Factors Influencing Level of Implementation of Physical Activity Interventions in Youth-Serving Organizations

Yuen Yan Lau

University of South Carolina

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Exercise Science Commons

Recommended Citation


This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.
FACTORS INFLUENCING LEVEL OF IMPLEMENTATION OF PHYSICAL ACTIVITY INTERVENTIONS IN YOUTH-SERVING ORGANIZATIONS

by

Yuen Yan Lau

Bachelor of Education
The Chinese University of Hong Kong, 2007

Master of Philosophy
Hong Kong Baptist University, 2009

Submitted in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy in
Exercise Science
The Norrman J. Arnold School of Public Health
University of South Carolina
2015

Accepted by:
Russell R. Pate, Major Professor
Ruth P. Saunders, Committee Member
Abraham Wandersman, Committee Member
Michael W. Beets, Committee Member
Lacy Ford, Senior Vice Provost and Dean of Graduate Studies
DEDICATION

To Raymond, Hailey, and my parents.
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Pate, for his excellent guidance, patience, encouragement, and for providing me with an excellent atmosphere for doing research. I would like to thank Dr. Saunders and Dr. Wandersman for their inspirations that shape my path to implementation research. I would also like to thank Dr. Beets for his guidance and his patience in correcting my writings. Thanks to Dr. Bo Cai for helping with my statistical analyses.

I would like to thank Dr. Jennifer O’Neill and Dr. Sharon Ross for their comments and suggestions to my dissertation. Many thanks to everyone in the Children’s Physical Activity Research group for their support.

I will never be able to thank my husband Raymond enough for everything he’s done and the sacrifices he made over the past four years. Thank you to my daughter Hailey; your smile is the best medication for stress relief.

Lastly, I would like to gratefully thank my parents for always believing in me, for the constant support, and unconditional love.
**ABSTRACT**

In the past decade, numerous interventions have been developed and tested to increase physical activity in children and adolescents. Among these interventions, those that reported higher levels of program implementation appear to have better program outcomes. However, relatively little is known about the specific factors that contributed to successful implementation in youth physical activity interventions. The overall purpose of this dissertation was to identify a set of core factors that are most important in explaining implementation of physical activity interventions in youth-serving settings. Three studies, an expert panel study and two prospective observational studies, were conducted to address the purpose of this dissertation.

In the first study, an expert panel was convened to identify factors that are most important in achieving successful implementation of physical activity interventions in youth-serving organizations. Five recognized experts participated in a four-round, modified Delphi process to identify factors related to implementation of youth physical activity interventions, and to quantify the importance of the identified factors. Experts’ opinions were translated into Bayesian predictive models for factor selections. These processes resulted in a final list of 15 factors, in which five factors were classified as organizational characteristics, six factors as implementation processes, two factors as provider
characteristics, and two factors as program characteristics and community-level factors, respectively.

The second study analyzed data from a previously completed preschool-based intervention. Participants were preschool classrooms enrolled in the first two years of the intervention (year 1: n= 19, year 2: n=17). The purpose was to examine the direct and indirect effects of preschool characteristics, teacher characteristics, and quality of implementation processes on level of implementation. The results of Bayesian path analysis show that the three selected factors were not significantly associated with level of implementation in year 1. Preschool characteristics were found to be directly associated with level of implementation in year 2. The third study analyzed data from a previously studied physical activity intervention carried out in 24 residential children’s homes (RCHs). The purpose was to examine the direct and indirect effects of RCH characteristics, wellness team characteristics, and quality of implementation processes on level of implementation. The results of Bayesian path analysis reveal that RCH characteristics and wellness team characteristics are directly associated with level of implementation. Overall, this dissertation found consistent evidence supporting the direct contribution of organizational characteristics in achieving successful implementation of physical activity interventions in youth-serving organizations. However, the influence of provider characteristics and quality of implementation processes on level of implementation appear to vary across interventions with different designs, at different implementation stages, and in different implementation settings.
TABLE OF CONTENTS

Dedication ............................................................................................................. iii
Acknowledgements ............................................................................................... iv
Abstract ................................................................................................................. v
List of Tables ........................................................................................................ x
List of figures ...................................................................................................... xi
Chapter 1: Overall Introduction ........................................................................... 1
Chapter 2: Manuscript 1: Factors influencing successful implementation of physical activity interventions in youth-serving organizations: An expert perspective ........................................................................................................ 9
  Abstract ........................................................................................................ 10
  Introduction ................................................................................................. 11
  Methods ..................................................................................................... 12
  Results ...................................................................................................... 20
  Discussion ................................................................................................ 22
  References ................................................................................................ 37

Chapter 3: Manuscript 2: Factors influencing implementation of a preschool-based physical activity intervention ................................................................. 41
  Abstract ................................................................................................ 42
  Introduction ............................................................................................. 43
  Methods .................................................................................................. 45
  Results .................................................................................................... 54
Appendix B – Observed degrees of implementation aspects of individually-oriented interventions .............................................................. 259

Appendix C – Observed degrees of implementation aspects of environmental interventions ................................................................. 265
LIST OF TABLES

Table 2.1 An example of using a three-factor Bayesian predictive model to estimate posterior odds of success ...................................................... 28

Table 2.2 Descriptions for the identified factors at the three factor-level: high, moderate, and low ...................................................................................... 29

Table 2.3 Likelihood ratios and diagnostic power scores for each factor-level ........................................... 36

Table 3.1 Descriptive statistics of preschool classrooms participated in year 1 and year 2 of the SHAPES intervention ................................................................. 59

Table 3.2 Direct and indirect effects of the selected factors on level of implementation of SHAPES-Move Outside .................................................................................. 60

Table 4.1 Descriptive statistics of residential children’s homes participated in ENRICH .................................................................................................................. 87

Table 4.2 Direct and indirect effects of the selected factors on level of implementation of ENRICH ................................................................................................. 89

Table 5.1. Summary of influences of program implementation identified in the three studies conducted in this dissertation ......................................................... 102
LIST OF FIGURES

Figure 2.1 Sample questions for the survey conducted in the third round ........27

Figure 3.1 A conceptual framework for examining factors influencing implementation of SHAPES-Move Outside.................................................................46

Figure 3.2 Final path models illustrating factors that influence implementation of the Move Outside component during year 1 and year 2 of the SHAPES intervention. .................................................................61

Figure 4.1 A conceptual framework for examining factors that influence level of implementation of ENRICH..............................................................................74

Figure 4.2 Final path model illustrating factors that influence level of implementation of ENRICH .................................................................88
CHAPTER 1

OVERALL INTRODUCTION

Promoting regular participation in physical activity among American children and adolescents is a national health priority.\textsuperscript{1} Over the years, researchers have developed and tested many interventions to increase youth physical activity in various settings. To date, these interventions have experienced very limited success.\textsuperscript{2-5} Although the majority of these interventions demonstrated positive effects on youth’s physical activity behaviors, the magnitude of the effects were modest.\textsuperscript{2-5} A systematic review and meta-analysis study\textsuperscript{4} showed that children and adolescents who participated in physical activity interventions engaged in approximately 4 minutes more walking or running per day.

Several researchers\textsuperscript{6,7} suggest that the lack of significant intervention effects can be attributed to three factors: invalid program theory, sub-optimal program implementation, and inappropriate outcome evaluation. An intervention is unlikely to yield the expected outcomes when guided by an invalid program theory because it is manipulating factors that have low relevance or have no causal connections with the program outcomes.\textsuperscript{8} An intervention is less likely to succeed with sub-optimal implementation because it is unlikely to exert a sufficient amount of influence to change the targeted program outcomes.\textsuperscript{6}
The chance of detecting a substantial intervention effect may greatly reduce when the intervention is assessed too soon, with poorly designed evaluation plans, or with unsuitable evaluation measures.\textsuperscript{9, 10}

While all three factors are important in influencing intervention effectiveness, research attention has been largely devoted to refining program theory and advancing outcome evaluation methods, but program implementation has been relatively neglected.\textsuperscript{9, 11} Thus far, only one systematic review\textsuperscript{12} focused on the implementation of youth physical activity interventions. Naylor et al.\textsuperscript{12} conducted a systematic review examining the relationships between implementation and effectiveness of school-based physical activity interventions. With so many school-based physical activity interventions available in the literature, this review only identified 15 eligible studies. Another ongoing review (E.Y. Lau, unpublished data, 2015) concentrated on the implementation of physical activity interventions in youth-serving organizations such as schools and childcare centers, and showed that only 48 of the 183 eligible interventions have assessed level of implementation.

Though there is limited number of included studies, both reviews\textsuperscript{12, 13} found evidence supporting the positive relationships between level of implementation and program outcomes. Moreover, level of implementation was highly variable across studies, with some studies demonstrating relatively low levels. In terms of completeness (i.e., proportions of intervention components being delivered), the range between studies reporting the highest and the lowest degrees was 75%, with the lowest endpoint being 22%. For providers’ adherence to the planned
protocol (i.e., fidelity), the range was 65%, with the lowest endpoint being 33% (E.Y. Lau, unpublished data, 2015). Given the positive relationships between program implementation and intervention effectiveness, it is important to examine the factors that contribute to the variation in implementation across studies, so that relevant strategies can be developed to improve the implementation of future youth physical activity interventions.

To date, there is a plethora of conceptual frameworks proposing a list of factors that are hypothesized to influence level of program implementation. Looking at three frequently used frameworks reveals that there are already over 100 potential factors. However, the factors identified in these frameworks were primarily based on the literature of health services and preventive interventions for youth (e.g., drugs abuse or tobacco prevention programs), which may not be fully applicable in youth physical activity interventions.

Currently, there is one systematic review that identifies 22 factors affecting implementation of youth physical activity interventions, but it is limited to school settings. While schools are important settings for promoting physical activity in children and adolescents, there have been an increasing number of studies focusing on other youth-serving settings, such as childcare centers and afterschool programs. There is a need to expand our understanding of the factors that influence implementation of physical activity interventions in broader youth-serving settings.

The purpose of this dissertation project was to identify a set of core factors that are most important in explaining implementation of physical activity
interventions in youth-serving organizations. To enhance the validity of the findings, this dissertation project employed a mixed-method approach to answer the research question of interest, which consisted of an expert panel study and two secondary data analyses of previously completed interventions.

The first study convened a panel of experts to identify factors that are most important to achieve successful implementation of physical activity interventions undertaken in youth-serving organizations. Five recognized experts engaged in a four-round, modified Delphi process to identify a list of potential factors and provide numerical estimates regarding the importance of these factors in contributing to the likelihood of a successful implementation. Experts’ opinions were then translated into a Bayesian predictive model for selecting the final list of factors.

The second and third studies analyzed data from two previously completed physical activity interventions undertaken in preschools and residential children’s homes (RCHs), respectively. The purpose of the second study was to examine the direct and indirect effects of three selected factors on level of implementation of a preschool-based physical activity intervention. The three selected factors were preschool characteristics, teacher characteristics, and quality of implementation processes, and were constructed by ten elements referenced in the literature. Bayesian path analyses were used to examine how the three factors influence level of implementation.

The purpose of the third study was to examine direct and indirect effects of three selected factors on level of implementation of an RCH-based physical
activity intervention. The three factors were RCH characteristics, wellness team characteristics, and quality of implementation processes, and were constructed by ten elements referenced in the literature. Bayesian path analyses were conducted to investigate the relationships between the three factors and level of implementation.

Cumulatively, the three studies conducted in this dissertation would provide important findings to expand our understanding of factors that influence program implementation specifically within the context of physical activity interventions undertaken in youth-serving organizations. These findings may have important implications for those who seeking to develop measures for assessing influences of program implementation and improve implementation planning of youth physical activity interventions. In turn, better understanding and planning of program implementation may result in improved intervention effectiveness.


CHAPTER 2

MANUSCRIPT 1: FACTORS INFLUENCING SUCCESSFUL IMPLEMENTATION OF PHYSICAL ACTIVITY INTERVENTIONS IN YOUTH-SERVING ORGANIZATIONS:

AN EXPERT PERSPECTIVE$^1$

---

$^1$ Lau EY, Saunders RP, Beets MW, Wandersman A, Cai B, Pate RR. To be submitted to Evaluation and Program Planning.
Abstract

**Background:** Little is known about the factors that influence implementation of physical activity interventions undertaken in youth-serving settings, thus impeding the development of effective implementation strategies. This study convened a panel of experts to identify factors that are most important in achieving successful implementation of physical activity interventions in youth-serving organizations.

**Methods:** Five recognized experts participated in a four-round, modified Delphi consensus process. The panelists were asked to achieve consensus on a list of potential factors that are most important in predicting successful implementation and to provide estimates regarding the individual contributions for each of the identified factors in predicting a successful implementation. These estimates were then translated into Bayesian predictive models for factor selection.

**Results:** During the first two rounds, the expert panel identified 23 factors. The factor selection procedures indicate that a final model containing 15 factors yielded the greatest contributions in predicting successful implementation of youth physical activity interventions. In this final model, five factors were classified as organizational characteristics, six factors as implementation processes, two factors as provider characteristics, and two factors as program characteristics and community-level factors, respectively.

**Conclusions:** The factors identified in this study provide important information to inform implementation planning and evaluation of future interventions.
Introduction

Achieving optimal program implementation is challenging in many field-based health promotion interventions, including physical activity interventions carried out in youth-serving organizations. A school-based intervention targeting physical activity in high school girls showed that 41% of the intervention schools did not achieve the intended levels of implementation.\textsuperscript{1} A preschool-based study found that 30% of the preschool teachers did not deliver the intervention components as planned.\textsuperscript{2} A community-based intervention targeting physical activity of children living in residential children’s homes also reported that 40% of the intervention homes did not meet the implementation criteria.\textsuperscript{3, 4} Researchers have suggested that sub-optimal program implementation may dilute intervention effects, thus masking the potential benefits of a program.\textsuperscript{5-9} Therefore, it is important to understand the factors that influence program implementation so that relevant strategies can be developed.

Reviews on implementation of health promotion programs have identified a list of potential factors that influence program implementation. Damschroder and colleagues\textsuperscript{10} synthesized 19 existing implementation frameworks and identified 31 factors related to program implementation in health service settings. Durlak and DuPre\textsuperscript{5} reviewed 59 intervention studies and identified 23 factors influencing implementation of health promotion interventions targeting children and adolescents. However, factors identified in these two reviews were based primarily on the literature of health services and preventive interventions (e.g.,
drugs abuse or tobacco prevention programs), which may not be fully applicable to youth physical activity interventions.

