 25.6 | 32.4 | 18.1 | 13.6 | 1.86 | 1.092 | .132  | -.565 |
| 2      | 7.8                        | 20.1 | 38.7 | 19.6 | 13.8 | 2.13 | 1.070 | .077  | -.538 |
| 3      | 11.7                       | 21.8 | 32.6 | 20.6 | 13.4 | 1.99 | 1.191 | -.041 | -.797 |
| 4      | 9.2                        | 25.8 | 32.5 | 20.2 | 12.2 | 2.08 | 1.134 | -.007 | -.710 |
| 5      | 13.9                       | 29.5 | 29.9 | 17.4 | 9.4  | 1.90 | 1.222 | .136  | -.948 |
| 6      | 5.3                        | 17.3 | 27.9 | 26.7 | 22.8 | 2.47 | 1.110 | -.260 | -.647 |
| 7      | 8.5                        | 24.6 | 35.5 | 21.0 | 10.4 | 2.09 | 1.110 | .017  | -.621 |
| 8      | 10.1                       | 26.0 | 31.7 | 19.8 | 12.4 | 2.02 | 1.114 | .039  | -.649 |
| 9      | 8.2                        | 25.4 | 36.7 | 19.8 | 10.0 | 2.02 | 1.125 | .111  | -.656 |
| 10     | 14.4                       | 22.8 | 36.5 | 18.3 | 8.0  | 1.95 | 1.110 | .141  | -.624 |
| 11     | 10.1                       | 19.2 | 29.1 | 23.6 | 18.0 | 2.23 | 1.116 | -.083 | -.717 |
| 12     | 7.2                        | 16.4 | 34.3 | 23.3 | 18.7 | 2.30 | 1.109 | -.097 | -.693 |
| 13     | 8.0                        | 18.2 | 36.0 | 24.5 | 13.4 | 2.20 | 1.096 | -.160 | -.634 |

Note. M = Mean SD = Standard deviation; Skew = Univariate Skewness; Kurt = Univariate Kurtosis; 0 = No Confidence; 1 = Low Confidence; 2 = Moderate Confidence; 3 = High Confidence; 4 = Complete Confidence; N ranges from 661 to 550.

Table 6. *Distribution of Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) items at Time 2*

| Time 2 |                            |      |      |      |      |      |       |       |       |
|--------|----------------------------|------|------|------|------|------|-------|-------|-------|
| Item   | Percent Observed Responses |      |      |      |      | M    | SD    | Skew  | Kurt  |
|        | 0                          | 1    | 2    | 3    | 4    |      |       |       |       |
| 1      | 11.0                       | 18.2 | 36.0 | 24.5 | 13.4 | 1.99 | 1.180 | .138  | -.798 |
| 2      | 5.8                        | 21.5 | 39.4 | 20.8 | 12.6 | 2.12 | 1.118 | .030  | -.597 |
| 3      | 13.2                       | 19.7 | 33.3 | 22.1 | 11.8 | 2.02 | 1.195 | .018  | -.819 |
| 4      | 8.7                        | 22.1 | 34.1 | 22.8 | 12.3 | 2.00 | 1.149 | .111  | -.760 |
| 5      | 13.4                       | 27.7 | 26.0 | 21.0 | 11.9 | 1.79 | 1.165 | .245  | -.733 |
| 6      | 4.5                        | 14.2 | 32.7 | 27.4 | 21.2 | 2.44 | 1.170 | -.258 | -.849 |
| 7      | 7.8                        | 21.5 | 36.7 | 21.6 | 12.4 | 2.00 | 1.102 | .083  | -.627 |
| 8      | 9.0                        | 23.3 | 35.3 | 21.7 | 10.7 | 1.98 | 1.166 | .118  | -.790 |

Table 6 (cont'd)

|    |     |      |      |      |      |      |       |       |       |
|----|-----|------|------|------|------|------|-------|-------|-------|
| 9  | 8.5 | 24.3 | 35.8 | 19.2 | 12.2 | 1.98 | 1.085 | .124  | -.566 |
| 10 | 9.3 | 26.2 | 35.1 | 19.3 | 10.2 | 1.83 | 1.131 | .078  | -.641 |
| 11 | 6.2 | 20.2 | 33.3 | 25.6 | 14.7 | 2.20 | 1.230 | -.133 | -.916 |
| 12 | 5.2 | 18.1 | 34.9 | 24.8 | 17.0 | 2.30 | 1.161 | -.157 | -.725 |
| 13 | 6.9 | 19.3 | 33.0 | 28.7 | 12.1 | 2.17 | 1.119 | -.109 | -.614 |

*Note.* M = Mean SD = Standard deviation; Skew = Univariate Skewness; Kurt = Univariate Kurtosis; 0 = No Confidence; 1 = Low Confidence; 2 = Moderate Confidence; 3 = High Confidence; 4 = Complete Confidence; N ranges from 661 to 550.

## STUDY 2A METHODS

ESEM was used to explore the factor structure of the baseline SERPA responses. Models with an increasing number of factors ( $m = 1, 2$ , etc.) were fit to the data. Nested model comparisons for better fitting models were evaluated using the difference in chi-square (robust) statistic ( $\Delta\chi_R^2$ ), where a significant test indicates that the simpler model fits significantly worse than the more complex model. Heuristic guidelines were used to judge change in model-data fit for nested model comparisons (e.g.,  $CFI_{\text{simple}} - CFI_{\text{complex}} = \Delta CFI$ ). The following were interpreted as evidence in favor of the more complex model,  $\Delta CFI \leq -.01$ ,  $\Delta TLI \leq .00$ , and  $\Delta RMSEA \geq .015$  (Marsh et al., 2010).

## STUDY 2A RESULTS

Examination of the model-data fit indices (see Table 7) for Model 1, the one-factor solution (i.e.,  $m = 1$ ) indicated that the null hypotheses for exact fit was rejected. The model-data fit indices for Model 1 ( $\chi^2(65) = 150.95, p < .001$ , RMSEA [CI<sub>90%</sub>] = .054 [.042, .065], CFI = .943, TLI = .931, SRMR = .041) indicated that the null hypothesis for exact fit was rejected but there was some evidence for close fit. The model-data fit indices for Model 2 ( $\chi^2(53) = 77.90, p < .05$ , RMSEA [CI<sub>90%</sub>] = .032 [.015, .046], CFI = .983, TLI = .976, SRMR = .027) indicated that the null hypothesis for exact fit was rejected but there was evidence for close fit. The model-data

fit indices for Model 3 ( $\chi^2(42) = 53.80, p > .05$ , RMSEA [CI<sub>90%</sub>] = .025 [.000, .042], CFI = .992, TLI = .985, SRMR = .021; see Table 7), indicated that the null hypothesis for exact fit was accepted and there was evidence for close fit. Observation of model-data fit indices suggested the two and three-factor solutions had better model-data fit when compared to the one-factor solution.

Table 7. Number of Factors Warranted to Explain Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 1 (N = 461)

| Model               | Goodness of fit |                            |      |      |      | Nested Model comparison |                           |
|---------------------|-----------------|----------------------------|------|------|------|-------------------------|---------------------------|
|                     | $\chi^2(df)$    | RMSEA [CI <sub>90%</sub> ] | CFI  | TLI  | SRMR | Model Compared          | $\Delta\chi^2(\Delta df)$ |
| M1:<br><i>m</i> = 1 | 150.95(65)***   | .054 [.042, .065]          | .943 | .931 | .041 | -                       | -                         |
| M2:<br><i>m</i> = 2 | 77.90(53)*      | .032 [.015, .046]          | .983 | .976 | .027 | M1 v M2                 | 65.28(12)***              |
| M3:<br><i>m</i> = 3 | 53.80(42)       | .025 [.000, .042]          | .992 | .985 | .021 | M2 v M3                 | 22.55(11)*                |

Note. *M* = Model, *v* = versus, \**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

The one-factor solution fit the data significantly worse ( $\Delta\chi^2(12) 65.28, p < .001$ ) than the two-factor solution (see Table 7). The latent construct reliability estimate of the one-factor solution was *H* = .89. The two-factor solution also fit the data significantly worse ( $\Delta\chi^2(11) 22.55, p < .05$ ) than the three-factor solution. The latent construct reliability estimates of the two-factor solution were *H* = .81, and *H* = .88. The inter-factor correlation,  $\psi$  for the two-factor solution was .58. The latent construct reliability estimates of the three-factor solution were *H* = .88, *H* = .68, and *H* = .84, respectively. For the three-factor solution one latent factor had a construct reliability estimate considerably below the point of desirability. The inter-factor correlations between the three factors comprising the three-factor solution were  $\psi = -.51$ , (between factors one and two),  $\psi = .67$  (between factors one and three) and  $\psi = -.40$  (between factors two and three) respectively. Though Model 3 had a better comparative fit than Model 2, evidence from

the reliability estimates and inter-factor correlations suggested that the three-factor model may not be appropriate for the data. Additionally, inspection of the item loadings for the three-factor model showed that only one item loaded significantly on the third factor. The three-factor model was consequently rejected. The majority of the evidence therefore indicated that a two-factor solution is empirically preferable to both the one and three-factor solutions.