Thus far, one systematic review\textsuperscript{11} has identified 22 factors affecting implementation of youth physical activity interventions, but it was limited to school settings. While schools are important settings for promoting physical activity in children and adolescents, there have been an increasing number of studies focusing on other youth-serving settings such as childcare centers and afterschool programs.\textsuperscript{12, 13} It is, therefore, important to identify a set of core factors that explain variations in implementation of physical activity interventions across youth-serving organizations for future studies to adopt, modify, and test.

The identification of such factors would ideally be based on empirical data. However, empirical studies are lacking in the field, thus another valuable source of information would be experts’ opinions. The purpose of the current study was to convene a panel of experts to identify factors that are most important in achieving successful implementation of physical activity interventions in youth-serving organizations.

Methods

\textit{A Bayesian predictive model for factor identification}

Researchers have suggested that program implementation is influenced by multiple factors. These factors can interact with each other, which creates a complex system that influences program implementation.\textsuperscript{5, 10, 14} To understand this complex system, we need to identify a set of factors that are collectively important in explaining variations in program implementation.
To systematically assess the contributions of a set of factors, the present study asked a panel of experts to identify a list of factors that are most important in explaining implementation of physical activity interventions in youth-serving organizations and quantified their importance. Then, we used Bayesian statistics to translate experts’ ratings into predictive models. By comparing different models, we determined which combination of factors most influences the outcome of interest. This approach has been successfully used to identify factors influencing implementation of tobacco prevention programs in schools\textsuperscript{15, 16} and health-related programs in other settings\textsuperscript{17-20}.

In the simplest form a Bayesian model assumes a dichotomous outcome, which is successful implementation or unsuccessful implementation for this study. To create the model, the following information was required from the expert panelists: 1) an operational definition of successful implementation, 2) a set of conditionally independent factors that are important in predicting successful implementation, 3) likelihood ratios of each of the identified factors, and 4) estimates for testing internal validity of the model. These components are described in detail in this paper.

**Expert panel participants**

Previous studies\textsuperscript{15, 17, 19} suggested that a panel size between five and seven members would provide optimal information for developing a Bayesian predictive model. A purposive sampling procedure was used\textsuperscript{21}. To ensure consistency of expertise levels within the panel, we targeted senior researchers who have substantial experience in implementing youth physical activity
interventions. Eligibility of the panelists were as follows: 1) academic appointment at the rank of associate professor or higher, 2) a track record of leading implementation of youth physical activity interventions, and 3) a demonstrated record of publications on process evaluation and implementation of youth physical activity interventions.

Based on the eligibility criteria, members of the study team generated a list of panelists by reviewing journal articles and faculties’ biographical descriptions on university websites, and consulting with senior researchers. Since the objective of this study was to identify a set of core factors that influence implementation of physical activity interventions across youth-serving organizations, we attempted to obtain a balance of individuals with expertise in various settings, such as schools and communities. Invitation letters were sent to six researchers. Five experts agreed to participate and one did not respond. The final list of panelists consisted of four professors and one professor emeritus. These panelists had an average of over 20 years of experiences in implementing, monitoring, evaluating, and publishing results of youth physical activity interventions in a variety of settings, including preschools, schools, afterschool programs, and communities (e.g., summer camps, troops).

**Data collection**

The five experts participated in a four-round, modified Delphi process. In the first round, the panelists completed an online survey to independently define successful implementation and suggest factors that influence successful implementation. The second round provided a group setting for the panelists to
elaborate and discuss their views on the definition and influences of successful implementation through a video conference. The third round required the panelists to complete another online survey to independently rate the importance of the suggested factors. The fourth round involved a final online survey that collected data for assessing test-retest reliability of panelists’ ratings. The design of the surveys and video conference were guided by previous studies.\textsuperscript{15, 17} Data were collected between February and May 2015. The Institutional Review Board at the University of South Carolina approved all study procedures.

First round

The online survey consisted of seven open-ended questions that required the panelists to 1) operationalize successful implementation of a physical activity intervention carried out in youth-serving organizations based on their own experience, 2) suggest six factors that are most important in predicting successful implementation, and 3) describe the suggested factors at the three factor-levels: high, moderate, and low. Responses were aggregated and summarized into a straw model containing all of the suggested factors and circulated among panelists for review before the video conference.

Second round

All panelists participated in a 90-minute video conference one month after the first survey. Section one of the video conference provided opportunities for the panelists to elaborate and discuss about 1) the definition of successful implementation, 2) importance of including the suggested factors, and 3) ways to improve descriptions for the suggested factors. Consensuses on these three
items were achieved through an iterative process of voting and discussion. In section two of the video conference, the panelists evaluated conditional independence of the potential factors. The panelists were told to assume that an organization had a successful implementation and that the organization was rated as having a high-level on a specific factor. They then discussed whether knowing this piece of information tells them a lot about how the organization might have responded to any of the other factors.\(^\text{17}\) If a factor violated the conditional independence, it was either rewritten or eliminated. This process was repeated for every suggested factor.

Refinements were made to the straw model in light of the discussion. The revised straw model was distributed among the panelists for final feedback. These procedures resulted in a final list of factors that were used to develop the surveys for the third round.

**Third round**

This survey consisted of two sections. In section one of this survey, the panelists estimated likelihood ratios of the final list of factors. The likelihood ratios are the weights of each identified factor in contributing to a successful implementation. Panelists were asked to assume that there are 100 hypothetical youth-serving organizations that had a successful implementation, and another 100 organizations had an unsuccessful implementation. They were told to distribute the 100 successful cases and the 100 unsuccessful cases among the three factor-level for each of the identified factors. A sample question is presented in Figure 1.1.
The likelihood ratios for each factor level were expressed in the ratios of conditional probability of observing the factor-level of a specific factor (a datum, D) given a successful implementation, to the conditional probability of that same datum given an unsuccessful implementation: \[ \frac{P(D|S)}{P(D|U)} \]. Using the example illustrates in Figure 1.1, the likelihood ratios for a factor called “implementer belief and motivation” would be 40/10=4/1 for the high factor-level, 30/30=1/1 for the moderate factor-level, and 30/60=1/2 for the low factor-level. Final likelihood ratios for each factor-level were obtained by averaging the individual estimates across the five panelists.

In section two, the panelists provided estimates for testing the internal validity of the predictive model. The predictive model would ideally be applied to predict a successful implementation in real cases, which is external validity. In the absence of a suitable empirical data base, however, we used experts’ opinions to generate a hypothetical data set for testing internal validity of the model. The panelists were asked to assume that a physical activity intervention was carried out in a sample of 60 youth-serving organizations. Then, they were provided with a set of computer-generated, hypothetical profiles reflecting how the 60 organizations rated on the factors identified in the second round. Every profile included all of the identified factors but with varying factor-level for each factor. Each panelist was asked to estimate how likely it was that organizations with a specific hypothetical profile would have a successful implementation, when taking all the identified factors into account at the same time. These estimates are called “holistic ratings.” The holistic ratings were estimated by using a 0-100
scale, where zero indicates absolutely no chance of having a successful implementation and 100 indicates 100% chance. A sample question is presented in Figure 1.1. A final group estimate for each profile was calculated by averaging the estimates across the five panelists.

Fourth round

Since the holistic ratings were used as a criterion for testing internal validity of the Bayesian predictive model, it was important to establish the reliability of these ratings. Two weeks after the third round, panelists completed a final online survey to re-rate 40 hypothetical profiles randomly selected from the original 60 profiles.

Analysis

Estimating posterior odds of success

Posterior odds of success were estimated for each of the 60 hypothetical profiles used in the third round. Posterior odds of success were calculated by multiplying the prior odds of success to the products of factor-level likelihood ratios. The prior odds of success are ratios of prior probability of successful implementation to probability of unsuccessful implementation: \([P(S)/P(U)]\). There are two types of priors: non-informative or informative. The informative priors are typically used when we have enough prior information about the estimating parameter, where the information would ideally be based on empirical studies. The non-informative priors are referred to as diffused priors. This type of prior is appropriate for estimating parameters that we may not have enough knowledge about its shape and scale of the distribution. Due to lack of previous research to guide the specification of an informative prior, this study assigned a non-
informative prior, which was 1/1 to the model. The factor-level likelihood ratios were obtained in round three of the modified Delphi process. An example on how one would use a three-factor Bayesian model to estimate posterior odds of successful implementation for a specific hypothetical profile is illustrated in Table 2.1. This example shows that an organization with a profile given in Figure 2.1 has about 57% chance of having a successful implementation.

Test-retest reliability

Intraclass correlations (ICC) with two-way random model were performed to examine test-retest reliability of holistic ratings obtained in the third and fourth rounds. ICC values of ≥0.75 indicates good reliability.

Internal validity

Pearson Product Moment Correlations were used to assess internal validity of the Bayesian model. The holistic ratings were correlated with model-derived posterior odds of success for each of the hypothetical profiles. A higher correlation value indicates better capability of the model in capturing the panelists' judgment.

Factor selection

First, a diagnostic power score was calculated for each factor to serve as a criterion for factor selection. The diagnostic power score refers to the range between the largest and the smallest likelihood ratio for that factor. If the highest and lowest likelihood ratios for a factor are 2.5/1 and 1/10, its diagnostic power would be 2.5+10=12.5. This score provides a crude measure regarding the amount of information that a certain factor can provide compared to other factors, with a larger value indicating a factor as more informative.
A backward factor selection procedure was used in attempt to reduce the final list of factors to those that are most important in predicting successful implementation. We started with a full model consisting of all factors identified in the third round and dropped one factor that had the lowest diagnostic power score at a time. A factor would be removed from the model if dropping it increased or did not reduce the internal validity. The procedures were repeated for every identified factor until internal validity had no further improvements. All data analyses were conducted using SPSS version 20.0 (IBM, Armonk, New York, USA).

Results

**Definition of successful implementation**

Based on the data collected during the first online survey and the video conference, the panelists indicated that successful implementation refers to “the intervention is carried out as planned as measured by fidelity to the protocol,” in which “protocol” refers to the quality elements specified by the intervention developer that are believed to be responsible for the intervention’s effects. Due to lack of consistent findings in the literature, the panelists decided not to determine a specific cut-point for fidelity. However, they indicated that researchers should explicitly define the quality elements and specify the criteria of successful implementation that are most relevant for their particular study.

**Factors identified by the expert panelists for predicting successful implementation**

The identified factors and their descriptions are presented in Table 2.2. The panelists identified 23 factors, containing 69 factor-levels, which are
important in predicting successful implementation. When categorized by the Durlak and DuPre ecological framework, seven factors were classified as organizational characteristics, including leadership motivation and engagement, physical activity culture, available space, available facilitates and equipment, available staff, communication, and competing program in the organization. Nine factors were categorized into implementation processes, including needs assessment, goal setting, engaging intervention staff, engaging youths, engaging program champion, training, technical assistance, reflecting and evaluating, and sustainability plans. Two factors were categorized as provider characteristics, including provider belief and motivation, and provider knowledge and skills. Three factors were related to program characteristics, including fun and inclusive design, empirical evidence, and adaptability. Finally, there were two community-level factors, including parental support for physical activity and competing programs in the community. Additionally, the expert panel also produced detailed descriptions for each of the identified factors at three levels of influence on successful implementation: high, moderate, and low.

The factor-level likelihood ratios and diagnostic power scores for the 23 identified factors are presented in Table 2.3. Factors with the highest diagnostic power scores were leadership motivation and engagement and engaging intervention staff; factors with the lowest scores were parental support for physical activity and competing programs in the community.
**Test-retest reliability**

Estimates of holistic ratings obtained in the third and fourth rounds were strongly correlated (ICC=0.88), indicating good reliability.

**Internal validity and factor selection**

The correlation between the posterior odds of success derived from the 23-factor full model and the holistic ratings was 0.65 (p<0.000), suggesting a moderate level of internal validity. With regard to factor selection, the backward selection procedures suggest that, among the comparison models, a final model that constituted of 15 factors yielded the highest internal validity (r=0.76, p<0.01.) The eight eliminated factors were: physical activity culture, communication, fun and inclusive design, empirical evidence, needs assessment, engaging youth, sustainability plans, and parental support for physical activity (Table 2.3.)

**Discussion**

This is the first study that utilized experts’ opinions to identify the influences of program implementation within the context of youth physical activity interventions. An accomplished panel of experts in implementation research of youth physical activity interventions engaged in a rigorous consensus process, which resulted in a carefully constructed definition of successful implementation, which is “the intervention was implemented as planned as measured by fidelity to the protocol.” This definition adds clarity to existing definitions. Previous studies generally define level of implementation as the extent to which the intervention is implemented as originally planned\textsuperscript{26} or the appropriate use of the intervention.\textsuperscript{27} However, the terms “planned” and “appropriate use” may still be too abstract.
Therefore, the current definition further elucidated “planned” refers to the protocol, which are the quality elements specified by the intervention developers.

During the first two rounds, the panelists identified 23 potential factors that are most important in achieving successful implementation of youth physical activity intervention. To further reduce the list of factors, experts’ ratings were translated into Bayesian predictive models. Among these models, a final model that retained 15 of the identified factors yielded the greatest contributions in predicting successful implementation of the hypothetical profiles. Although external validity remains to be established, this 15-factor final model had good internal validity.

When looking at the composition of the 15-factor final model, 80% of the factors concentrated on organizational characteristics and implementation processes, with the remaining factors distributed across the categories of provider characteristics, program characteristics, and community-level factors. This composition is consistent with the proposition of the Durlak and DuPre ecological framework that states that characteristics of the organization and implementation processes are central to program implementation, but successful implementation is dependent on factors from all five categories. This finding stresses the importance of balancing efforts to address the influencing factors at different ecological levels.

Additionally, it was expected that the 15-factor final model would be composed of factors with the highest diagnostic power because these factors are posited to be most informative in explaining variations in successful
implementation.\textsuperscript{15} However, this was not the case in the present study. The final model eliminated three major factors with relatively high diagnostic power (i.e., needs assessment, physical activity, and fun and inclusive intervention design), but retained two factors with lowest diagnostic power (i.e., competing programs within the organization and competing programs in the community.) A plausible explanation is that the two factors with lowest diagnostic power may have interacted synergistically with other factors in the model, thus outweighing the effects of the three major factors on successful implementation.\textsuperscript{14} These findings also indicate the importance of considering collective contributions rather than individual contributions of these factors in explaining successful implementation.