### **Interpreting the Two-Factor Solution**

The geomin-rotated pattern coefficients, the inter-factor correlation, and coefficient  $H$  values for the accepted two-factor SERPA solution at T1 are presented in Table 8. Items 6, 11 and 12 had significant ( $p < .05$ ) moderate to large positive loadings on Factor 1. The other items had significant ( $p < .01$ ), larger positive loadings on Factor 2 (than Factor 1) which ranged from .39 (item 11) to .91 (item 5). The correlation between Factor 1 and Factor 2 ( $\psi = .58$ ) was significant ( $p < .001$ ), moderately large and positive. The construct reliability estimate, coefficient  $H$  of both factors was above the point of desirability. Factor 1 was conceptualized as self-efficacy to regulate barriers to physical activity participation based on social considerations. Factor 2 was conceptualized as self-efficacy to regulate internally perceived barriers to physical activity participation. The magnitude of Item 11 loadings on Factor 1 and Factor 2 were fairly close; they were .31 ( $p < .05$ ) and .39 ( $p < .01$ ), respectively. From a conceptual perspective however, Item 11 appears to refer to self-efficacy to regulate externally determined barriers to physical activity participation and was therefore interpreted as comprising Factor 1 and not Factor 2.

Table 8. *The Accepted Geomin-Rotated ( $\epsilon = .1$ ) Pattern Coefficients ( $\Lambda^*$ ), Inter-Factor Correlation ( $\psi$ ), and Coefficient H, for the Self-efficacy to Regulate Physical Activity Scale (SERPA) Factors at Time 1 ( $N = 461$ )*

| Item | $\Lambda^*$ |          | $\psi$ | $H$      |          |
|------|-------------|----------|--------|----------|----------|
|      | Factor 1    | Factor 2 |        | Factor 1 | Factor 2 |
| 1    | .12         | .44***   |        |          |          |
| 2    | .28*        | .48***   |        |          |          |
| 3    | .15         | .42***   |        |          |          |
| 4    | .06         | .60***   |        |          |          |
| 5    | -.34**      | .91***   |        |          |          |
| 6    | .76***      | .01      |        |          |          |
| 7    | .00         | .67***   | .58*** | .81      | .88      |
| 8    | .05         | .61***   |        |          |          |
| 9    | -.06        | .72***   |        |          |          |
| 10   | .11         | .52***   |        |          |          |
| 11   | .31*        | .39**    |        |          |          |
| 12   | .37**       | .32**    |        |          |          |
| 13   | .21         | .46***   |        |          |          |

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## STUDY 2A CONCLUSION

In response to Research question 1, the ESEM factor structure deemed most interpretable was a two-dimensional model. There was therefore substantial evidence for the validity of the SERPA measurement model. This finding emphasizes the claim that assuming a one-dimensional structure under a traditional observed score approach may compromise both reliability and validity as suggested by Myers and Feltz (2007). A conceptual analysis of the items comprising the SERPA dimensions of latent self-efficacy to regulate physical activity suggests that Factor 1, comprising 3 items (Items 6, 11 and 12), consists of barriers to physical activity concerned with the perceived judgment or support of external others (e.g., feeling self-conscious about appearance or receiving encouragement). Factor 2 consisting of the 10 other items, comprises barriers to physical activity directly related to the subjective evaluation by the individual in question (e.g., bad weather, or feeling physical discomfort). Factor 1 was therefore

conceptualized as self-efficacy to regulate barriers to physical activity participation based on social considerations. Factor 2 was therefore conceptualized as self-efficacy to regulate internally perceived barriers to physical activity participation.

## **STUDY 2B METHODS**

Four increasingly restricted models were fitted to determine longitudinal measurement invariance across T1 and T2 measurement points. The first model (Model 1) imposed the constraints required for identification (the baseline model). Model 2 added the constraint of invariant pattern coefficients (invariant  $\Lambda$ ) to Model 1. Model 3 added the constraint of invariant thresholds (invariant  $\Lambda$  and  $\tau$ ) to Model 2. Model 4 added the constraint of an invariant residual covariance matrix (invariant  $\Lambda$ ,  $\tau$ , and  $\Theta$ ) to Model 3. The fit of the increasingly restricted models was compared using the  $\Delta\chi^2_R$  statistic. In specific cases of possible misfit, the heuristic guidelines for a nested model comparison were used to provide evidence for the magnitude of the misfit. Heuristic guidelines were used to judge change in model-data fit for nested model comparisons (e.g.,  $CFI_{\text{simple}} - CFI_{\text{complex}} = \Delta CFI$ ). The following were interpreted as evidence in favor of the more complex model,  $\Delta CFI \leq -.01$ ,  $\Delta TLI \leq .00$ , and  $\Delta RMSEA \geq .015$  (Marsh et al., 2010). Modification indices were used post-hoc as necessary to determine possible locations of model misspecification (Saris et al., 2009). Given the design of the WBPA Study (i.e., intervention delivered from baseline until 30-days post-baseline), FFW assignment was specified as a covariate (i.e., exerting a direct effect on each of the two factors) in each model to account for any differences that may have been attributable to the study design.

## **STUDY 2B RESULTS**

Table 9 presents the results of measurement invariance analyses by time for responses to the SERPA at T1 and T2. The baseline model (Model 1) showed evidence for close model-data

fit ( $\chi^2(130) = 194.74, p < .001$ , RMSEA [CI<sub>90%</sub>] = .033 [.023, .043], CFI = .979, TLI = .971, SRMR = .029). Model 2 (invariant  $\Lambda$ ) showed evidence for good model-data fit ( $\chi^2(152) = 214.37, p < .001$ , RMSEA [CI<sub>90%</sub>] = .030 [.020, .039], CFI = .980, TLI = .976, SRMR = .033). Model 2 fit the data as well as Model 1 ( $\Delta\chi^2(22) = 20.23, p = .569$ ). Model 3 (invariant  $\Lambda$  and  $\tau$ ) showed evidence for good model-data fit ( $\chi^2(163) = 231.28, p < .001$ , RMSEA [CI<sub>90%</sub>] = .031 [.021, .039], CFI = .978, TLI = .976, SRMR = .036). Model 3 fit the data as well as Model 2 ( $\Delta\chi^2(11) = 17.17, p = .103$ ). Model 4 (invariant  $\Lambda, \tau$  and  $\Theta$ ) showed evidence for good model-data fit ( $\chi^2(176) = 259.02, p < .001$ , RMSEA [CI<sub>90%</sub>] = .033 [.024, .041], CFI = .974, TLI = .973, SRMR = .041). Model 4 however, fit the data significantly worse than Model 3 ( $\Delta\chi^2(13) = 28.43, p = .008$ ). The heuristic guidelines for the comparative fit of Models 3 and 4 however suggest that the magnitude of the comparative misfit may be small ( $\Delta\text{CFI} = -.004, \Delta\text{TLI} = -.003$ , and  $\Delta\text{RMSEA} = .002$ ). A post hoc inspection of the modification indices indicated that constraining the residual variance of item 1 to invariance by time may have been primarily responsible for the observed comparative misfit.

Model 4b (Table 9) therefore allowed the residual variance of Item 1 to vary by time. Model 4b showed evidence for good model-data fit ( $\chi^2(175) = 252.54, p < .001$ , RMSEA [CI<sub>90%</sub>] = .032 [.022, .040], CFI = .975, TLI = .974, SRMR = .041). Model 4b however fit the data significantly worse than Model 3 ( $\Delta\chi^2(12) = 21.53, p = .043$ ), though the misfit appeared to be marginal. The heuristic guidelines for the comparative fit of Models 3 and 4b ( $\Delta\text{RMSEA} = .001, \Delta\text{CFI} = -.003, \Delta\text{TLI} = -.002$ ) provide further evidence that the magnitude of the comparative misfit is likely to be very small. Model 4b was retained based on its good model-data fit, the marginal misfit of Model 4b compared to Model 3, and the presumed small initial misfit between Model 4 and Model 3. Model 4b imposed invariance for almost all (50 of 51) of the



measurement parameters (i.e., 25 of 25  $\lambda$ , 13 of 13  $\tau$ , and 12 of 13  $\theta$ ). There was consequently strong evidence for at least partial longitudinal measurement invariance of the SERPA responses.

Table 9. *Longitudinal Measurement Invariance for Responses to the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 1 (N = 461) and Time 2 (N = 424)*

| Model (M) | Model-data fit |                               |      |      |          | Nested model comparison |                           |
|-----------|----------------|-------------------------------|------|------|----------|-------------------------|---------------------------|
|           | $\chi^2(df)$   | RMSEA<br>[CI <sub>90%</sub> ] | CFI  | TLI  | SR<br>MR | Model compared          | $\Delta\chi^2(\Delta df)$ |
| M1        | 194.74(130)*** | .033<br>[.023, .043]          | .979 | .971 | .029     | -                       | -                         |
| M2        | 214.37(152)*** | .030<br>[.020, .039]          | .980 | .976 | .033     | M2 v M1                 | 20.23(22)                 |
| M3        | 231.28(163)*** | .031<br>[.021, .039]          | .978 | .976 | .036     | M3 v M2                 | 17.17(11)                 |
| M4        | 259.02(176)*** | .033<br>[.024, .041]          | .974 | .973 | .041     | M4 v M3                 | 28.43(13)**               |
| M4        | 252.54(175)*** | .032<br>[.022, .040]          | .975 | .974 | .041     | M4b v M3                | 21.53(12)*                |

Note. M1 = Baseline, M2 = Inv.  $\Lambda$ , M3 = Inv.  $\Lambda$  &  $\tau$ , M4 = Inv.  $\Lambda$ ,  $\tau$ , &  $\Theta$ , M4b = Inv.  $\Lambda$ ,  $\tau$ , &  $\Theta^a$ ,  $\Lambda$  = pattern coefficient matrix;  $\tau$  = thresholds vector;  $\Theta$  = residual covariance matrix.

<sup>a</sup>Freed residual variance for item 1 to vary by time.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## STUDY 2B CONCLUSION

In response to Research question 2, the responses to the SERPA showed evidence of at least partial longitudinal measurement invariance. There was only a modest degree of non-invariance observed. Though specific empirical implications of partial measurement invariance may be unclear (Millsap & Kwok, 2004), the location and degree of non-invariance is important. The constraints for 1 (i.e., 1  $\Theta$ ) parameter was freed and allowed to vary by time to allow for measurement non-invariance. Additionally, partial measurement invariance gives support for latent means comparisons, especially since structured means can be mathematically expected to be unaffected by non-invariance in  $\Theta$  (Byrne et al., 1989; Sörbom, 1974) which justifies investigating the FFW intervention exerting a positive direct effect on latent self-efficacy to regulate physical activity in Study 2c.

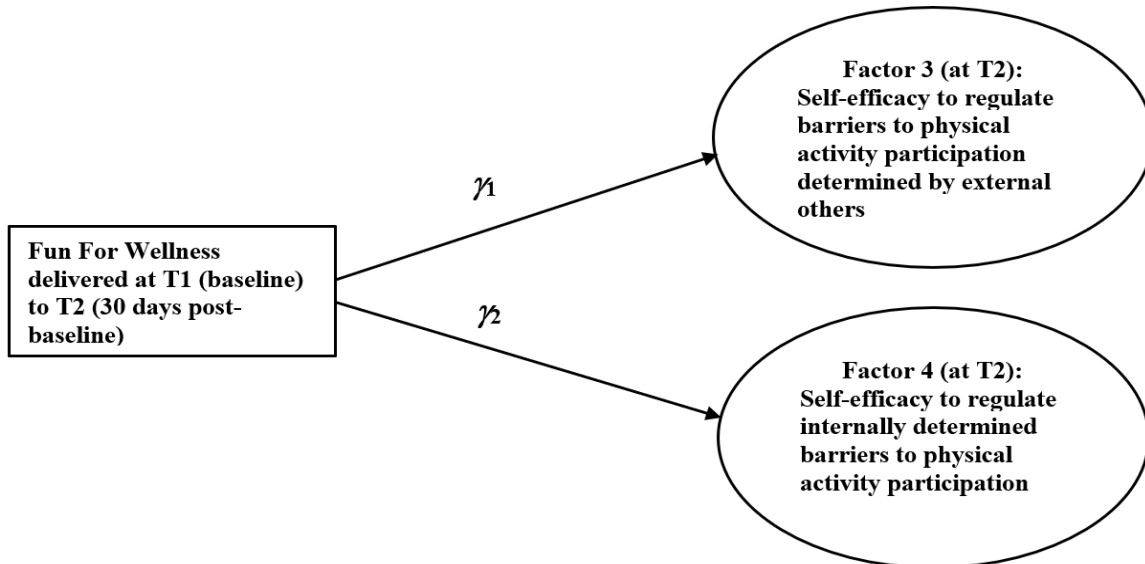
## STUDY 2C METHODS

An over-identified ( $df = 568$ ) path model was fit separately for the SERPA under an intent to treat approach (Hollis & Campbell, 1999). A total of 173 parameters (i.e., 24  $\tau$ , 50  $\lambda$ , 39  $\theta$ , 2 latent variable intercepts, and 58 direct effects) were estimated in the rotated solution for the model. SERPA items measured at T1 were specified as ESEM indicators of two continuous latent variables (i.e., Factor 1 and Factor 2) at T1. SERPA items measured at T2 were specified as ESEM indicators of two factors T2 (i.e., Factor 3 and Factor 4). Factor 1 and Factor 3 were labeled self-efficacy to regulate barriers to physical activity participation determined by external others. Factor 2 and Factor 4 were labeled self-efficacy to regulate internally determined barriers to physical activity participation. Factor 1 and Factor 2 were regressed on the demographic covariates. Factor 3 and Factor 4 were regressed on FFW (i.e., 0 = UC, 1 = FFW), Factor 1 and Factor 2, along with the demographic covariates. The residual for each SERPA item measured at T1 was free to covary with the corresponding residual at T2.

Two direct effects (see Figure 2) were focal parameters in each path model. The first focal parameter (i.e.,  $\gamma_1$ ) was the direct effect of FFW on Factor 3. The second focal parameter (i.e.,  $\gamma_2$ ) was the direct effect of FFW on Factor 4. The first focal parameter was interpreted as the adjusted (i.e., accounting for the covariates) latent mean difference on self-efficacy to regulate barriers to physical activity participation determined by external others at T2 for the FFW group as compared to the UC group. The second focal parameter was interpreted as the adjusted latent mean difference on self-efficacy to regulate internally determined barriers to physical activity participation at T2 for the FFW group as compared to the UC group. One-tailed hypothesis tests of statistical significance were used in the analysis of the focal parameters based on evidence for FFW to promote self-efficacy beliefs (Lee et al., 2020; Myers et al., 2017;

Myers, Prilleltensky, et al., 2020). Latent mean difference (Hancock, 2001), an analog to Cohen's  $d$  (Cohen, 1988), was used as an index of effect size for both focal parameters. The latent mean difference is denoted as  $d$  from this point forward. Commonly used heuristics were used to assist in the interpretation of an absolute value of Cohen's  $d$ : .20 (small effect), .50 (medium effect), and .80 (large effect). A focal parameter was considered meaningful if (a) it was statistically significant and (b) the magnitude of the effect was at least small in size.