It is notable that the interrelationships among the identified factors will vary across interventions with different designs, at different implementation stages, and in different implementation settings.\textsuperscript{5, 10, 14, 28} Therefore, the list of factors identified in here is not intended to be a prescriptive formula. Rather, it is intended to provide researchers with a set of core elements that they can adopt, modify, and test.

\textit{Limitations}

Several limitations of this study should be considered. First, we are not able to incorporate the perspectives from the front-line staff that were responsible for day-to-day intervention operations, such as project coordinators or interventionists. We acknowledge that these individuals could provide valuable insights regarding factors that are most important in achieving successful implementation. Due to high turnover rate in this population, however, it was
difficult to recruit individuals who possessed optimal amounts and diversity of experiences enabling them to identify influences of program implementation that are commonly observed across multiple youth physical activity interventions. Second, the predictive model is considered preliminary because its external validity has not yet been established. However, previous studies\textsuperscript{15, 17-19} indicated that models developed through this systematic Bayesian approach could have good external validity. Third, as suggested in previous studies,\textsuperscript{15, 17, 18} the expert consensus process would ideally be condensed into a two-day intensive in-person meeting. To ensure optimal participation rate of the expert panelists, we employed a modified Delphi approach and the final model had good internal validity.

**Implications and future studies**

The current findings have immediate implications. First, the current study standardizes the terminology and descriptions for a set of factors that influences program implementation within the context of youth physical activity interventions. These standardized terminologies and descriptions can promote more consistent conceptualizations on these influencing factors in the field, thus allowing meaningful comparison across future studies. Second, the list of factors and their descriptions can be used by researchers and youth-serving organizations as a formative assessment tool to guide the planning and evaluation of their implementation efforts. For example, the factor “training” listed in Table 2.2 provides clear descriptions of what would a high-quality training session looks like, these could be used to guide the development of staff training, and those descriptions can be translated into a rubric for evaluation. Moreover,
researchers can adopt or modify the list of identified factors to suit their respective interventions.

This study is a step in the direction toward enhancing implementation of physical activity interventions carried out in youth-serving organizations. Future studies will continue to refine the descriptions of the identified factors and establish external validity of the predictive model. Our goal is to produce a valid diagnostic tool, with a set of well-defined factors accompanying a predictive model, which can be used by researchers and local staff in youth-serving organizations to systematically assess and identify factors that may assist or impede implementation before the intervention begins. With such information, necessary resources can be made available to the organizations and capacity building strategies can be tailored accordingly. Ultimately, better implementation planning may result in enhanced implementation fidelity, which may in turn improve program effectiveness.
A sample question for estimating likelihood ratios

Please distribute the 100 SUCCESSFUL cases and the 100 UNSUCCESSFUL cases among the different factor levels of “implementer beliefs and motivation” (total must sum to 100).

<table>
<thead>
<tr>
<th>have high levels of implementer beliefs and motivation</th>
<th>Of the 100 SUCCESSFUL cases, how many of them...</th>
<th>Of the 100 UNSUCCESSFUL cases, how many of them...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>have moderate levels of implementer beliefs and motivation</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>have low levels of implementer beliefs and motivation</td>
<td>30</td>
<td>60</td>
</tr>
</tbody>
</table>

A sample question for estimating holistic ratings

Below is a profile of a youth-serving organization. With the consideration of all the listed factors in total, please provide an overall estimate of the probability of a successful implementation on a scale of 0 to 100 (0=0% probability of success, 100=100% probability of success).

Profile 001

Factors with high ratings
- Leadership motivation & engagement
- Available staff
- Implementer belief and motivation
- Implementer knowledge & skills
- Empirical evidence
- Reflecting and evaluating

Factors with moderate ratings
- PA culture
- Available space
- Available facilities & equipment
- Competing organization programs
- Fun & inclusive design
- Adaptability
- Engaging program champion
- Training
- Technical assistance
- Parental PA support

Factors with low ratings
- Communication
- Needs assessment
- Goal settings
- Engaging youths
- Engaging staff
- Sustainability plans
- Competing community programs

Please click on the slider to provide your estimate:

0% probability of success 10 20 30 40 50 60 70 80 90 100% probability of success

Profile 001

Figure 2.1. Sample questions for the survey conducted in the third round
Table 2.1. An example of using a three-factor Bayesian predictive model to estimate posterior odds of success

<table>
<thead>
<tr>
<th>Factors</th>
<th>Factor-level</th>
<th>Factor-level likelihood ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative support</td>
<td>Moderate</td>
<td>1.5/1</td>
</tr>
<tr>
<td>Physical activity resources</td>
<td>High</td>
<td>2.07/1</td>
</tr>
<tr>
<td>Implementer enthusiasm</td>
<td>Low</td>
<td>1/2.33</td>
</tr>
</tbody>
</table>

Hypothetically, the expert panel suggested a three-factor model. Organization A rated at moderate level for administrative support, high level for physical activity resources, and low level for implementer enthusiasm. The mathematical form of the Bayesian model for predicting posterior odds of success is described as:

\[
\text{Prior odds of success} \times \text{products of factor-level likelihood ratios}
\]

Given a non-informative prior odds of success (1/1), the posterior odds of success for organization A is calculated as:

\[
1/1 \times 1.5/1 \times 2.07/1 \times 1/2.33 = 3.11/2.33 \text{ or } 1.33/1
\]

This is a 57% probability of successful implementation for organization A.
Table 2.2. Descriptions for the identified factors at the three factor-level: high, moderate, and low

<table>
<thead>
<tr>
<th>Factor</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
</table>
| 1. Leadership motivation and engagement | • Organization administrators are motivated to implement the intervention.  
• Organization administrators actively engage in:  
  ▶ Planning the intervention.  
  ▶ Participating in staff training.  
  ▶ Establishing policies (e.g., accountability system.)  
  ▶ Committing other staff to support intervention-related activities. | • Organization administrators are somewhat motivated to implement the intervention.  
• Organization administrators are supportive of:  
  ▶ Coordinating staff training.  
  ▶ Allocating resources.  
  ▶ Encouraging other staff to support intervention-related activities. | • Organization administrators are not motivated to implement the intervention.  
• Organization administrators are not involved and do not encourage other staff to support intervention-related activities. |
| 2. PA culture                 | • PA is central to the organization’s mission.  
• The organization currently offers PA programs. | • PA is not central to the organization’s mission.  
• The organization currently offers PA programs. | • PA is not central to the organization’s mission.  
• The organization currently does not offer PA programs. |
| 3. Available space            | • Adequate indoor and outdoor spaces are available for the intervention.  
• They are not taken away because of other reasons (e.g., inclement weather or organizational events). | • Adequate indoor and outdoor spaces are available.  
• They are occasionally not available because of other reasons. | • The organization has inadequate indoor and outdoor space.  
• If adequate, the space is often taken away because of other reasons. |
<table>
<thead>
<tr>
<th>Organizational characteristics</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Available facilities and equipment</td>
<td>• The organization has all the necessary facilities (e.g., basketball court) and equipment (e.g., basketball, hoops) for both children and implementers.</td>
<td>• The organization has some but not all the necessary facilities and equipment available for children and implementers.</td>
<td>• The organization has none of the necessary facilities and equipment for children and implementers.</td>
</tr>
<tr>
<td>5. Availabl staff</td>
<td>• Staff are experienced in delivering PA programs.</td>
<td>• Staff have some experience in delivering PA programs.</td>
<td>• Staff have no experience in delivering PA programs.</td>
</tr>
<tr>
<td></td>
<td>• Delivery of PA programs is part of their job description.</td>
<td>• Delivery of PA programs is part of their job description.</td>
<td>• Delivery of PA programs is not part of their job description.</td>
</tr>
<tr>
<td>6. Communication</td>
<td>• Their organization has effective communication mechanisms (formal and informal) to encourage frequent and open communication.</td>
<td>• The organization has acceptable communication mechanisms.</td>
<td>• The organization has poor communication mechanisms.</td>
</tr>
<tr>
<td>7. Competing programs in the organization</td>
<td>• The organization has no major competing programs or requirements of staff.</td>
<td>• The organization has implemented another PA program in the last year. The program was institutionalized in the organization over 6 months ago.</td>
<td>• The organization is required to focus on non-PA programs and recently reorganized its schedule accordingly.</td>
</tr>
</tbody>
</table>
### Table 2.2. (Continued)

<table>
<thead>
<tr>
<th>Implementation processes</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8. Needs assessment</strong></td>
<td>Qualitative data are collected with the target population to understand their desires and barriers, and the intervention is designed accordingly.</td>
<td>Either qualitative data are not collected with the target population to understand their desires and barriers, or the intervention is not designed accordingly.</td>
<td>Qualitative data are not collected with the target population to understand their desires and barriers.</td>
</tr>
<tr>
<td><strong>9. Goal setting</strong></td>
<td>The organization has clear short- and long-term intervention goals.</td>
<td>The organization has a general idea, but is not precise on short- and long term intervention goals.</td>
<td>The organization has no clear short- or long-term intervention goals.</td>
</tr>
<tr>
<td></td>
<td>The goals are clearly communicated and acted upon.</td>
<td>The goals are not clearly communicated to staff.</td>
<td>No one in the organization is aware of the goals.</td>
</tr>
<tr>
<td><strong>10. Engaging intervention staff</strong></td>
<td>Organization consults with and obtains consensus from staff when adopting the intervention.</td>
<td>The administrators who adopt the intervention notify the staff and give them opportunities to ask questions.</td>
<td>This is an organizational decision; the staff are notified.</td>
</tr>
<tr>
<td></td>
<td>Staff engage actively in developing the intervention and implementation protocols.</td>
<td>Staff are allowed some autonomy to modify the implementation protocol.</td>
<td>Staff are not involved in developing the intervention and implementation protocols.</td>
</tr>
<tr>
<td><strong>11. Engaging youths</strong></td>
<td>Organization engages the youths who are the target population in the adoption and development of the intervention (e.g., design, goal settings).</td>
<td>The organization somewhat involve the youths in the adoption and development of the intervention.</td>
<td>The organization does not involve the youths in any of the intervention process.</td>
</tr>
<tr>
<td>Implementation processes</td>
<td>High</td>
<td>Moderate</td>
<td>Low</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 12. Engaging program champion | • Individuals who are willing to invest and advocate for the intervention implementation dedicate themselves to serve in this role.  
• The program champion is well-respected by staff, youths and parents. | • The program champion is identified by the organization and the individuals agreed to participate.  
• Although not fully endorsed by staff, the program champion has adequate communication skills for parents or children. | • The program champion is not identified or used in the implementation of the intervention.  
• No program champion has been identified. |
| 13. Training              | • The intervention staff perceive the training as of good quality:  
  ❖ Adequate length and intensity.  
  ❖ Skills and content are provided in sequential steps (shaping procedures)  
  ❖ Adequate demonstrations on integrating intervention components into local contexts.  
  ❖ Adequate opportunities for the staff to try out, ask questions, and receive feedback. | • The intervention staff perceive the training as of moderate quality:  
  ❖ Acceptable length and intensity.  
  ❖ A few demonstrations show how to integrate intervention component into local context.  
  ❖ Opportunities are available for staff to ask questions. | • The intervention staff have no training or perceive the training as of poor quality:  
  ❖ Inadequate length and intensity.  
  ❖ Only general information about the intervention is provided. |
<table>
<thead>
<tr>
<th>Implementation processes</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Technical assistance</td>
<td>• Adequate follow-up visits are available.</td>
<td>• There were some follow-up visits available.</td>
<td>• Follow-up visits are not available.</td>
</tr>
<tr>
<td></td>
<td>• Efforts are made to support the intervention being implemented</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>beyond the intervention staff.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Reflecting and</td>
<td>• Organization has a quality improvement system to monitor the</td>
<td>• Organization has an evaluation system to monitor the implementation process and to provide performance feedback to staff.</td>
<td>• Organization has only a trust-based system and staff are not monitored or evaluated.</td>
</tr>
<tr>
<td>evaluating</td>
<td>implementation process.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Process data are utilized to provide performance feedback to</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>staff, inform strategies in making systematic, and continue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>implementation improvement.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Sustainability cells</td>
<td>• Organization has precise plans to sustain the intervention in an</td>
<td>• Some attentions are paid to intervention sustainability. However, the organization has no precise sustainability plans.</td>
<td>• There is no attention paid to intervention sustainability.</td>
</tr>
<tr>
<td></td>
<td>innovative and a fun way in the long run.</td>
<td></td>
<td>• Funding for sustaining the intervention is uncertain.</td>
</tr>
<tr>
<td></td>
<td>• Funding is available for sustaining the intervention.</td>
<td></td>
<td>• Funding is not available for sustaining the intervention.</td>
</tr>
</tbody>
</table>
Table 2.2. (Continued)

<table>
<thead>
<tr>
<th>Provider characteristics</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
</table>
| 17. Provider beliefs and motivation | • Intervention staff value child physical activity.  
• Staff are motivated to implement the intervention. | • Intervention staff accept the importance of child physical activity.  
• Staff are somewhat motivated to implement the intervention. | • Intervention staff prioritize non-PA related activities (e.g. homework, arts).  
• Staff are not motivated to implement the intervention. |
| 18. Provider knowledge and skills | • Intervention staff have all the necessary knowledge and skills to implement the intervention. | • Intervention staff have some of the necessary knowledge and skills to implement the intervention. | • Intervention staff have none of the necessary knowledge and skills to implement the intervention. |
| 19. Fun and inclusive design | • The intervention is fun and appeals to children of different ages and genders, and with different interests and levels of physical competence. | • The intervention is fun, but appeals only to a certain group of children. | • The intervention is not fun and appeals only to a limited number of children with specific characteristics (e.g., age, gender, or physical competence levels). |
Table 2.2. (Continued)

<table>
<thead>
<tr>
<th>Program characteristics</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
</table>
| 20. Empirical evidence   | • The intervention protocol and/or materials are tested and determined to be efficacious with youths similar to those in the intervention.  
                          | • The intervention has been found to be feasible in a variety of settings by diverse implementers. | • Part of the intervention protocol and/or materials are tested and determined to be efficacious with youths similar to those in the intervention.  
                          |                                                                       | • The intervention is feasible in a few settings. | • The intervention protocol and/or materials are not tested. |
| 21. Adaptability         | • Specific guidelines and materials are available to guide effective adaptation. | • General guidelines are available for effective adaptation. | • No guideline or materials are available to guide effective adaptation. |
| 22. Parental support for PA | • Parents value physical activity.  
                          | • They are willing to support increasing PA levels for their child. | • Parents think that PA programs could be beneficial to their child.  
                          |                                                                       | • They are somewhat supportive to increasing PA levels for their child. | • Parents prioritize non-PA events (e.g., homework, arts).  
                          |                                                                       |                                                                       | • They do support increasing PA levels for their child. |
| 23. Competing programs in the community | • Competing programs are not available in the community. | • The community has a few programs that are similar to the intervention. | • Several competing programs are offered in the community. |