Figure 2. *Focal Parameters (i.e.,  $\gamma_1$  and  $\gamma_2$ ) from the Path Model for the Self-Efficacy to Regulate Physical Activity Scale (SERPA) at T1 (Baseline) and T2 (30 Days Post-Baseline)*



## STUDY 2C RESULTS

Table 10 presents the results for the analysis of the direct effect of FFW on the proposed two-factor latent self-efficacy to regulate physical activity structure at T2 (see Appendix D for a full set of unstandardized parameter estimates). The model-data fit indices for the path model ( $\chi^2(568) = 681.55, p < .001, RMSEA [CI_{90\%}] = .021 [.014, .027], CFI = .971, TLI = .964, SRMR = .033$ ) indicated that the null hypothesis for exact fit was rejected but there was evidence for

close fit. The  $R^2$  estimates were 12.1%, 7.8%, 51.3%, and 51.2% for Factors 1 to 4, respectively. Only estimates for the focal parameters are discussed in this paper (the full set, that is, 171 of the unstandardized parameter estimates for the SERPA is available upon request). Some of the parameter estimates not discussed in this paper (e.g., direct effect of a demographic variable on self-efficacy to regulate physical activity) are similar to estimates reported in Table 2 of Myers et al. (2020).

Table 10. *Model-Data Fit, Percentage of Latent Variable Variance Accounted for ( $R^2$ ), and Unstandardized Direct Effects ( $\gamma$ ) in the Path Model for Factor 1 and Factor 2 of the Self-efficacy to Regulate Physical Activity Scale (SERPA) at Time 2 ( $N = 424$ ) Regressed on FFW*

| Scale      | $\chi^2(df)$   | Model-data fit        |      |      |      | $R^2$            |                  |
|------------|----------------|-----------------------|------|------|------|------------------|------------------|
|            |                | RMSEA [ $CI_{90\%}$ ] | CFI  | TLI  | SRMR | Factor 3         | Factor 4         |
| SERPA (T2) | 681.55(568)*** | .021 [.014, .027]     | .971 | .964 | .033 | 51.3%            | 51.2%            |
|            |                |                       |      |      |      | $\gamma$         |                  |
|            |                |                       |      |      |      | $d[CI_{95\%}]$   |                  |
|            |                |                       |      |      |      | .28 <sup>c</sup> | .24 <sup>d</sup> |
|            |                |                       |      |      |      | [.05 .50]        | [.04 .45]        |

*Note.* Factor 3 = self-efficacy to regulate barriers to physical activity participation determined by external others; Factor 4 = self-efficacy to regulate internally determined barriers to physical activity participation. <sup>c</sup> $p = .024$  (one-tailed). <sup>d</sup> $p = .027$  (one-tailed). \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

Both focal parameter estimates of the SERPA were meaningful. The adjusted latent mean difference on self-efficacy to regulate barriers to physical activity participation determined by external others (Factor 1) was statistically significant and approximately small in size,  $\eta_1 = .28$ ,  $p = .024$ ,  $d = .28$  [.05, .50], for the FFW group as compared to the UC group. The adjusted latent mean difference on self-efficacy to regulate internally determined barriers to physical activity participation (Factor 2) was statistically significant and approximately small in size,  $\eta_2 = .24$ ,  $p = .027$ ,  $d = .24$  [.04, .45], .28 for the FFW group as compared to the UC group. This pair of findings provided evidence for the effectiveness of the FFW intervention to promote self-efficacy to regulate physical activity in adults with obesity.

## **STUDY 2C CONCLUSION**

In response to Research question 3, the FFW intervention exerted positive direct effects on both Factor 1, self-efficacy to regulate barriers to physical activity participation determined by external others and Factor 2, self-efficacy to regulate internally determined barriers to physical activity participation. The model also explained approximately 50% of the change in latent self-efficacy to regulate physical activity between the FFW and the UC groups even though the FFW effect size was small. This is evidence for the external validity of the measures produced by the SERPA and further provides support for the hypothesized FFW conceptual model (i.e., that the FFW intervention exerts a direct effect on self-efficacy to regulate physical activity at T2 while controlling for demographic covariates and self-efficacy to regulate physical activity at T1; Myers et al., 2020). Additionally, this finding extends findings conducted using a traditional observed score approach in Myers et al. (2020) and has important conceptual implications.

## **BRIEF GENERAL DISCUSSION**

The overarching aim of this exploratory study was to improve the measurement of self-efficacy to regulate physical activity as measured by the SERPA (Myers et al., 2020) using a latent variable framework. The current study has uncovered three major outcomes. First, the SERPA appears to have a two-dimensional factor structure. Second, there is evidence of at least partial temporal measurement invariance of the SERPA. Third, there is evidence of external validity of the scores produced by the SERPA and therefore evidence for the effectiveness of the FFW intervention. The following paragraphs briefly discuss these major outcomes.

The majority of the evidence supported a two-dimensional model for the SERPA responses. Scores from scales measuring self-efficacy beliefs in a specified domain however, are

often treated as unidimensional under a traditional observed score approach (Bandura, 1997, 2006; Myers & Feltz, 2007). For example, most scales purporting to measure a construct that could similarly be conceptualized as self-regulatory efficacy associated with physical activity appear to be treated as unidimensional (Bateman et al. 2021). Two scales however, the self-efficacy for exercise behaviors scale (SEB; Sallis et al., 1988) and the self-efficacy sub-scale of the health behavior scale (HBS; Anderson et al., 2007) appear to consist of two-dimensional structures. The dimensions reported for these two scales however appear to be conceptually different from the SERPA as determined in the current study. The SEB dimensions are labelled “resisting relapse” and “making time for exercise”. The HBS dimensions are labelled “self-efficacy for overcoming barriers to increasing physical activity” and “self-efficacy for integrating physical activity in the daily routine”. The dimensions of the SERPA however, appear to be based on an evaluation of social or interpersonal involvement in one’s appraisal of their personal capability to overcome barriers to physical activity. Evidence for multidimensionality of the SERPA along with similar findings using other scales suggests that the self-efficacy to regulate physical activity construct may have a multidimensional structure in general. The two-dimensional SERPA (i.e., self-efficacy to regulate physical activity) structure also coincides with the two-dimensional PASE (i.e., self-efficacy for varying levels of physical activity) scales structure identified in Myers et al. (2021). These similar findings provide growing evidence that different forms of self-efficacy associated with physical activity - often assumed to be unidimensional – may be more accurately viewed as multidimensional constructs.

The majority of the evidence suggested at least partial longitudinal measurement invariance of the SERPA responses. Evidence for measurement invariance is an important component in providing evidence for construct validity (AERA, APA & NCME, 2014).

Invariance across measurement occasions for which scores on a construct are expected to change over time is also of particular importance (Millsap, 2012). Finding evidence for at least partial longitudinal invariance of the SERPA is evidence for the generalizability of SERPA measurements and its latent construct across the two measurement occasions in the WBPA study. The SERPA constructs therefore appear to have approximately similar measurement properties over time (i.e., measured on the same metric). Evidence for at least partial measurement invariance provides further justification for the construct validity of the SERPA measurement model. This finding is of practical importance as it is evidence for the SERPA construct maintaining its theoretical structure across multiple measurement occasions in an intervention study.

There was evidence for the effectiveness of the FFW intervention to exert a direct effect on both dimensions of latent self-efficacy to regulate physical activity participation at T2. This finding is supported by self-efficacy theory (Bandura, 1997) because exposure to sources of self-efficacy information (i.e., the BET I CAN challenges) related to the physical activity domain would be expected to promote growth in self-efficacy to regulate physical activity. Conceptually, FFW appears to influence self-efficacy to regulate physical activity through self-efficacy to regulate barriers to physical activity based on social considerations (Factor 1) and self-efficacy to regulate internally determined barriers to physical activity participation (Factor 2). The Factor 1 conceptualization provides evidence for a social and/or interpersonal component (e.g., offering moral support in the form of encouragement) influencing growth in self-efficacy to regulate physical activity behavior in adults with obesity. One's perception of another person's evaluation of them (e.g., their physical appearance), another person's offering of moral support (e.g., encouragement), and another person's physical support (e.g., joining them) within the context of

physical activity participation may therefore be an important contributor to the strength of one's self-regulatory capability beliefs relating to physical activity participation. The Factor 2 conceptualization provides evidence for internalized personal beliefs and judgments (e.g., an assessment of the ideality of the weather conditions for comfortably engaging in physical activity) influencing growth in self-efficacy to regulate physical activity behavior in adults with obesity. The direct effect of FFW on the two-factor conceptualization of the SERPA provides evidence for targeting not only intrapersonal (e.g., personal beliefs), but also interpersonal (e.g., social support) elements when developing programs for promoting growth in self-efficacy to regulate physical activity in adults with obesity. This approach is congruent with recommendations that different levels of the socio-ecological model should be targeted when developing programs for promoting physical activity (e.g., Sallis et al., 2006).

There are some noteworthy limitations to the current study. First, the exploratory approach taken in this paper does not allow for confirmation of the validity of the SERPA responses. Second, the exploratory approach taken in the current study may also capitalize on chance. Third, Myers et al. (2020) also examined the data analyzed in this manuscript and therefore the two sets of results are not independent. Future studies should be designed within a more confirmatory framework to test similar research questions to the ones posed in the current study to verify the proposed measurement model, temporal invariance, and external validity of the SERPA in an independent sample of adults with obesity. Finally, the current study only examined the effectiveness of the FFW intervention by testing its effect on the proposed two-dimensional structure of latent self-efficacy to regulate physical activity and not the direct effect of latent self-efficacy to regulate physical activity on physical activity behavior measured at T3.



Future studies should explore the relationship between this two-dimensional structure of latent self-efficacy to regulate physical activity and physical activity behavior in adults with obesity.

In summary, the analyses conducted in this study are important as they respond directly to recommendations regarding the valid measurement of psychological constructs (AERA, APA & NCME, 2014), longitudinal measurement invariance (Millsap, 2012), the measurement of self-efficacy beliefs in general (Bandura 1997), and the measurement of self-efficacy in physical human performance (Bandura 2006; Feltz et al., 2008). The present paper explored and found evidence for a two-factor structure, at least partial temporal measurement invariance, and external validity of latent self-efficacy to regulate physical activity as measured by the SERPA (Myers et al., 2020). The SERPA appears to have a two-dimensional factor structure. Factor 1 was conceptualized as self-efficacy to regulate barriers to physical activity participation determined by external others. Factor 2 was conceptualized as self-efficacy to regulate internally determined barriers to physical activity participation. The residual variance for only one item had to be freed to provide strong evidence for partial temporal measurement invariance. The FFW intervention exerted meaningful, positive direct effects on both Factor 1 and Factor 2 of the latent self-efficacy to regulate physical activity construct, which was evidence for the effectiveness of FFW in promoting self-efficacy to regulate physical activity in adults with obesity.

## APPENDICES

## APPENDIX A: Search Strategy: Embase

1. 'self-efficacy':ab,ti
2. 'physical activity':ab,ti
3. 1 and 2
4. 'trial':ti
5. 'intervention':ti
6. 4 or 5
7. 3 and 6
8. [english]/lim
9. [adult]/lim
10. [2010-2020]/py
11. 7 and 8 and 9 and 10

## APPENDIX B: Reviewer Rating Scale

For each physical activity self-efficacy scale please respond to the items below (with yellow highlights) using the response categories provided based on the extent each scale meets each criterion. Your responses to each item should be based on the current study.

1. The self-efficacy instrument measures beliefs about an appropriate judgement (e.g., “can do” not “will do”)

| No | Somewhat | Yes |
|----|----------|-----|
| 0  | 1        | 2   |

2. The self-efficacy instrument measures beliefs about a current (e.g., right now) capability

| No | Somewhat | Yes |
|----|----------|-----|
| 0  | 1        | 2   |

3. The self-efficacy instrument measures beliefs about a future (e.g., in the next week) capability

| No | Somewhat | Yes |
|----|----------|-----|
| 0  | 1        | 2   |

4. The self-efficacy instrument measures beliefs about the strength of a personal capability (e.g., “how confident are you...”)

| No | Somewhat | Yes |
|----|----------|-----|
| 0  | 1        | 2   |

5. The self-efficacy instrument measures beliefs about a specific level (e.g., frequency, intensity, duration, etc.) of physical activity

| No | Somewhat | Yes |
|----|----------|-----|
| 0  | 1        | 2   |

6. The self-efficacy instrument measures beliefs about a clear frequency level (e.g., times per week) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

7. The self-efficacy instrument measures beliefs about a clear intensity level (e.g., low, moderate, vigorous, etc.) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

8. The self-efficacy instrument measures beliefs about a clear duration level (e.g., minutes per day/week) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

9. The self-efficacy instrument measures beliefs specific to the physical activity domain

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

10. The self-efficacy instrument measures beliefs about a personal capability to overcome barriers (e.g., “when the weather is bad”) to engaging in physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

11. The self-efficacy instrument is determined by a conceptual analysis (e.g., experts create and/or modify items)

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

12. The self-efficacy instrument is concordant with the physical activity instrument regarding the frequency level (e.g., times per week) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

13. The self-efficacy instrument is concordant with the physical activity instrument regarding the intensity level (e.g., low, moderate, vigorous, etc.) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

14. The self-efficacy instrument is concordant with the physical activity instrument regarding the duration level (e.g., minutes per day/week) of physical activity

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

15. Empirical evidence (e.g., Cronbach's alpha) is provided for the internal consistency of scores derived from responses to the self-efficacy instrument

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

16. Empirical evidence (e.g., factor analysis) is provided for the dimensionality of measures produced from responses to the self-efficacy instrument

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
| 0  | 1        | 2   |

17. The response scale for the self-efficacy instrument is optimally categorized (i.e., 5 categories or less)

|    |          |     |
|----|----------|-----|
| No | Somewhat | Yes |
|----|----------|-----|

0

1

2

## APPENDIX C: Self-Efficacy to Regulate Physical Activity (SERPA) Scale

Think about how confident you are in your current ability to overcome possible barriers to engagement in a recommended amount of weekly physical activity for health. Examples of a recommended amount of physical activity for health (counting only those physical activities that you do for at least 10 minutes at a time) include:

- at least 150 minutes per week of moderate physical activity;
- or at least 75 minutes per week of vigorous physical activity;
- or an equivalent combination of the two recommendations listed above.

Moderate physical activity refers to activities (e.g., carrying light loads; raking in the garden or yard; bicycling at a moderate intensity; etc.) that take moderate physical effort and make you breathe somewhat harder than normal. Vigorous physical activity refers to activities (e.g., heavy lifting; chopping wood; highly intense bicycling class; etc.) that take hard physical effort and make you breathe much harder than normal.

Rate your confidence for each of the items below.