35
<table>
<thead>
<tr>
<th>Factors</th>
<th>Likelihood ratios for each factor-level</th>
<th>Diagnostic power scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>Organizational characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership motivation and engagement</td>
<td>8.75/1</td>
<td>1/1.05</td>
</tr>
<tr>
<td>Physical activity culture</td>
<td>4.69/1</td>
<td>1.07/1</td>
</tr>
<tr>
<td>Available facilitates and equipment</td>
<td>4/1</td>
<td>1/1.12</td>
</tr>
<tr>
<td>Available space</td>
<td>3.05/1</td>
<td>1.19/1</td>
</tr>
<tr>
<td>Available staff</td>
<td>5/1</td>
<td>1/1.17</td>
</tr>
<tr>
<td>Communication</td>
<td>2.55/1</td>
<td>1.05/1</td>
</tr>
<tr>
<td>Competing programs within the organization</td>
<td>1.77/1</td>
<td>1.19/1</td>
</tr>
<tr>
<td><strong>Implementation processes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needs assessment</td>
<td>7.75/1</td>
<td>1/1.23</td>
</tr>
<tr>
<td>Goal setting</td>
<td>4.31/1</td>
<td>1.12/1</td>
</tr>
<tr>
<td>Engaging intervention staff</td>
<td>5.57/1</td>
<td>1.08/1</td>
</tr>
<tr>
<td>Engaging youths</td>
<td>3.16/1</td>
<td>1.17/1</td>
</tr>
<tr>
<td>Engaging program champion</td>
<td>4.43/1</td>
<td>1/1.15</td>
</tr>
<tr>
<td>Training</td>
<td>3.39/1</td>
<td>1/1</td>
</tr>
<tr>
<td>Technical assistance</td>
<td>3.85/1</td>
<td>1.27/1</td>
</tr>
<tr>
<td>Reflecting and evaluating</td>
<td>4.92/1</td>
<td>1.44/1</td>
</tr>
<tr>
<td>Sustainability plans</td>
<td>2.23/1</td>
<td>1.36/1</td>
</tr>
<tr>
<td><strong>Provider characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provider belief and motivation</td>
<td>4.54/1</td>
<td>1.13/1</td>
</tr>
<tr>
<td>Provider knowledge and skills about the intervention</td>
<td>5.16/1</td>
<td>1.03/1</td>
</tr>
<tr>
<td><strong>Program characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fun and inclusive design</td>
<td>4.47/1</td>
<td>1.10/1</td>
</tr>
<tr>
<td>Empirical evidence</td>
<td>2.67/1</td>
<td>1/1</td>
</tr>
<tr>
<td>Adaptability</td>
<td>4.54/1</td>
<td>1/1.19</td>
</tr>
<tr>
<td><strong>Community-level factor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competing programs in the community</td>
<td>1.96/1</td>
<td>1.13/1</td>
</tr>
<tr>
<td>Parental support for physical activity</td>
<td>2.15/1</td>
<td>1.06/1</td>
</tr>
</tbody>
</table>

Note: Diagnostic power refers to the range between the largest and smallest likelihood ratios for a specific factor. For example, the largest and smallest likelihood ratios for “physical activity culture” are 4.69/1 and 1/6.56, respectively. The diagnostic power for this factor is calculated as 4.69+6.56=11.25. Boldface indicates factors retained in the final model.
References


CHAPTER 3

MANUSCRIPT 2: FACTORS INFLUENCING IMPLEMENTATION OF A PRESCHOOL-BASED PHYSICAL ACTIVITY INTERVENTION\textsuperscript{2}

\textsuperscript{2} Lau EY, Saunders RP, Beets MW, Wandersman A, Cai B, Pate RR. To be submitted to Health Education Research.
Abstract

**Background:** The Study of Health and Activity in Preschool Environments (SHAPES) was a 3-year multi-component randomized trial designed to increase physical activity in preschoolers. The purpose of this study was to examine the factors that influence implementation of the Move Outside component of SHAPES.

**Methods:** This study analyzed process evaluation data from preschool classrooms that participated in year 1 and year 2 of SHAPES. Implementation of the Move Outside component was assessed by direct observations. The three selected factors, preschool characteristics, teacher characteristics, and quality of implementation processes, were measured by direct observations, interviews, evaluation forms, and surveys. Bayesian path analyses were used to test the hypothesized direct and indirect effects between the three selected factors and level of implementation.

**Results:** Level of implementation across preschool classrooms ranged from 0% to 186% (M=62.9%, SD=55.4%) in year 1 and from 0% to 185% (M=56.7%, SD=40.6%) in year 2. The selected factors were not significantly associated with level of implementation in year 1. Preschool characteristics were found to be directly associated with level of implementation ($\beta=0.528$, 95% CI: 0.134, 0.827) in year 2.

**Conclusion:** The present findings provide preliminary evidence suggesting that factors that influence level of implementation may differ depending on stages of implementation.
**Introduction**

Physical activity is important to the growth and development of preschool-age children (ages 3-5 years).\(^1\) National guidelines\(^2\) on physical activity recommend that children in this age group engage in at least 60 minutes of structured moderate-to-vigorous physical activity (MVPA) per day and an additional 60 minutes of unstructured MVPA per day. However, studies that used objective measures of physical activity consistently showed that most preschool-age children do not meet these guidelines.\(^3\)\(^-\)\(^6\) These data indicate the need for developing effective interventions to increase physical activity among children in this age group.

Given that over 60% of American children ages 3 to 5 years are enrolling in center-based preschools,\(^7\) preschools offer a potential point of intervention. To date, a limited number of randomized controlled trials (RCTs) have been conducted to test the effectiveness of preschool-based physical activity interventions,\(^8\)\(^-\)\(^10\) and several effective programs were identified.\(^11\)\(^-\)\(^14\) However, the identification of evidence-based interventions is only the first step. For such interventions to have broad and lasting impact on the population prevalence of physical activity, widespread dissemination is needed.\(^15\)\(^,\)\(^16\) To be recognized as ready for broad dissemination, a program must identify the key components that determine its success.\(^17\) Moreover, it is essential to examine the factors that may potentially influence the implementation of such components, so that relevant implementation strategies can be developed accordingly.\(^18\)\(^,\)\(^19\)
The Study of Health and Activity in Preschool Environments (SHAPES) intervention was a 3-year RCT that was found to be effective in increasing physical activity in preschool-age children. A complete description of the outcome evaluation is currently under development (R. R. Pate, unpublished data, 2015.) SHAPES consisted of three main components that aimed at increasing preschoolers' MVPA by offering physical activity opportunities through indoor playtime (Move Inside), recess (Move Outside), and active learning (Move To Learn). An ongoing process evaluation analysis (R.P. Saunders, unpublished data, 2015) showed that implementation of the Move Outside component was significantly associated with preschool day MVPA over the intervention period in girls. Girls who attended classrooms classified as high-implementers of the Move Outside component engaged in significantly more MVPA than girls in low-implementer classrooms or in the control classrooms; no difference was found between the low-implementer and the control classrooms. These findings indicate that improving implementation of the Move Outside component may further increase the effectiveness of SHAPES in future dissemination efforts. However, factors that influence the implementation of this important component have not been identified.

Therefore, the purpose of the current study was to examine the effects of preschool characteristics, teacher characteristics, and quality of implementation processes on implementation of the MO component of SHAPES.
Methods

Conceptual model of the present study

This study was guided by the ecological framework developed by Durlak and DuPre,\textsuperscript{19} which has been used to study the influences of implementation of health promotion interventions\textsuperscript{19} and school-based physical activity interventions.\textsuperscript{21} The framework posits that level of intervention implementation is influenced by five types of factors, including organizational characteristics, implementation processes (e.g., training), provider characteristics, program characteristics, and community-level factors, such as politics. The framework also suggests that these factors may interact to influence implementation.

In the current study, ten elements referenced in previous studies\textsuperscript{19,21} were included to predict level of implementation of Move Outside component. Given the lack of research into the pathways through which these elements interact to influence implementation outcomes, the ten elements were grouped into three reasonably independent factors based on the Durlak and DuPre framework: preschool characteristics, teacher characteristics, and quality of implementation processes. Then, a model was proposed to test how these three factors interact to influence level of implementation (Figure 3.1). It was hypothesized that the three selected factors will be directly related to level of implementation. As there is limited empirical evidence suggesting that characteristics of the provider may mediate the influences of the other factors,\textsuperscript{22-26} we further hypothesized that teacher characteristics will mediate the influence of the other two factors.
**Study Design**

A prospective observational study design was used to address the aim of the present study. Data were taken from the process evaluation of SHAPES collected over the 3-year intervention period. These included data related to implementation of the Move Outside component assessed during each intervention year, and preschool characteristics, teacher characteristics, and quality of implementation processes measured at baseline and annual process evaluation assessments.

In the current study, year 3 process data were excluded from the analyses due to the large amount of missing data in preschool characteristics. Due to changes in process evaluation methodology and teacher turnover, we treated the data collected from year 1 and year 2 as two cross-sectional samples and
analyzed the data by year. The current study was approved by the University of South Carolina’s Institutional Review Board.

**Participants**

The participants were preschool classrooms randomized into the intervention group that had implemented SHAPES and had complete process evaluation data. For year 1, there were 20 intervention classrooms, 19 of which had complete process evaluation data. For year 2, there were 17 intervention classrooms and all had complete process evaluation data. These classrooms were nested in eight intervention preschools and the characteristics of these preschools varied, with the number of enrolled students ranging from 199 to 870. Fifty percent of the preschools were public schools. Thirty-eight percent of the preschools predominantly served Caucasian children, 38% served predominately African American children, and 24% had an equal distribution in race/ethnicity. Sixty-three percent of them offered full time programs, and 75% of them provided physical education.

Characteristics of the intervention classrooms did not differ between the two years, with the number of children per classroom ranging from 14 to 20. All classrooms had a female teacher. For year 2, 14 classrooms were led by teachers that had implemented SHAPES in year 1 and three classrooms were led by new teachers.

**Overview of SHAPES-Move Outside**

Based on the evidence that children are more likely to be active when they are outdoors, the Move Outside component was designed to increase
opportunities for outdoor physical activity through outdoor recess. The goal was for teachers to provide two 20-minute outdoor recesses per day whenever possible. Each session should include at least one 5-minute teacher-led physical activity. If weather was not conducive for outdoor recess, teachers were encouraged to provide an indoor recess of equal duration.$^{20}$

To assist teachers in achieving intervention goals, training and ongoing technical assistance (i.e., workshops, site visits, and newsletters) were provided to increase teachers' knowledge, skills, and confidence to implement SHAPES. Pfeiffer et al$^{20}$ provide detailed descriptions of the multi-component intervention.

**Measures**

Items used to assess study variables in the current study were selected from multiple instruments used in the SHAPES process evaluation, including survey, interview, and evaluation form. A complete description of the process evaluation is currently under development.

**Level of implementation**

The current study defined level of implementation as the extent to which the Move Outside component was delivered as planned (i.e., fidelity). The process evaluation methodology differed between year 1 and year 2 due to resources constraints. In year 1, observations on the intervention implementation were conducted on four fall days and four spring days, and observers sampled children from multiple classrooms in each observation day. For year 2, observations were conducted on one fall day and one spring day, and observers sampled children from a single classroom on each day.
During each observation session, the evaluator recorded the total minutes of physical activity opportunities provided through the Move Outside component. For each year, mean daily minutes of physical activity opportunity observed through Move Outside was calculated across the observation days. Level of implementation was calculated as observed daily total minutes of physical activity opportunities provided through Move Outside divided by the prescribed daily minutes for the Move Outside component (i.e., 40 minutes recess time per day for full-day programs or 20 minutes recess time per day for half-day programs). A higher score indicates a higher adherence to the planned Move Outside component.

**Preschool characteristics**

Seven elements of preschool characteristics were assessed, including preschools’ physical activity policies and practices, structural characteristics, organizational climates, physical activity resources, organizational functioning, leadership support, and community connections.

At baseline, the preschool directors completed five items regarding preschools’ physical activity policies and practices, such as time for active free play. Responses were recorded on a four-point Likert-scale. The directors also completed three items related to structural characteristics, including the number of children served, teachers’ education levels, teachers’ training on physical activity- or exercise-related aspects in the past year.

The preschool directors also completed 13 items taken from the Early Childhood Environment Rating Scale, Revised Edition (ECERS-R)\textsuperscript{28} regarding
preschools’ organizational climates. The ECERS-R is a standardized rating scale that has been widely used to evaluate the resources and quality of early childhood education programs.\textsuperscript{29-32} The full ECERS-R scale consisted of 43 seven-point Likert-type items, and the scale has been demonstrated to be reliable at the individual item and total scale score levels previously.\textsuperscript{28,29} Ten items were used to assess preschools’ physical activity resources. The preschool directors responded to eight items taken from the ECERS-R regarding space, equipment, and scheduling for physical activity. The intervention staff responded to two items assessing the classroom sizes (square feet) and playground sizes (square feet).

During the end of each year, interventionists completed one item evaluating preschools’ functioning of each classroom. Each classroom teacher completed one item assessing their perceived leadership support, with response options ranging from 1 “not inadequate” to 4 “very adequate.” Community connection, defined as parental support for physical activity and coordination with community agencies, was also measured at baseline. The observer completed one item from the ECERS-R to rate parental involvement in school activities. Additionally, the preschool directors responded to one item related to the type of community program/activity provided on preschool grounds. Responses to the items were summed to create a composite score for preschool characteristics. A higher score indicated a more supportive preschool environment for implementing physical activity interventions.
**Teacher characteristics**

Two elements of teacher characteristics were assessed, including self-efficacy and skills proficiency. At the end of each year, teachers responded to one item regarding their self-efficacy of implementing the intervention on a four-point Likert scale, with the endpoint ranging from “no effect” to “a big effect” or “very unprepared” to “very prepared.” The interventionists completed one item evaluating teachers’ skills in resolving implementation issues, with response options ranging from 1 “poor” to 4 “strong.” Responses to each item were summed to create a composite score for teacher characteristics, with higher scores indicating teacher characteristics as being more favorable to a successful implementation.

**Quality of implementation processes**

Quality of implementation processes was measured, defined as quality of training and technical assistance provided by the SHAPES staff. At the end of each intervention year, each classroom teacher completed two items measuring the quality of the training and technical assistance as reflected by perceived ease of implementation the Move Outside component and adequacy of support received from the SHAPES staff. Responses were recorded on a four-point Likert scale, with 1 indicating “very inadequate” and 4 indicating “very adequate.” A composite score for quality of implementation processes was calculating by adding the two items, with higher scores indicating the implementation processes being perceived as more favorable to a successful implementation.
Analysis

Bayesian path analysis was conducted to assess the direct and indirect effects of preschool characteristics, teacher characteristics, and quality of implementation processes on level of implementation. The Bayesian approach was selected because it is more appropriate for modeling data based on a small sample, which is often the case in process evaluation data.

In the current study, a prior hypothesized path model was tested (see Figure 3.1) using M-plus software (version 6.11) with the Bayesian estimation method. In the path diagram, preschool characteristics and quality of implementation processes are the predictor variables, teacher characteristics is the mediator, and level of implementation is the outcome variable. Following Wang and Preacher’s conceptual approach to describing mediation, path $a$ and path $z$ are the relationships between the predictor variables and the mediator variable. Path $b$ is the direct effect of the mediator on the outcome variable. Path $c'$ and $e'$ are the direct effects of the predictor on the outcome variable after controlling for the mediator.

The average indirect effects were calculated as $a*b$ (preschool characteristics $\rightarrow$ teacher characteristics $\rightarrow$ level of implementation) and $z*b$ (quality of implementation processes $\rightarrow$ teacher characteristics $\rightarrow$ level of implementation). The total indirect effect (c) was calculated as $a*b + z*b$. The estimated direct and indirect effect would be determined to be significant if the 95% CI did not include zero.

To apply the Bayesian approach, we need to specify a model and assign
prior distribution to all unknown parameters in the model. The mathematical form of the model is given as follows:

\[ M = d_M + aX_{1i} + zX_{2i} + e_m \]

\[ Y = d_Y + c'X_{1i} + e'X_{2i} + bM_i + e_Y \]

In the model, \( d \) represents the intercept and \( e \) represents error. The term \( Y \) denotes the outcome variable (IMTLEVEL); \( M \) is the mediating variable (TEACHER); and \( X_1 \) and \( X_2 \) are the independent variables (PRESCH and PROCESS). The term \( e_{mi} \) and \( e_{yi} \) are the residuals of \( M \) and \( Y \); the parameter \( d_M \) and \( d_Y \) are the intercepts, and \( a, b, c', z, \) and \( e' \) are slopes. As accurate prior information is not available, all the unknown parameters were assigned independent non-informative uniform priors. The regression coefficients \( \beta = (a, b, c', z, e', d_M, d_Y)' \) were assigned to follow a normal distribution, and the variance parameters \( \sigma^2 = (\sigma^2_{eM}, \sigma^2_{eY})' \) follow an inverse-gamma distribution. As suggested in previous studies, the prior distributions can be specified as follows:

\[ \beta_p \sim N (0, 1.0+6E) \]

\[ \sigma_i^2 \sim IG (0.0001, 0.0001) \]

where \( \beta_p \) represents the \( p \)th element of \( \beta \) and \( \sigma^2_i \) represents the \( l \)th elements of \( \sigma^2 \). A large variance in the normal prior above implies a non-informative prior. Similarly, the small hyper-parameters in the inverse prior provide a diffuse prior for \( \sigma^2_i \).

Bayesian estimates of all parameters and variance components in the framework were calculated based on 10,000 samples after 1000 burn-in
iterations. Posterior mean, posterior standard error and 95% credibility interval (CI) at the 0.025 and 0.975 quantiles of the average direct and indirect effects were also obtained. The convergence of the final model was assessed by multiple indices. The Proportional Scale Reduction (PSR) with a value equal to 1 indicates convergence. The second index is the stability of trace plots for the posterior samples of the parameters with a tight and horizontal shape suggesting reliable estimations of the parameters. The third index is the autocorrelation plot with a small value (≤0.1) indicating a good model convergence.36

Model-to-data fit was evaluated based on the 95% confidence interval for the difference between the observed and the replicated chi-square values (95% CI for chi-square value) and the posterior predictive p-value (PPP). Both indices evaluated the discrepancy between the observed data and data generated by the model, with a smaller discrepancy indicating a good model-data fit. A lower negative value of 95% CI for chi-square value and a PPP value close to 0.5 are preferred.

Results

Descriptive statistics of the study variables are presented in Table 3.1. Level of implementation across preschool classrooms ranged from 0% to 186% (M=62.9%, SD=55.4%) for year 1 and from 0% to 185% (M=56.7%, SD=40.6%) for year 2. The hypothesized models for both years had a good convergence as the PSR values were close to 1 for all the estimated parameters, the trace plots illustrated a tight and horizontal shape, and the autocorrelation plots showed all parameters had value ≤0.1. The models also provided a good fit as the lower
bound of the 95% CI for chi-square value achieved a negative value and the PPP value was close to 0.5 (Table 3.2).

The final path diagram with standardized path coefficients is presented in Figure 3.2. The hypothesized models explained 15% and 53% of the variance in level of implementation for year 1 and year 2, respectively. For year 1, none of the selected factors had significantly direct or indirect associations with level of implementation. The only significant association was the direct effect of preschool characteristics on teacher characteristics ($\beta=0.797$, 95% CI: 0.588, 0.951.) For year 2, preschool characteristics were found to be directly associated with level of implementation ($\beta=0.528$, 95% CI: 0.134, 0.827) (Table 3.2). The result is interpreted as follows: for every one standard deviation unit increase in the score of preschool characteristics, level of implementation increases by 0.53 standard deviation units.

**Discussion**

This study examined the direct and indirect effects of preschool characteristics, teacher characteristics, and quality of implementation processes on implementation of the Move Outside component during year 1 and year 2 of the SHAPES intervention. Due to differences in process evaluation methodology, the two years were treated as two cross-sectional samples and analyzed separately.

The present findings show that the pattern of associations between the selected factors and level of implementation differed by implementation stages. Year 1 of SHAPES was considered the initial implementation stage, where the intervention was being used the first time. During this stage, the three selected
factors collectively explained only 15% of the variance in level of implementation and none of the selected factors had significant associations with level of implementation. This finding suggests that level of implementation during year 1 is mainly explained by other unmeasured factors. There is one study\(^{37}\) suggesting that factors that influence adoption of the intervention may also be influential to implementation in initial stage. Future studies should continue to explore other factors that specifically influence implementation in the initial stage.

During year 2, SHAPES was in a transition from initial implementation stage to full implementation stage. As published previously,\(^ {38}\) modifications were made to the Move Outside component after year 1 based on teachers’ and interventionists’ feedback. Such adaptations resulted in enhanced acceptability and feasibility of SHAPES; in turn SHAPES progressed into a fully operational program in year 2. The present findings show that, during this stage, the three selected factors collectively explained more than 50% of the variance in level of implementation, with preschool characteristics being significantly and directly associated with level of implementation. This finding provides empirical evidence supporting the positive association between organizational characteristics and implementation suggested in existing conceptual frameworks.

Quality of implementation processes and teacher characteristics were neither directly nor indirectly associated with level of implementation. The lack of significant association may be a result of lack of variability in these two variables, with most of the participating classrooms have high ratings for both variables. Although the lack of variability in these two variables might be due to insufficient
sensitivity of the instrument in recognizing variability at the upper end, it is more likely that the consistently high ratings accurately reflect the positive impact of the intervention (i.e., training and on-going technical assistance) on teachers’ perceptions of the programs, self-efficacy, and skills.

Strengths of the current study include objective measures of level of implementation and comprehensive assessment of the preschools’ characteristics. Moreover, the use of the Bayesian estimation method provided the statistical power to test the proposed relationships that were usually lacking in implementation research. However, there are a number of limitations of the present study that warrant further explorations in future research. Findings from the path analysis can only disprove our hypothesized causal relations among variables, but it cannot prove causality. Future studies should employ experimental study design to examine whether manipulating the factors identified in the current study is associated with implementation outcomes. With a small sample, the stability of the path model should be viewed with caution; this study should be replicated in a larger sample. Although the participating classrooms are nested in eight preschools, we were unable to analyze the data with a multilevel model due to limited statistical power associated with the number of preschools and the number of classrooms per preschool (ranging from 1 to 4). Also, the level of implementation of classrooms within the same preschool was highly variable. These variations suggest that modeling the data at the classroom level may be more appropriate than aggregating the data to a higher order.
Due to difference in process evaluation methodology, data collected in the two years were treated as two cross-sectional samples. This prevents the examination of longitudinal associations between the selected factors measured in year 1 and level of implementation in year 2. Given the potential of longitudinal associations,\textsuperscript{37} future studies should examine how prior levels of the influencing factors affect level of implementation in the later stage of implementation. Additionally, some potential determinants of the level of implementation are likely to be omitted in this study although multiple factors were measured. However, our study provides a set of core constructs and measures specific to preschool-based physical activity interventions that future studies may adopt and test. Future studies should take the current findings into consideration and continue to explore other contextual factors that influence the level of implementation, especially for the initial stage of implementation.

In conclusion, the present study shows that the three selected factors, preschool characteristics, teacher characteristics, and quality of implementation processes, were not significantly associated with level of implementation in year 1. However, preschool characteristics were found to be significantly associated with level of implementation in year 2. These findings provide preliminary evidence suggesting that factors that influence level of implementation may differ depending on stages of implementation.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Year 1 (n=19)</th>
<th>Year 2 (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Preschool characteristics</td>
<td>154.61 (10.82)</td>
<td>134.8-163.85</td>
</tr>
<tr>
<td>Teacher characteristics</td>
<td>7.35 (1.53)</td>
<td>3.0-8.85</td>
</tr>
<tr>
<td>Quality of implementation processes</td>
<td>15.26 (1.19)</td>
<td>12.0-16.0</td>
</tr>
<tr>
<td>Level of implementation</td>
<td>62.94 (55.39)</td>
<td>0-186.9</td>
</tr>
</tbody>
</table>
### Table 3.2. Direct and indirect effects of selected factors on level of implementation of SHAPES-Move Outside

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th></th>
<th>Year 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimates (SD)</td>
<td>95% CI</td>
<td>Estimates (SD)</td>
<td>95% CI</td>
</tr>
<tr>
<td><strong>Direct effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRESCH → IMLEVEL <em>(path c')</em></td>
<td>0.013 (0.410)</td>
<td>-0.795 to 0.833</td>
<td><strong>0.528 (0.178)</strong></td>
<td><strong>0.135 to 0.827</strong></td>
</tr>
<tr>
<td>TEACHER → IMLEVEL <em>(path b)</em></td>
<td>0.105 (0.497)</td>
<td>-0.883 to 1.075</td>
<td>0.153 (0.278)</td>
<td>-0.402 to 0.689</td>
</tr>
<tr>
<td>PROCESS → IMLEVEL <em>(path e')</em></td>
<td>-0.089 (0.288)</td>
<td>-0.620 to 0.508</td>
<td>-0.372 (0.204)</td>
<td>-0.784 to 0.029</td>
</tr>
<tr>
<td>PRESCH → TEACHER <em>(path a)</em></td>
<td><strong>0.797 (0.096)</strong></td>
<td><strong>0.588 to 0.951</strong></td>
<td>0.141 (0.211)</td>
<td>-0.291 to 0.541</td>
</tr>
<tr>
<td>PROCESS → TEACHER <em>(path d)</em></td>
<td>0.251 (0.141)</td>
<td>-0.036 to 0.529</td>
<td>0.490 (0.255)</td>
<td>-0.130 to 0.862</td>
</tr>
<tr>
<td><strong>Indirect effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRESCH → TEACHER → IMLEVEL <em>(path a</em>b)*</td>
<td>0.083 (0.411)</td>
<td>-0.748 to 0.884</td>
<td>0.024 (0.094)</td>
<td>-0.148 to 0.243</td>
</tr>
<tr>
<td>PROCESS → TEACHER → IMLEVEL <em>(path d</em>b)*</td>
<td>0.023 (0.124)</td>
<td>-0.230 to 0.285</td>
<td>0.078 (0.166)</td>
<td>-0.232 to 0.477</td>
</tr>
</tbody>
</table>

**Model fit**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>95% CI for difference between observed and replicated chi-square values</td>
<td>-16.360 to 17.844</td>
<td>-15.843 to 17.887</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior Predictive P-value</td>
<td>0.470</td>
<td>0.454</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Boldfaced indicates a significant effect with 95% CI not including zero. Estimates: average of the posterior means, SD: average of the posterior standard deviations, 95% CI: lower and upper bounds of the 95% credibility interval. IMLEVEL=implementation levels; PROCESS=quality of implementation processes; PRESCH=preschool characteristics; PROCESS =quality of implementation processes, SHAPES=Study of Health and Activity in Preschool Environments; TEACHER=teacher characteristics.
**Figure 3.2.** Final path models illustrating factors that influence implementation of the Move Outside component during year 1 and year 2 of the SHAPES intervention.

Bolded solid path indicates significant paths. Standardized path coefficients are shown. Asterisk indicates significant association with 95% CI not including zero. For year 1, preschool characteristics had significant direct associations with teacher characteristics. For year 2, preschool characteristics had significant direct associations with level of implementation.
References


CHAPTER 4

MANUSCRIPT 3: FACTORS INFLUENCING IMPLEMENTATION OF A PHYSICAL ACTIVITY INTERVENTION IN RESIDENTIAL CHILDREN’S HOMES

3 Lau EY, Saunders RP, Beets MW, Wandersman A, Cai B, Pate RR. To be submitted to Health Promotion Practice.
Abstract

**Background:** The Environmental Intervention in Children’s Homes (ENRICH) study was the first and only published physical activity intervention undertaken in residential children’s homes (RCHs). The study revealed that differential implementation across sites appears to be one of the key players that affect program effectiveness. The purpose of this study was to examine the direct and indirect effects of RCH characteristics, wellness team characteristics, and quality of implementation processes on level of implementation of the ENRICH intervention.