No Confidence = 0    Low Confidence = 1    Moderate Confidence = 2    High Confidence = 3

Complete Confidence = 4

1. the weather is very bad
2. you are bored with the physical activities available to you
3. you are on vacation
4. you are uninterested in the physical activities available to you
5. you feel physical discomfort while being physically active
6. you have to be physically active by yourself
7. you do not enjoy the physical activities available to you



8. it is difficult to get to a location suitable for being physically active
9. you do not like the physical activities available to you
10. your schedule conflicts with being physically active
11. you feel self-conscious about your appearance while being physically active
12. you do not receive any encouragement for being physically active
13. you are under personal stress

APPENDIX D: Unstandardized Direct Effects ( $\gamma$ ) in the Path Model for Factor 1 and Factor 2 of the demographic covariates at Time 2 (N = 424) Regressed on FFW

Table 11. *Unstandardized Direct Effects ( $\gamma$ ) in the Path Model for Factor 1 and Factor 2 of the demographic covariates at Time 2 (N = 424) Regressed on FFW*

| Predictor                       | Factor 1   |     | Factor 2   |     |
|---------------------------------|------------|-----|------------|-----|
|                                 | $\gamma_1$ | SE  | $\gamma_2$ | SE  |
| Female                          | -.27       | .13 | -.19       | .15 |
| Black                           | -.15       | .22 | -.08       | .21 |
| Hispanic                        | -.37       | .36 | .14        | .25 |
| Vocational or technical school  | .35        | .36 | .24        | .35 |
| Some college                    | -.19       | .35 | -.06       | .32 |
| Undergraduate degree            | .25        | .34 | .09        | .33 |
| Graduate or professional degree | .39        | .36 | -.07       | .35 |
| Married                         | .17        | .17 | .08        | .17 |
| Part-time employment            | .29        | .36 | .64        | .38 |
| Full-time employment            | .62        | .29 | .93        | .27 |
| Retired                         | .16        | .38 | .28        | .39 |
| Age in years                    | -.02       | .01 | -.03       | .01 |
| Income in thousand dollars      | -.00       | .00 | -.00       | .00 |

*Note.* SE = standard error.

## REFERENCES

## REFERENCES

- Ainsworth, B. E., Bassett, D. R., Strath, S. J., Swartz, A. M., O'Brien, W. L., Thompson, R. W., Jones, D. A., Macera, C. D., & Kimsey, D. C. (2000). Comparison of three methods for measuring the time spent in physical activity. *Medicine & Science in Sports & Exercise*, 32, S457–S464. <https://doi.org/10.1097/00005768-200009001-00004>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7<sup>th</sup> ed.). Washington, DC: American Psychological Association.
- Anderson, E. S., Winett, R. A., & Wojcik, J. R. (2007). Self-regulation, self-efficacy, outcome expectations, and social support: social cognitive theory and nutrition behavior. *Annals of behavioral medicine*, 34(3), 304-312. <https://doi.org/10.1007/BF02874555>
- Anderson-Bill, E. S., Winett, R. A., Wojcik, J. R., & Winett, S. G. (2011). Web-based guide to health: relationship of theoretical variables to change in physical activity, nutrition and weight at 16-months. *Journal of medical Internet research*, 13(1), e27. <https://doi.org/10.2196/jmir.1614>
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, 16(3), 397-438. <https://doi.org/10.1080/10705510903008204>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26. doi:10.1146/annurev.psych.52.1.1
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.
- Bateman, A., Myers, N. D., Chen, S., & Lee, S. (2021). Measurement of Physical Activity Self-Efficacy in Physical Activity-Promoting Interventions in Adults: A Systematic Review. *Measurement in Physical Education and Exercise Science*. Advance online publication. <https://doi.org/10.1080/1091367X.2021.1962324>
- Bauman, A. E, Reis R. S., Sallis, J. F., Wells, J. C., Loos, R. J. F., & Martin, B. W. (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet*, 380, 258-271. [https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)

- Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456–466.  
<https://psycnet.apa.org/doi/10.1037/00332909.105.3.456>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., . . . Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*, 1381-1395.  
<https://doi.org/10.1249/01.MSS.0000078924.61453.FB>
- Curry, S. J., Krist, A. H., Owens, D. K., Barry, M. J., Caughey, A. B., Davidson, K. W., ... & Kubik, M. (2018). Behavioral weight loss interventions to prevent obesity-related morbidity and mortality in adults: US Preventive Services Task Force recommendation statement. *Jama*, *320*(11), 1163-1171.  
<http://jamanetwork.com/journals/jama/fullarticle/10.1001/jama.2018.13022>
- Donnelly, J. E., Blair, S. N., Jakicic, J. M., Manore, M. M., Rankin, J. W., & Smith, B. K. (2009). Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Medicine & Science in Sports & Exercise*, *41*(2), 459-471. <https://doi.org/10.1249/mss.0b013e3181949333>
- Feltz, D. L., & Chase, M. A. (1998). The measurement of self-efficacy and confidence in sport. In J. L. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 63- 78). Morgantown, WV: Fitness Information Technology.
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport: Research and strategies for working with athletes, teams, and coaches*. Champaign, IL: Human Kinetics.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. *Structural equation modeling: A second course*, *10*(6), 269-314.
- Gilson, T. A., & Feltz, D. L. (2012). Self-efficacy and motivation in physical activity and sports: Mediating processes and outcomes. In G. Roberts & D. Treasure (Eds.), *Advances in motivation in sport and exercise* (3rd ed., pp. 271–297). Champaign, IL: Human Kinetics.
- Gourlan, M. J., Trouilloud, D. O., & Sarrazin, P. G. (2011). Interventions promoting physical activity among obese populations: a meta-analysis considering global effect, long-term maintenance, physical activity indicators and dose characteristics. *Obesity reviews*, *12*(7), e633-e645. <https://doi.org/10.1111/j.1467-789X.2011.00874.x>
- Graham, J. M., Guthrie, A. C., Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling*, *10*, 142-153. [https://doi.org/10.1207/S15328007SEM1001\\_7](https://doi.org/10.1207/S15328007SEM1001_7)

- Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and mimic approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika*, *66*, 373-388. doi:10.1007/BF02294440
- Hancock, G. R., & Mueller, R.O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S.H.C. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Past and present. A festschrift in honor of Karl G. Jöreskog* (pp. 195–261). Chicago: Scientific Software International, Inc.
- Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ: British Medical Journal*, *319*, 670-674. <https://doi.org/10.1136/bmj.319.7211.670>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55. <https://doi.org/10.1080/10705519909540118>
- Ilaqqa, A., Emerenziani, G. P., Guidetti, L., & Baldari, C. (2019). The role of physical activity in adult obesity (Ed.), In *Nutrition in the Prevention and Treatment of Abdominal Obesity* (pp. 123-128). San Diego, CA: Academic Press.
- Jackson, B., Beauchamp, M. R., & Dimmock, J. A. (2020). Efficacy beliefs in physical activity settings: Contemporary debate and unanswered questions. In G. Tenenbaum & R.C. Eklund (Eds.), *The Handbook of Sport Psychology* (4th ed., pp. 57-80). New York: Wiley. <https://doi.org/10.1002/9781119568124.ch4>
- Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., . . . Lancet Physical Activity Series Working Group. (2012). The pandemic of physical inactivity: Global action for public health. *The Lancet (British Edition)*, *380*(9838), 294-305. [https://doi.org/10.1016/S0140-6736\(12\)60898-8](https://doi.org/10.1016/S0140-6736(12)60898-8)
- Lee, S., McMahon, A., Prilleltensky, I., Myers, N. D., Dietz, S., Prilleltensky, O., Pfeiffer, K. A., Bateman, A. G., & Brincks, A. M. (2020). Effectiveness of the fun for wellness online behavioral intervention to promote well-being actions in adults with obesity: A randomized controlled trial. *Journal of Sport & Exercise Psychology*. Advance online publication. <https://doi.org/10.1123/jsep.2020-0049>
- Little, R.J.A. and Rubin, D. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471-491. <https://psycnet.apa.org/doi/10.1037/a0019227>

- McAuley, E. (1992). The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. *Journal of Behavioral Medicine*, *15*(1), 65-88.  
<https://psycnet.apa.org/doi/10.1007/BF00848378>
- McAuley, E., Poag, K., Gleason, A., & Wraith, S. (1990). Attrition from exercise programs: Attributional and affective perspectives. *Journal of Social Behavior and Personality*, *5*(6), 591.
- Millsap, R. E. (2012). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge.
- Millsap, R. E., & Kwok, O-M. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods*, *9*, 93–115.  
<https://psycnet.apa.org/doi/10.1037/1082-989X.9.1.93>
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Myers, N. D., Bateman, A. G., McMahon, A., Prilleltensky, I., Lee, S., Prilleltensky, O., ... & Brincks, A. M. (2021). Measurement of Physical Activity Self-Efficacy in Adults With Obesity: A Latent Variable Approach to Explore Dimensionality, Temporal Invariance, and External Validity. *Journal of Sport and Exercise Psychology*, *43*(6), 497-513.
- Myers, N. D., & Feltz, D. L. (2007). From self-efficacy to collective efficacy in sport: Transitional methodological issues. In G. Tenenbaum & R.C. Eklund (Eds.), *The handbook of sport psychology* (pp. 799–819). Wiley.
- Myers, N. D., Jin, Y., Ahn, S., Celimli, S., & Zopluoglu, C. (2015). Rotation to a partially specified target matrix in exploratory factor analysis in practice. *Behavior research methods*, *47*(2), 494-505. <https://doi.org/10.3758/s13428-014-0486-7>
- Myers, N. D., McMahon, A., Prilleltensky, I., Lee, S., Dietz, S., Prilleltensky, O., ... & Brincks, A. M. (2020). Effectiveness of the fun for wellness web-based behavioral intervention to promote physical activity in adults with obesity (or overweight): Randomized controlled trial. *JMIR formative research*, *4*(2), e15919. <https://doi.org/10.2196/15919>.
- Myers, N. D., Prilleltensky, I., Hill, C. R., & Feltz, D. L. (2017). Well-being self-efficacy and complier average causal effect modeling: A substantive-methodological synergy. *Psychology of Sport & Exercise*, *30*, 135-144.  
<https://doi.org/10.1016/j.psychsport.2017.02.010>
- Myers N. D., Prilleltensky I., Lee S., Dietz S., Prilleltensky O., McMahon A., ... Brincks, A. M. (2019). Effectiveness of the fun for wellness online behavioral intervention to promote well-being and physical activity: Protocol for a randomized controlled trial. *BMC Public Health*, *19*:737. doi:10.1186/s12889-019-7089-2

- Rubenstein, C. L., Duff, J., Prilleltensky, I., Jin, Y., Dietz, S., Myers, N. D., & Prilleltensky, O. (2016). Demographic group differences in domain-specific well-being. *Journal of Community Psychology, 44*, 499-515. <https://doi.org/10.1002/jcop.21784>
- Sallis, J. F., Bull, F., Guthold, R., Heath, G. W., Inoue, S., Kelly, P., . . . Lancet Physical Activity Series 2 Executive Committee. (2016). Progress in physical activity over the olympic quadrennium. *The Lancet (British Edition), 388*(10051), 1325-1336. [https://doi.org/10.1016/S0140-6736\(16\)30581-5](https://doi.org/10.1016/S0140-6736(16)30581-5)
- Sallis, J. F., Cervero, R. B., Ascher, W., Henderson, K. A., Kraft, M. K., & Kerr, J. (2006). An ecological approach to creating active living communities. *Annu. Rev. Public Health, 27*, 297-322. <https://doi.org/10.1146/annurev.publhealth.27.021405.102100>
- Sallis, J. F., Pinski, R. B., Grossman, R. M., Patterson, T. L., & Nader, P. R. (1988). The development of self-efficacy scales for healthrelated diet and exercise behaviors. *Health Education Research, 3*(3), 283-292. <https://doi.org/10.1093/her/3.3.283>
- Saris, W.E., Satorra, A., & van der Veld, W. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling, 16*, 561-582. <https://doi.org/10.1080/10705510903203433>
- Scarpa, M. P., Prilleltensky, I., McMahon, A., Myers, N. D., Prilleltensky, O., Lee, S., Pfeiffer, K. A., Bateman, A. G., & Brincks, A. M. (2021). Is fun for wellness engaging? Evaluation of user experience of an online intervention to promote well-being and physical activity. *Frontiers in Computer Science*. Advance online publication. <https://doi.org/10.3389/fcomp.2021.690389>
- Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods, 7*(2), 147. <https://psycnet.apa.org/doi/10.1037/1082-989X.7.2.147>
- Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27*, 229-239. <https://doi.org/10.1111/j.2044-8317.1974.tb00543.x>
- Spink, K. S., & Nickel, D. (2010). Self-regulatory efficacy as a mediator between attributions and intention for health-related physical activity. *Journal of Health Psychology, 15*(1), 75-84. <https://doi.org/10.1177%2F1359105309342308>
- Tran, L., Tran, P., & Tran, L. (2020). A cross-sectional examination of sociodemographic factors associated with meeting physical activity recommendations in overweight and obese US adults. *Obesity research & clinical practice, 14*(1), 91-98. <https://doi.org/10.1016/j.orcp.2020.01.002>
- Tudor-Locke, C., Brashear, M. M., Johnson, W. D., & Katzmarzyk, P. T. (2010). Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese U.S.



men and women. *International Journal of Behavioral Nutrition and Physical Activity*, 7, 60. <https://doi.org/10.1186/1479-5868-7-60>

United States Department of Health and Human Services. (2013). *Managing overweight and obesity in adults: Systematic evidence review from the obesity expert panel*. Retrieved from <https://www.nhlbi.nih.gov/sites/default/files/media/docs/obesity-evidence-review.pdf>

World Health Organization. (2018). *Obesity and overweight fact sheet*. Retrieved from <http://www.who.int/mediacentre/factsheets/fs311/en/>

## CHAPTER IV: SUMMARY & CONCLUSIONS

Increasing self-efficacy associated with physical activity is a frequently employed mechanism for promoting physical activity in adults (Bauman et al., 2012). The Well-Being and Physical Activity Study (WBPA) utilized self-efficacy as the behavior change theory for increasing physical activity in adults with obesity (Myers et al., 2019). The WBPA study employed two forms of self-efficacy, physical activity self-efficacy (PASE) and self-efficacy to regulate physical activity (SERPA), as potential modifiable mediators of physical activity in deploying the Fun For Wellness (FFW) intervention (Myers et al., 2020). However, valid measurement of psychological constructs is important for the advancement of legitimate science in health psychology and related disciplines (Williams & Rhodes, 2016). The valid measurement of self-efficacy constructs is therefore important to the efficacy of the physical activity-promoting interventions in which they are used. The purposes of this dissertation were to (a) identify the issues in the measurement of self-efficacy associated with physical activity in physical activity promoting interventions for adults and to (b) explore evidence for the validity of the measurement of self-efficacy to regulate physical activity in the WBPA study.

A systematic review (Study 1) of the theoretical and measurement quality of scales measuring physical activity self-efficacy in physical activity-promoting interventions identified fourteen distinct multi-item scales with varying numbers of items and other properties. The review highlighted some important theoretical and measurement-related issues associated with the use of these scales. First, many studies did not consistently ensure concordance between self-efficacy measurement and physical activity measurement (e.g., not specifying levels of physical activity associated with self-efficacy measurement). Second, many studies did not consistently emphasize some essential conceptual components of self-efficacy measurement such as not

emphasizing the timeframe (e.g., current or future) for which the self-efficacy measurement applies. Third, many studies did not accurately distinguish and label the form (e.g., task or self-regulatory efficacy) of self-efficacy beliefs being measured by the scale. This issue is of particular importance because most of the scales examined in the review were identified as measuring self-efficacy to regulate physical activity and not physical activity self-efficacy (i.e., task-related). Fourth, many studies did not provide complete evidence for the dimensionality, or validity (e.g., only reporting Cronbach's alpha as evidence of scale reliability) of the physical activity self-efficacy scale used. Fifth, many studies did not use self-efficacy scales with optimally categorized (e.g., not more than five) Likert response scales. Considering these issues, Study 2 explored validity evidence for the measurement of self-efficacy to regulate physical activity using SERPA), a modified version of the barriers self-efficacy scale (McAuley, 1992) which was developed according to self-efficacy theory and measurement recommendations to fit the FFW context (Myers et al., 2020).