**Methods:** This study analyzed the ENRICH process evaluation data collected from 29 RCHs. Bayesian Path analysis was used to examine the direct and indirect effects of RCH characteristics, wellness team characteristics, and quality of implementation processes on level of implementation.

**Results:** Level of implementation across RCHs was variable, ranging from 38% to 97% (M=68.3, SD=14.45). Results revealed that RCH characteristics and wellness team characteristics had significant direct association with level of implementation. Neither direct nor indirect associations between quality of implementation processes and level of implementation reached statistical significance.

**Conclusion:** Organizational contexts and providers’ attitudes and skills played an important role in influencing implementation of the ENRICH intervention. Incorporating information about both factors in implementation planning may increase the likelihood of achieving higher levels of implementation in future studies.
Introduction

Obesity, once a rare condition in foster children, is now recognized as one of the primary medical problems among children in the foster care system. A study involving 6,177 children who entered foster care in Utah found that 35% of them were overweight or obese, which was higher than the National Health and Nutrition Examination Survey (NHANES) prevalence (31.8%). In another study of 2,078 Hispanic children who entered foster care in Los Angeles, nearly 40% of the sample were overweight or obese, which was also slightly higher than the national prevalence for Hispanic children (38.9%). These data suggest that there is an urgent need to address the obesity problem in this population.

There is consistent evidence indicating that low physical activity is associated with the development of excessive fatness in children. Foster children are at higher risk of being underactive than the general population, because many of them have physical and mental health conditions. These conditions may have limited the children’s ability or willingness to participate in physical activity, in turn increasing their risk of obesity. Hence, promoting a more physically active lifestyle may be an important strategy in addressing the obesity problem in foster children.

Research on promoting physical activity among foster children is in its infancy. One study conducted in residential children’s homes (RCHs) found that several features in the homes are positively correlated with children’s physical activity, including the presence of a recreation director for physical activity
programming, availability of physical activity opportunities, quality of the physical activity opportunities, and accessibility of physical activity resources. These findings suggest that the RCH settings may be an alternative point of intervention.

To date, the Environmental Intervention in Children’s Homes (ENRICH) study was the first and only published randomized controlled trial designed to test the effectiveness of an environmental intervention in increasing physical activity levels among children residing in RCHs.\(^9,10\) A primary component of ENRICH was utilizing a wellness team formed by RCH adult staff as organizational change agents to create an RCH environment that supports physical activity. Outcome evaluation of the ENRICH intervention showed that there were no statistically significant differences in children’s physical activity levels between the intervention group and the control group.\(^9\)

However, process evaluation indicates that the lack of between-group differences in ENRICH is likely to be partially explained by secular trends, in which some homes in the control group implemented unexpected improvements in selected environmental features promoted in the ENRICH framework. These secular trends reduced the magnitude of the differences between the intervention and control groups, thus masking intervention effects.\(^11\) Furthermore, process data revealed that about 40% of the RCHs in the intervention group did not implement the intervention at intended levels,\(^9,10\) and that may have further diluted the intervention effects. It is possible that positive changes in the
intervention homes could have outpaced the secular trends if more intervention homes had carried out the intervention at optimal levels.

Although we previously found that level of implementation was not significantly associated with program outcomes in ENRICH, it is likely that the results were confounded by the secular trend. It is possible that positive changes in the intervention homes could have outpaced the secular trends if more intervention homes had carried out the intervention at optimal levels. Also, it is possible that level of implementation influenced the ENRICH outcome through indirect pathways that were not examined in the previous analysis.

Furthermore, given the lack of research in RCH-based physical activity interventions, identifying the factors that contributed to differential implementation of ENRICH will provide valuable information to inform implementation planning in future studies, which may increase their chance of attaining desired program outcomes. Therefore, the purpose of this study was to examine the effects of RCH characteristics, wellness team characteristics, and quality of implementation processes on level of implementation of the ENRICH intervention.

Methods

Conceptual model of the present study

This study was guided by the ecological framework proposed by Durlak and DuPre, which has been used to study the influences of implementation of health promotion interventions and school-based physical activity interventions. The framework posits that level of intervention implementation is influenced by factors present in five categories, including organizational
characteristics, implementation processes (e.g., training), provider characteristics, program characteristics, and community-level factors, such as politics. The framework also suggests that elements in the five categories may interact to influence implementation.

In the current study, ten elements referenced in previous studies\textsuperscript{13, 14} were included to predict level of implementation of ENRICH. As there is a dearth of research exploring the pathways through which these elements interact to influence implementation outcomes, the ten elements were grouped into three reasonably independent factors: RCH characteristics, wellness team characteristics, and quality of implementation processes. Then, a model was proposed to test how these three factors interact to affect level of implementation (Figure 4.1.) We hypothesized that the three factors will be directly related to level of implementation. As there is limited empirical evidence suggesting that provider characteristics may mediate the influences of the other factors,\textsuperscript{15-19} we further hypothesized that wellness team characteristics will mediate the influence of the other two factors.

\textit{Study design}

A prospective observational study design was used to examine factors that influenced implementation levels of the ENRICH intervention. ENRICH employed randomized controlled trial design to test the effectiveness of a 2-year environmental intervention. Twenty-nine RCHs were matched on organizational characteristics and randomly assigned to either the Early intervention group (active intervention 2004-2006, n=17) or the Delayed intervention group (waiting
control with active intervention 2006-2008). The present study analyzed process evaluation data related to implementation of the ENRICH physical activity component collected from both the Early and the Delayed group. These data included level of implementation assessed at the end of the 2-year intervention, and RCH characteristics, wellness team characteristics, and quality of implementation processes measured at baseline and annual process evaluation assessments. The current study was approved by the University of South Carolina’s Institutional Review Board.

**Figure 4.1.** A conceptual framework for examining factors influencing level of implementation of ENRICH.

**Overview of ENRICH**

The ENRICH intervention was guided by the Structural Ecologic Model. The ENRICH physical activity component aimed at increasing physical activity levels among children residing in the RCHs by modifying the structural, policies,
physical, and social environments in the homes. A primary component of ENRICH was utilizing a wellness team formed by RCH adult staff as organizational change agents to plan and execute environmental changes. The wellness team was provided with a framework as a guide to assist them in developing strategic plans for environmental changes. ENRICH employed a flexible and adaptive intervention approach to assist the wellness team in attaining environmental changes. The framework consisted of four essential elements for creating a health-promoting RCH environment: (1) providing opportunities for enjoyable physical activity, which could be achieved through scheduling and provision of equipment; (2) developing, strengthening, and/or enforcing policies that support physical activity; (3) strengthening adult social support and modeling for physical activity; and (4) increasing positive media messages. To facilitate implementation, the investigative team worked in partnership with the RCH staff to identify goals and develop strategic plans for effective environmental changes. Training and ongoing technical assistance, such as workshops, in-service training, consultations, site visits, and resources were provided to the wellness teams to increase their capacity of assessing environments in the homes, and developing and executing the strategic plans. Saunders et al.\textsuperscript{10} and Dominick et al.\textsuperscript{9} provide detailed descriptions of the intervention.

**Participants**

The participants were 29 RCHs that implemented the ENRICH physical activity component. The RCHs varied on organizational characteristics, in which
31% were located in South Carolina and 69% in North Carolina. Sixty-seven percent had a complex structure, which was classified based on whether they had multiple locations, served multiple populations, and/or provided multiple types of care. Thirty percent were participating in the National Breakfast and Lunch Program and 67% were accredited. The current analysis excluded five RCHs as they had incomplete data on the study variables, resulting in a total of 24 RCHs in the final sample.

**Measures**

Items used to assess study variables in the current study were selected from multiple instruments used in the ENRICH process evaluation. Saunders et al.\textsuperscript{10} and Dominick et al.\textsuperscript{9} provide detailed descriptions of the process evaluation measures.

**Level of implementation**

The current study employed the same definition of level of implementation as described in Saunders et al., which is the extent to which the wellness teams changed the RCH environment in accordance to the ENRICH essential elements (fidelity and completeness).\textsuperscript{10} Previously, Saunders et al.\textsuperscript{10} and Dominick et al.\textsuperscript{9} determined level of implementation by triangulating data collected from multiple sources. In this study, level of implementation was determined based primarily on data collected from direct observations. Eight items were used to assess level of implementation. The evaluator completed a seven-item environmental checklist assessing the homes on selected environmental features promoted in the ENRICH framework, including characteristics of physical activity opportunities,
written physical activity policies, social environments, and media environments. Reponse options ranged from 1 “needs improvement” to 3 “excellent.” The wellness team coordinators’ completed one item to report their perception of overall implementation of their strategic plans, with response option ranging from 0 “no, not at all” to 2 “yes completely.” A composite implementation score was calculated by adding the responses to the eight items divided by the highest possible scores. The implementation score was expressed as a percentage and used as a continuum in the analysis.

**RCH characteristics**

Five elements of RCH characteristics were assessed including leadership support for ENRICH, structural characteristics, physical activity resources, policy for committing staff time for physical activity programming, and community connections. The ENRICH staff completed one item assessing their perceptions of leadership support in each home, with response option ranging from 1 “not very well” to 4 “outstanding.” The designated RCH representative completed one item to report RCHs’ structural characteristics based on homes’ complexity (yes/no). The RCH representative also completed four items related to availability of physical activity resources, RCHs’ policy to support staff to use work hours for physical activity programming, and coordination with other agencies. Response options ranged from 0 “does not exist” or “never” to 3 “fully in place” or “always.” Responses to each item were summed to create a composite score for RCH characteristics. A higher score indicates a more conducive RCH environment for implementing physical activity interventions.
Wellness team characteristics

Three elements of wellness characteristics were assessed, including wellness teams’ perceived benefit of their strategic plans on children’s physical activity levels, self-efficacy in implementation, and skill proficiency. Wellness team coordinators responded to two items assessing teams’ perceived benefit of the plans and self-efficacy for implementation on a four-point Likert scale, with the endpoint ranging from “no effect” to “a big effect” or “very unprepared” to “very prepared.” The ENRICH intervention coordinators responded to two items evaluating the wellness teams’ understanding of the environmental change process and ability to engaging other team members. Response options ranged from 0 “no understanding” to 3 “thorough understanding.” The evaluator also rated the wellness teams’ capability in leading the home to make environmental changes, with 1 indicating “needs improvement” and 3 indicating “excellent.” A composite score for wellness team characteristics was calculated by adding the five items, with higher scores indicating wellness characteristics as being more favorable to a successful implementation.

Quality of implementation processes

The ENRICH implementation processes entailed establishing working relationships with the homes and providing training and technical assistance to assist the wellness team in utilizing local physical activity resources and securing leadership support, and increasing teams’ confidence and skills in implementation. The quality of these elements was assessed by seven items. The ENRICH staff completed one item to rate the quality of the working
relationships between ENRCH staff and wellness teams, with response options ranging from 1 “not very well” to 4 “outstanding.” In terms of the quality of training and technical assistance, the ENRICH intervention coordinator completed to two items measuring wellness teams’ effectiveness in utilizing local physical activity resources and maintaining leadership support, with response options ranging from 1 “not at all effective” to 4 “very effective.” Wellness team coordinators completed four items evaluating the overall quantity and quality of the training and the technical assistance. Responses were recorded on a 4-point Likert scale with endpoint ranging from “no support” to “more than adequate” or “very unsupportive” to “very supportive.” A composite score for implementation processes was calculating by adding the seven items, with higher scores indicating the implementation processes being perceived as more favorable to a successful implementation.

**Covariates**

The designated RCH representative responded to 15 items measuring four aspects of the RCHs’ physical activity environments at baseline, including characteristics of physical activity opportunities (8 items), physical activity social environment (1 item), written physical activity policy (2 items), and physical activity media environment (2 items).

**Analysis**

Bayesian path analysis was used to assess the direct and indirect effects of RCH characteristics, wellness team characteristics, and quality of implementation processes in influencing level of implementation of ENRICH. As
the RCHs’ physical activity environments at baseline could be associated with the environmental features at post-intervention, this was adjusted in the model as a covariate. The current study tested a prior hypothesized path model (see Figure 4.2) using M-plus software (version 6.11) with the Bayesian estimation method.

In the path diagram, RCH characteristics and quality of implementation processes are the predictor variables, wellness characteristics are the mediator, RCHs’ physical activity environment at baseline is the covariate, and level of implementation is the outcome variable. Following Wang and Preacher’s conceptual approach to describing mediation, path $a$ and path $z$ are the relationships between the predictor variables and the mediator variable. Path $b$ is the direct effect of the mediator on the outcome variable. Paths $c'$ and $e'$ are the direct effects of the predictor on the outcome variable after controlling for the mediator, and path $f'$ is the effects of RCHs' baseline physical activity environment scores on the level of implementation.

The average indirect effects were calculated as $a*b$ ($RCH$ characteristics $\rightarrow$ wellness team characteristics $\rightarrow$ level of implementation) and $d*b$ ($implementation$ $processes$ $\rightarrow$ wellness team characteristics $\rightarrow$ level of implementation). The total indirect effect ($c$) was calculated as $a*b + d*b$. Statistical significance of the direct and indirect effects was determined based on the 95% Credibility Interval (CI). If the 95% CI for the average direct and indirect effect did not include zero, we would conclude the effects were significant.
The mathematical form of the model is given as follow:

\[
M = d_M + aX_1i + zX_2i + e_m
\]

\[
Y = d_Y + c'X_1i + e'X_2i + bM_i + f'X_3i + e_Y
\]

In the model, \(d\) represents the intercept and \(e\) represents error. The term \(Y\) denotes the outcome variable (IMTLEVEL); \(M\) is the mediating variable (TEAM); \(X_1\) and \(X_2\) are the independent variables (RCH and PROCESS); and \(X_3\) is the covariate (i.e., RCHs’ baseline physical activity environments.) The term \(e_{mi}\) and \(e_{yi}\) are the residuals of \(M\) and \(Y\); the parameter \(d_M\) and \(d_Y\) are the intercepts, and \(a, b, c', e', f', and z'\) are slopes. As accurate prior information is not available, all the unknown parameters were assigned independent non-informative uniform priors. The regression coefficients \(\beta = (a, b, c', e', f', z, d_M, d_Y)'\) were assigned to follow a normal distribution and the variance parameters \(\sigma^2 = (\sigma^2_{e_M}, \sigma^2_{e_Y})'\) follow an inverse-gamma distribution. A large variance in the normal prior above implies a non-informative prior.