Study 2 found evidence for a two-dimensional structure, temporal measurement invariance, and external validity of the SERPA scale (Study 2). The two dimensions were conceptualized as, 1) self-efficacy to regulate barriers to physical activity participation based on social considerations and 2) self-efficacy to regulate internally perceived barriers to physical activity participation. There was evidence of external validity (i.e., effectiveness) of the FFW intervention to exert a direct effect on the proposed two-dimensional structure of latent SERPA in adults with obesity at Time 2 (T2). Study 2 therefore provided strong cumulative evidence for the validity of scores produced by the SERPA along with evidence for the effectiveness of the FFW intervention.

The findings of Study 2 have important implications for valid measurement of self-efficacy within the context of human physical performance, especially considering the conceptual and measurement-related issues uncovered in Study 1. The findings of Study 1 imply that adhering to the conceptual and measurement-related recommendations for the measurement of self-efficacy will result in increased validity of measurement. The SERPA (Appendix C) was therefore developed based on recommendations (e.g., measuring strength of current capability judgment) for valid self-efficacy measurement according to self-efficacy literature. The SERPA was therefore accurately labeled as a self-regulatory efficacy construct measuring confidence in overcoming barriers to physical activity using an optimally categorized response scale (i.e., 5 response categories). Importantly, the SERPA was also designed to be concordant (i.e., capability judgement aligned with levels of physical activity) with the measure of physical activity in terms of frequency, intensity, and duration. Consequently, it follows that the overall findings of Study 2 strongly supported valid measurement of self-efficacy to regulate physical activity in the WBPA study. There is therefore strong empirical support that following the recommendations (e.g., ensuring concordance between self-efficacy and physical activity measurement) for valid self-efficacy measurement (i.e., as outlined in Study 1) within the physical activity context is associated with valid measurement. Scales measuring various forms of self-efficacy associated with physical activity should therefore be meticulously designed according to self-efficacy theory and measurement literature to facilitate improved validity.

### **STRENGTHS AND LIMITATIONS**

Increasing physical activity on a population level is a complex problem requiring solutions across multiple sectors (Pratt et al., 2020). One part of this complex problem is ensuring the validity of the measurement of the psychosocial constructs targeted in physical

activity interventions. This dissertation provides evidence that there are some issues in the measurement of self-efficacy associated with physical activity in physical activity interventions. Fortunately, strong evidence was found that the SERPA, which was carefully developed based on self-efficacy theory and measurement principles, is valid for measuring self-efficacy to regulate physical activity in adults with obesity. Another part of the complex problem of increasing physical activity on a population level is the paucity of intervention studies testing the effectiveness of physical activity promoting programs (Pratt et al., 2020). This dissertation also provides evidence for the effectiveness of the FFW intervention to increase latent self-efficacy to regulate physical activity in adults with obesity. These major findings are significant and therefore make major contributions to the fields of physical activity-promotion and self-efficacy theory and measurement.

The current study has a few noteworthy limitations. First, the scope of the systematic review (Study 1) did not allow for associating quality self-efficacy measurement with the effectiveness of the physical activity intervention outcomes. Second, the specific focus of Study 1 on the initial publication of the physical activity self-efficacy scales identified did not completely account for subsequent modifications to the scales and did not account for validity evidence of the measurement external to the content of the scale (e.g., the population being measured). Third, Study 2 did not use a confirmatory framework to examine the validity of the SERPA scale because there was no known a priori factor structure for the newly modified scale. Fourth, Study 2 did not examine the effect of latent SERPA on physical activity participation at T3 in the WBPA study.

## **CONCLUSIONS**

At least three major conclusions can be made based on the combined results of Study 1 and Study 2. First, though there may be issues in the current measurement of various forms of self-efficacy associated with physical activity in physical activity-promoting interventions, there is considerable evidence for the effectiveness of the FFW intervention increasing self-efficacy to regulate physical activity in adults with obesity in the WBPA study. Second, self-efficacy to regulate physical activity appears to be an important modifiable conceptual target mediator of physical activity participation in physical activity-promoting interventions. Third, self-efficacy to regulate physical activity measurement under a traditional observed score approach may not be ideal because it generally assumes a one-dimensional construct and therefore does not account for multi-dimensionality nor measurement error. The SERPA scale appears to be two-dimensional and disregarding its factor structure is likely to comprise valid measurement of the construct.

## **FUTURE RESEARCH DIRECTIONS**

Future intervention studies that promote physical activity should therefore target self-efficacy to regulate physical activity as a mediator of physical activity in different sub-populations. The SERPA could be modified in these studies based on empirical evidence for unique barriers to physical activity that might be relevant to the population being targeted in the intervention. It is important however, that the version of the self-efficacy to regulate physical activity scale being used conforms to self-efficacy theory and recommendations for valid self-efficacy measurement (e.g., as outlined in Bateman et al., 2021).

Determining and implementing evidence-based strategies and implementing these strategies is important for improved physical activity promotion (Pratt et al., 2020). Future

studies should therefore conduct meta-analyses of self-efficacy mediated physical activity-promoting interventions in different populations. Such studies would contribute significantly to public health by providing evidence (e.g., effect sizes) of effective intervention strategies and therefore justify implementing these strategies at scale in order to achieve population-level impact. Additionally, a future study should be conducted to determine if latent SERPA had a positive direct effect on physical activity participation in the WBPA study as further evidence for both valid SERPA measurement and effectiveness of the FFW intervention.

Study 2 of this dissertation focused on testing the validity of the SERPA in an exploratory methodological framework. Future research would benefit from testing validity of SERPA in a more confirmatory framework. This may involve using confirmatory factor analysis to test the two-dimensional structure determined in the current study. Confirming the factor structure conforms with recommendations from self-efficacy literature (e.g., Bandura 1997, 2006; Feltz et al., 2008) and the measurement of psychological constructs in general (AERA, APA, & NCME, 2014). Confirmation of the SERPA factor structure would further improve the measurement of the construct, and therefore by extension the efficacy of future physical activity-promoting interventions targeting the construct as a malleable mediator of physical activity participation.

## REFERENCES



## REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.
- Bateman, A., Myers, N. D., Chen, S., & Lee, S. (2021). Measurement of Physical Activity Self-Efficacy in Physical Activity-Promoting Interventions in Adults: A Systematic Review. *Measurement in Physical Education and Exercise Science*. Advance online publication. <https://doi.org/10.1080/1091367X.2021.1962324>
- Bauman, A. E, Reis R. S., Sallis, J. F., Wells, J. C., Loos, R. J. F., & Martin, B. W. (2012). Correlates of physical activity: why are some people physically active and others not? *The Lancet*, 380, 258-271. [https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport: Research and strategies for working with athletes, teams, and coaches*. Champaign, IL: Human Kinetics.
- McAuley, E. (1992). The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. *Journal of Behavioral Medicine*, 15(1), 65-88. <https://psycnet.apa.org/doi/10.1007/BF00848378>
- Myers, N. D., McMahon, A., Prilleltensky, I., Lee, S., Dietz, S., Prilleltensky, O., ... & Brincks, A. M. (2020). Effectiveness of the fun for wellness web-based behavioral intervention to promote physical activity in adults with obesity (or overweight): Randomized controlled trial. *JMIR formative research*, 4(2), e15919. <https://doi.org/10.2196/15919>.
- Myers N. D., Prilleltensky I., Lee S., Dietz S., Prilleltensky O., McMahon A., ... Brincks, A. M. (2019). Effectiveness of the fun for wellness online behavioral intervention to promote well-being and physical activity: Protocol for a randomized controlled trial. *BMC Public Health*, 19:737. doi:10.1186/s12889-019-7089-2
- Pratt, M., Varela, A. R., Salvo, D., Kohl III, H. W., & Ding, D. (2020). Attacking the pandemic of physical inactivity: what is holding us back? *British journal of sports medicine*, 54(13), 760-762. <http://dx.doi.org/10.1136/bjsports-2019-101392>

Williams, D. M., & Rhodes, R. E. (2016). The confounded self-efficacy construct: Conceptual analysis and recommendations for future research. *Health Psychology Review, 10*(2), 113-128. <https://doi.org/10.1080/17437199.2014.941998>