Bayesian estimates of all parameters and variance components in the framework were calculated based on 10,000 samples after 1000 burn-in iterations. Posterior mean, posterior standard error and 95% CI at the 0.025 and 0.975 quantiles of the average direct and indirect effects were also obtained. The convergence of the final model was assessed by trace plots, Proportional Scale Reduction (PSR) index, and the autocorrelation plot. A tight and horizontal shape of the trace plots, values for PSR of 1, and values for autocorrelation plots of \(
\leq 0.1 \) indicate good model convergence.\textsuperscript{21} We evaluated the model-to-data fit
based on the 95% confidence interval for the difference between the observed and the replicated chi-square values (95% CI for chi-square value) and the posterior predictive p-value (PPP). A lower negative value of 95% CI for chi-square value and a PPP value close to 0.5 are ideal.

**Results**

Descriptive statistics of the study variables are presented in Table 4.1. Across the 24 RCHs, level of implementation ranged from 38.1% to 97.6% (M=68.35, SD=14.45). The hypothesized model had a good convergence with all the estimated parameters: PRS values were close to 1 (ranged from 1.000 to 1.015); trace plots illustrated a tight and horizontal shape; and the autocorrelation plot showed all parameters had value ≤0.1. The hypothesized model also provided a good fit to the data according to the 95% CI for chi-square value (-17.260, 26.221) and the PPP (p= 0.400).

The path diagram of the hypothesized model tested in this study with standardized path coefficients is presented in Figure 4.2. The direct and indirect effects of the final path model are summarized in Table 4.2. The model accounted for significant variance in levels of implementation ($R^2=0.527$). The results showed that RCH characteristics had significant direct effects on level of implementation ($path c'$) after controlling for other variables. For every standard deviation unit increment in RCH characteristics scores, level of implementation improved by 0.43 standard deviation units, which is approximately 6.2%. Also, wellness team characteristics were found to have significant direct effects on level of implementation ($path b$) after adjusting for other variables. Every one
standard deviation unit increment in wellness team characteristics scores improved level of implementation by 0.36 standard deviation units, which is approximately 5.2%. Quality of implementation processes was not significantly associated with level of implementation (path $d$). Additionally, neither of the hypothesized indirect effects attained statistical significance (path $a*b$ and $d*b$.)

**Discussion**

The present study examined the direct and indirect effects of RCH characteristics, wellness characteristics, and quality of implementation processes on level of implementation of the ENRICH intervention. Results from path analyses showed that both RCH characteristics and wellness team characteristics had direct effects on level of implementation. Though many several theoretical frameworks implicitly suggest that these two factors are directly associated with level of implementation,$^{13, 22, 23}$ this is one of the few empirical studies$^{18, 19, 24, 25}$ demonstrating these direct associations. Also, this is the first study to show these direct effects within the context of a community-based children’s physical activity intervention.

Our data revealed that RCH characteristics and wellness team characteristics had approximately the same amount of influence on level of implementation, which is 6% and 5%, respectively. This finding is congruent with the ENRICH intervention design that an equal emphasis was placed on making changes at the RCH level and the wellness team level. Two of the four essential elements of ENRICH emphasized changing RCHs’ structural and policy environments. It is recognized that changeability of these elements is likely
determined by RCH characteristics, such as leadership support and organizational complexity. Meanwhile, degrees of change in the other two essential elements that focused on the homes’ social and media environments are likely dependent on the wellness team characteristics. Hence, it is not surprising to see that the two factors have a similar amount of influence on level of implementation. Given this equal emphasis design, both the RCH and the wellness team are required to put in efforts in order to achieve high levels of implementation. This may also explain why wellness team characteristics did not have significant mediating effects on the relationships between RCH characteristics and level of implementation.

Several conceptual frameworks have posited that quality of implementation processes, including working relationship, training, and technical assistance are necessary to program implementation. However, none of these frameworks have specified the underlying pathways in which these elements influence level of program implementation. The current study provide empirical evidence demonstrating that quality of implementation processes was not directly associated with level of implementation after controlling the effects of RCH characteristics and wellness team characteristics. This finding indicates that, though necessary, only having high quality implementation processes, comprised of two elements, may not be sufficient to directly influence level of implementation.

There is limited evidence suggesting that quality of implementation processes may indirectly influence program implementation through the
enhancement of providers’ capacities.\textsuperscript{27,28} However, this association was not observed in the current study. Our findings showed that the indirect effect from quality of implementation processes through wellness team characteristics to level of implementation did not attain statistical significance. The non-significant indirect effects could be due to many factors. It is possible that the implementation processes employed in ENRICH were not intensive enough to produce sizable changes in wellness teams’ capacities, thus resulting in the null effects. It may be that program implementation is only significantly associated with some of the elements (e.g., technical assistance) but not all of them. In this study, level of implementation is limited to completeness and fidelity. It is also probable that these implementation processes may be associated with other implementation outcomes. However, given the lack of research in this area, more studies are needed to confirm or refute these speculations.

Strengths of the current study include objective measures of level of implementation and comprehensive assessment of RCH characteristics. Moreover, the use of the Bayesian estimation method provides us the statistical power to test the proposed relationships that are usually lacking in implementation research. There are a number of limitations of the present study that warrant further explorations in future research. Findings from the path analysis do not imply causality.\textsuperscript{29} Future studies should employ experimental study design to examine whether manipulating the factors identified in the current study is associated with implementation outcomes.
The current study employed a holistic approach to construct the influences of implementation, in which 10 elements were grouped into three global factors. This approach is different from most of the previous studies that focused on the effect of individual elements. In fact, neither the holistic approach nor the individual element approach alone would be sufficient to help us understand how multiple elements function as a complex system to influence level of implementation. Due to a small sample size, we are not able to combine the two approaches in the present study. Future studies should employ both approaches in the same study to allow researchers to understand the system as an integrated whole.

Although RCH characteristics, wellness team characteristics, and quality of implementation processes were defined by 10 elements referenced in the literature, it is possible that some potential elements are omitted. Future studies should continue to explore other elements to improve the definition of the three factors. Though model fit was adequate, due to the small sample size, the stability of the path model should be confirmed in future studies with a larger sample.

Studies examining factors that influence implementation of youth physical activity interventions is in its infancy. The current study contributes valuable empirical evidence supporting the important role of the organizational contexts and providers’ attitudes and skills in influencing implementation of a physical activity intervention undertaken in community-based organizational settings. Assessing these characteristics in formative assessments and immediately after
the initial training sessions would provide important information to inform implementation planning, such as developing context-specific technical assistance. Better implementation planning may increase the chance of achieving higher levels of intervention implementation, which may in turn improve program effectiveness.

Table 4.1. Descriptive statistics of residential children's homes participated in ENRICH

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total sample (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Level of implementation</td>
<td>68.35 (14.45)</td>
</tr>
<tr>
<td>RCH characteristics</td>
<td>13.92 (2.99)</td>
</tr>
<tr>
<td>Wellness team characteristics</td>
<td>10.85 (1.97)</td>
</tr>
<tr>
<td>Quality of implementation processes</td>
<td>8.60 (1.59)</td>
</tr>
<tr>
<td>Baseline PA environment</td>
<td>31.67 (4.99)</td>
</tr>
</tbody>
</table>

Note: PA=physical activity, RCH=residential children’s home.
Figure 4.2. Final path model illustrating factors that influence level of implementation of ENRICH. Standardized path coefficients are shown. Asterisk indicates significant association with 95% CI not including zero. Bolded solid path indicates a significant direct effect. RCH=residential children’s homes. The model shows that RCH characteristics and wellness characteristics directly influenced level of implementation, after adjusting for other covariates.
Table 4.2. Direct and indirect effects of the selected factors on level of implementation of ENRICH

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimates (SD)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direct effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCH → IMLEVEL (path c')</td>
<td>0.431 (0.201)</td>
<td>0.015 to 0.801</td>
</tr>
<tr>
<td>TEAM → IMLEVEL (path b)</td>
<td>0.356 (0.172)</td>
<td>0.005 to 0.676</td>
</tr>
<tr>
<td>PROCESS → IMLEVEL (path e')</td>
<td>-0.287 (0.189)</td>
<td>-0.664 to 0.089</td>
</tr>
<tr>
<td>RCH → TEAM (path a)</td>
<td>0.245 (0.240)</td>
<td>-0.258 to 0.668</td>
</tr>
<tr>
<td>PROCESS → TEAM (path d)</td>
<td>0.065 (0.242)</td>
<td>-0.416 to 0.529</td>
</tr>
</tbody>
</table>

**Indirect effects**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimates (SD)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCH → TEAM → IMLEVEL (path a*b)</td>
<td>0.089 (0.112)</td>
<td>-0.100 to 0.347</td>
</tr>
<tr>
<td>PROCESS → TEAM → IMLEVEL (path d*b)</td>
<td>0.025 (0.103)</td>
<td>-0.177 to 0.251</td>
</tr>
</tbody>
</table>

**Model fit**

- 95% CI for difference between observed and replicated chi-square values: -17.260 to 26.221
- Posterior Predictive P-value: 0.400

Notes: Boldfaced indicates a significant effect with 95% CI not including zero. Estimates: average of the posterior means, SD: average of the posterior standard deviations, 95% CI: lower and upper bounds of the 95% credibility interval. ENRICH=Environmental Intervention in Children’s Home; IMLEVEL=implementation levels; PROCESS=implementation processes; TEAM=wellness team characteristics; RCH=residential children’s home characteristics.
References


CHAPTER 5

OVERALL DISCUSSION

Significance

A substantial percentage of American children and adolescents fail to meet the physical activity guideline.\(^1\)\(^,\)\(^2\) Over the years, numerous interventions have been developed and implemented to increase youth’s physical activity. However, these interventions have achieved limited success.\(^3\) Emerging evidence shows that sub-optimal program implementation is one of the important factors that contribute to the lack of significant intervention effects.\(^4\)\(^,\)\(^5\) This suggests that future interventions may be more likely to succeed if they are implemented optimally. To date, very little is known about the specific factors that influence implementation of youth physical activity interventions. The lack of knowledge in this area is a major limitation to the development of effective implementation strategies. This dissertation is significant because it identifies and describes a comprehensive list of factors that influence program implementation, specifically within the context of youth physical activity interventions. This thereby provides the needed knowledge to develop effective strategies for improving program implementation. Better program implementation may improve the effectiveness of youth physical activity interventions, which may ultimately help increase physical activity levels among children and adolescents.
Purpose

The overall purpose of this dissertation was to identify a set of core factors that are most important in explaining implementation of physical activity intervention undertaken in youth-serving organizations. The purpose of the first study was to convene a panel of experts to identify factors that are most important in achieving successful implementation of physical activity interventions carried out in youth-serving organizations. The purpose of the second study was to examine the direct and indirect effects of preschool characteristics, teacher characteristics, and quality of implementation processes on level of implementation of a previously completed physical activity intervention carried out in preschool settings. The purpose of study three was to examine the direct and indirect effects of residential children’s home (RCH) characteristics, wellness team characteristics, and quality of implementation processes on level of implementation of a previously studied physical activity intervention carried out in RCHs.

Design and Methods

This dissertation employed two study designs. The first was a cross-sectional expert panel study. Five recognized experts engaged in a four-round, modified Delphi process to identify factors related to implementation of youth physical activity interventions, and quantify the importance of the identified factors. Experts’ opinions were translated into Bayesian predictive models for factor selections.
The second and third studies used a prospective observational study design. In the second study, data were taken from the first two years of the Study of Health and Activity in Preschool Environments (SHAPES) intervention. The participants were 19 and 17 preschool classrooms enrolled in year 1 and year 2 of SHAPES, respectively. The outcome variable, level of implementation, was measured by direct observation. The three selected factors, including preschool characteristics, teacher characteristics, and quality of implementation processes, were measured by direct observations, interviews, evaluation forms, and surveys. In the third study, data were taken from the Environmental Intervention in Children’s Homes (ENRICH) study. The participants were 29 RCHs enrolled in ENRICH. The outcome variable, level of implementation, was measured by direct observation. The three selected factors, including organizational characteristics, wellness team characteristics, and quality of implementation processes, were measured by interviews, evaluation forms, and surveys. Both the second and third studies used Bayesian Path analysis to test the hypothesized direct and indirect associations between the selected factors on the level of implementation.

**Major findings**

Overall, this dissertation identified a list of factors that are likely to influence implementation of physical activity interventions undertaken in youth-serving organizations, which is presented in Table 5.1. In the first study, the experts identified 15 factors that are most important in achieving successful implementation of physical activity interventions in youth-serving organizations, in which five factors were classified as organizational characteristics, six factors as
implementation processes, two factors as provider characteristics, and two factors as program characteristics and community-level factors, respectively.

In the second study, the three selected factor explained 15% of the variance of level of implementation, but the selected factors were not significantly associated with level of implementation in year 1. In year 2, the three selected factor explained 53% of the variance in level of implementation, preschool characteristics were found to be directly associated with level of implementation. Teacher characteristics and quality of implementation processes were not significantly associated with level of implementation. These findings suggest that factors that influence level of implementation differed by stages of implementation. In the third study, the contribution of the three selected factors in explaining level of implementation was 53%. The results showed that RCH characteristics and wellness team characteristics were directly associated with level of implementation. Neither direct nor indirect associations between quality of implementation processes and level of implementation attained statistical significance. However, more studies are needed to investigate in what contexts do quality of implementation processes influence level of implementation.

Due to differences in definitions and measures of the selected factors, it is impossible to make direct comparisons on the individual factors across the three studies. Nonetheless, the three studies demonstrate consistent evidence supporting the direct contribution of organizational characteristics to program implementation. However, the effects of provider characteristics and quality of implementation processes on level of implementation appear to differ depending
on implementation stages and intervention designs. Additionally, the findings from the second and third studies appear to suggest that quality of implementation processes comprising of training, technical assistance, and working relationships, may not be sufficient to influence level of implementation. These findings suggest the needs of considering the comprehensive list of factors identified in the first study.

**Limitations**

This dissertation has several limitations that should be considered. A limitation of the first study was that the expert panel did not incorporate the perspectives from the front-line intervention staff, such as project coordinators. However, this is based on the consideration that these individuals, employed on a project-by-project basis, may not possess optimal amounts and diversity of experiences enabling them to identify influences of program implementation that are commonly observed across multiple youth physical activity interventions. Also, the predictive model developed for the first study has not yet been tested for external validity.

A major limitation of the second and third studies is that both were secondary data analyses. Therefore, the definitions and measures of the selected factors in these two studies were different from the first study, thus precluding direct comparisons on each individual factor. Additionally, findings from the Bayesian path analyses do not imply causality. Future experimental studies are needed to further investigate the causal pathways between the influencing factors and level of implementation.
Practical implications

This dissertation has implication to investigators that seeking to developing strategies for improving implementation of youth physical activity interventions. The list of well-described factors identified in this dissertation provides important information to guide the development of implementation strategies for future intervention studies. Investigators can also use the factors identified in this dissertation to develop theoretical models for explaining program implementation of youth physical activity interventions.

Findings from this dissertation also have implication to investigators who are interested in assessing factors that affect implementation of youth physical activity interventions. The list of factors accompanying with the Bayesian predictive model produced in the first study can immediately be used an instrument. Researchers can validate this instrument in different settings; they can also use the factors identified in this dissertation to develop other measures.

Considerations for future research

Program implementation is a complex process. It is unlikely that the same factors will uniformly explain the variation in implementation in all interventions. It is much more likely that the salient factors will vary across interventions with different designs, at different implementation stages, and in different implementation settings. However, there is currently lacking of theories or empirical evidence to explain such variations. Therefore, future studies should continue to examine why and when the factors identified in this dissertation affect
implementation of physical activity interventions in different youth-serving settings.

**Conclusion**

Overall, this dissertation found consistent evidence supporting the direct contribution of organizational characteristics in achieving successful implementation of physical activity interventions in youth-serving organizations. However, the effects of provider characteristics and quality of implementation processes on level of implementation appear to differ across interventions with different intervention designs, at different implementation stages, and in different implementation settings. Further studies are needed to test the validity of the identified factors in different youth-serving settings.
Table 5.1. Summary of influences of program implementation identified in the three studies conducted in this dissertation

<table>
<thead>
<tr>
<th>Study 1: Expert panel study</th>
<th>Study 2: Preschool study</th>
<th>Study 3: RCH study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational characteristics</strong></td>
<td><strong>Organizational characteristics</strong></td>
<td><strong>Organizational characteristics</strong></td>
</tr>
<tr>
<td>Leadership engagement &amp; motivation</td>
<td>Leadership support</td>
<td>Leadership support</td>
</tr>
<tr>
<td>PA resources: facilities and equipment, space, staff</td>
<td>PA resources</td>
<td>PA resources</td>
</tr>
<tr>
<td>Competing programs in the organization</td>
<td>Structural characteristics</td>
<td>Structural characteristics</td>
</tr>
<tr>
<td></td>
<td>PA policies and practices</td>
<td>PA policies</td>
</tr>
<tr>
<td></td>
<td>Community connection</td>
<td>Community connection</td>
</tr>
<tr>
<td></td>
<td>Organizational functioning</td>
<td>Organizational functioning</td>
</tr>
<tr>
<td></td>
<td>Organizational climates</td>
<td>Organizational climates</td>
</tr>
<tr>
<td><strong>Quality of implementation processes</strong></td>
<td><strong>Quality of implementation processes</strong></td>
<td><strong>Quality of implementation processes</strong></td>
</tr>
<tr>
<td>Training and technical assistance</td>
<td>Training and technical assistance</td>
<td>Training and technical assistance</td>
</tr>
<tr>
<td>Needs assessment</td>
<td></td>
<td>Working relationships</td>
</tr>
<tr>
<td>Engaging intervention staff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engaging program champion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reflecting and evaluating</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Provider characteristics</strong></td>
<td><strong>Provider characteristics</strong></td>
<td><strong>Provider characteristics</strong></td>
</tr>
<tr>
<td>Belief and motivation</td>
<td>Self-efficacy</td>
<td>Self-efficacy</td>
</tr>
<tr>
<td>Knowledge and skills</td>
<td>Problem solving skills</td>
<td>Skills</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perceived program benefits</td>
</tr>
<tr>
<td><strong>Program characteristics</strong></td>
<td><strong>Program characteristics</strong></td>
<td><strong>Program characteristics</strong></td>
</tr>
<tr>
<td>Adaptability</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Community-level factors</strong></td>
<td><strong>Community-level factors</strong></td>
<td><strong>Community-level factors</strong></td>
</tr>
<tr>
<td>Competing programs in the community</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note: Boldface indicates the factors had statistically significant associations with level of implementation in the respective study. PA=physical activity, RCH=residential children’s’ homes
References


Introduction

Regular physical activity is associated with numerous health benefits in children and adolescents. Despite the well-recognized health benefits, less than 50% of the U.S. children and adolescents obtained the recommended amount of physical activity. To address the low prevalence, many physical activity interventions have been developed and implemented. Overall, these interventions have demonstrated positive effects on youth’s physical activity behaviors. However, the overall magnitudes of change were small.\textsuperscript{1,2} The modest effects have prompted researchers, health agencies, and program funders to examine factors influencing program effectiveness. Emerging evidence has suggested that inadequate levels of implementation have been linked with poor program outcomes. A 2008 review\textsuperscript{3} of health promotion and preventive interventions showed that the effect sizes were two to three times lower in programs with lower levels of implementation than those with higher levels of implementation.

Several reviews of the literature\textsuperscript{3-7} indicate that levels of implementation are influenced by factors operating at multiple levels: macro level (e.g., consistency with federal policies community partnership); organizational level
(e.g., organizational capacities, climates, leadership, administrative support, resources); implementer level (e.g., professional characteristics and implementer's perception of the innovation); and program level (e.g., structural characteristics and implementation processes). It is important to note that these potential factors taken from the literature of youth preventive interventions (e.g., mental health and substance abuse programs) may not be applicable to physical activity interventions delivered in youth-serving organizations (e.g., schools, childcare centers, and community recreation centers). Also, it is not known which of these factors are most important in predicting a successful implementation.

**Statement of the problem**

The overarching goal of the proposed study is to examine how specific characteristics of the organization, implementer, and program influence levels of implementation of physical activity interventions targeting youth-serving organizations. The specific aims and objectives of the proposed study are outlined below.

**Aim 1: To develop a Bayesian model for predicting successful implementation of physical activity interventions in schools.**

Objective 1a: Convene a panel of experts to identify factors influencing the successful implementation of youth physical activity interventions. Panel members will have expertise in physical activity interventions in children and adolescents, and implementation research.
Objective 1b: The information obtained in Objective 1a will inform the development of an initial Bayesian model to predict the probability that a school will successfully implement a physical activity intervention.

Objective 1c: To examine the internal validity of the model developed in Objective 1b.

**Aim 2: To examine the effects of specific characteristics of the organization, implementer, and program on levels of implementation of a physical activity intervention delivered in a preschool setting.**

Objective 2: To assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in influencing levels of implementation of a physical activity intervention delivered in a preschool setting.

Hypotheses 2a: The specific characteristics of the organization, implementer, and program will have significant direct effects on levels of implementation.

Hypotheses 2b: The specific characteristics of the organization and program will have significant indirect effects on levels of implementation mediated through the characteristics of the implementers.

**Aim 3: To examine the effects of specific characteristics of the organization, implementer, and program on levels of implementation of a physical activity intervention delivered in a children’s group home setting.**

Objective 3: To assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in
influencing levels of implementation of a physical activity intervention delivered in a children’s group home setting.

Hypotheses 3a: The specific characteristics of the organization, implementer, and program will have significant direct effects on levels of implementation.

Hypotheses 3b: The specific characteristics of the organization and program will have significant indirect effects on levels of implementation mediated through the characteristics of implementers.

Scope

With the exception of Objectives 1a and 1b, all objectives of the proposed study will be addressed by analyzing existing data sets. Specifically, the scope of Objective 2 will be limited to the characteristics of organization, implementer, and program that were measured in Study of Health and Activity in Preschool Environments (SHAPES). SHAPES was a one-year group randomized trial which aimed to increase the physical activity levels of 3- to 5-year old children in preschool settings through increasing physical activity promoting practices and policies (i.e., instructional and environmental factors) in the preschool classrooms. The scope of Objective 3 will be limited to the characteristics of organizations, implementers, and programs that were measured in Environmental Interventions in Children’s Homes (ENRICH) study. ENRICH was a two-year group randomized trial with the overall goal to promote physical activity and healthful nutrition behavior in a population of children residing in group homes. Within both the SHAPES and ENRICH studies, a comprehensive
process evaluation was conducted to assess levels of implementation.

Information related to the specific characteristics of the organizations, implementers and programs were obtained by using multiple approaches (i.e., direct observation, interview, review of documentation, and self-report survey) and drawn from multiple sources (i.e., primary implementers, target audiences, administrators and intervention staffs).

**Significance of the proposed study**

The overall findings of the proposed study will help to identify the critical characteristics of successful implementation of physical activity interventions in youth-serving organizations. The research findings will have the following important implications for implementation: The identified characteristics will provide researchers with a conceptual framework to design context-specific service delivery protocol, thus increasing the likelihood of effective implementation. The findings can be used by youth-serving organizations to identify their specific strengths and weaknesses to implement a physical activity intervention, so that adequate resources can be allocated to assist an organization to achieve desired levels of implementation. Funding agencies can also take the identified characteristics into consideration when interpreting the effectiveness of funded physical activity interventions.

**Limitations**

There are several limitations of the proposed study. First, the use of the Bayesian statistical method in all three aims incorporates prior subjective information in model estimations, which could make it a less objective approach.
However, given the lack of empirical evidence in this area, a Bayesian model that incorporates prior subjective estimates from the expert panel is considered to be a time efficient and cost-effective method to answer the research questions in the proposed study. The specific characteristics examined in Aim 2 and 3 will be limited to those that were measured in the two existing data sets. It is possible that other important predictors will not be included in the model due to this limitation. Also, the proposed study will examine the predictors in two settings only, preschools and children’s group homes. Conducting comprehensive process evaluations in large-scale intervention studies requires extensive resources; therefore, analyzing existing data sets with detailed process data can provide timely and valuable information to inform the design of future studies.

**Operational definitions**

**Physical activity interventions**

An intervention is defined as “any activity of a program that aims directly at changing the target behavior or its related determinants.”\(^5\) Interventions can include specific programs, policies, practices, or principals.\(^8\) In the proposed study, physical activity interventions are operationally defined as structured programs or practices that aim at improving physical activity participations in children and adolescents. Policies are excluded from the definition because policy implementation involves a set of factors that are different from structured programs or practices. For example, factors that are considered as distal to implementation success of structured programs would become proximal factors to successful policy implementation (e.g., community-university partnership and
Throughout this proposal, the terms intervention, program, innovation, are used interchangeably.

**Preventive interventions**

The proposed study operationally defined preventive interventions as structured programs that aim at preventing the development of a disease, disorder or health risk behavior, reducing its complication and lowering its negative influence on an individual’s quality of life. For example, prevention programs can be targeting substance abuse, cigarette smoking, mental health, immunization, hypertension or cancer.

**Youth**

In this proposal, youth is used as a general term referring to children and adolescents. Youth is operationally defined as individuals aged 2.0 to 18.9 years. The terms children, adolescents, and youth are used interchangeably throughout this proposal.

**Youth-serving organizations**

Youth-serving organization can be operationally defined as any organization that provides services to a group of children and adolescents. This can include childcare centers, schools, churches, neighborhoods, or local recreation centers. These organizations may operate on a local, national, or even international level; and can be developed either by young people themselves or adults such as coaches, ministers, or staff of the local YWCA.

**Review of the literature**

**Physical activity and health in youth**
The Surgeon General’s Report on Physical Activity and Health\textsuperscript{12} was the first report that documented the health benefits of regular physical activity in children and adolescents. The report concluded that higher levels of physical activity can favorably influence youth’s blood pressure, high-density lipoprotein cholesterol, and blood lipid profile in children and adolescents with high risk of developing coronary heart diseases. Also, regular, weight-bearing physical activity plays a substantial role in the development of bone mass during childhood and adolescence. The health benefits of regular physical activity in youth were further summarized in several subsequent review studies.

In 2005, the Centers for Disease Control and Prevention and the Constella Group convened a multi-disciplinary expert panel to review the evidence on physical activity and health in school-age children.\textsuperscript{13} The expert panels reviewed over 850 articles published before 2004. The panel concluded that the evidence strongly supports the beneficial effects of physical activity on musculoskeletal health, adiposity in overweight youth, and blood pressure in mildly hypertensive youth. Also, there is adequate evidence supporting the positive effects of regular physical activity on lipid and lipoprotein levels, adiposity in normal weight youth, blood pressure in normotensive youths, self-concept, anxiety, depression, and academic performance. In 2008, the Physical Activity Guidelines Advisory Committee Midcourse Report\textsuperscript{14} also reviewed the health benefits of physical activity in youth based on publications from 1995 to 2007. The report concluded that regular physical activity has beneficial effects on adiposity, physical fitness (both cardiorespiratory fitness and muscular strength), cardiovascular and
metabolic disease risk profiles, bone health, and depression and anxiety symptoms. In 2010, Janssen and colleagues\textsuperscript{15} reviewed a total of 86 observational and experimental studies to examine the health benefits obtained from different types of physical activity. The evidence strongly supports that aerobic exercise has been shown to reduce cholesterol, blood lipid, fasting insulin, insulin resistance, total and abdominal fat, blood pressure, depressive symptoms, and increase bone mineral density. Additionally, the authors highlighted the importance of muscle strengthening exercises in maintaining bone health among youths.

\textbf{Prevalence of physical activity in youth}

The 2008 Federal Physical Activity Guidelines for Americans recommend that children and adolescents participate in at least 60 minutes of moderate-to-vigorous physical activity (MVPA) per day, on most days of the week.\textsuperscript{14} However, population surveillance systems show that only 41.1\% of U.S. children and adolescents met the recommendations.\textsuperscript{16} The National Health and Nutritional Examination Survey 2003-2004 showed that the percentage of American youths meeting the physical activity recommendation was 42\%, 8\%, and 7.6\% for boys and girls ages 6 to 11 years, 12 to 15 years, and 16 to 19 years, respectively, as measured by accelerometry.\textsuperscript{17} Self-reported physical activity of American youth from the 2013 Youth Risk Behavior Survey showed that only 27.1\% of 9\textsuperscript{th} to 12\textsuperscript{th} graders reported being physically active at the recommended level on a daily basis; and only 47.3\% were active five days of the week.\textsuperscript{18} The prevalence of physical activity varies by age, gender, and ethnicity. Thirty percent of 9\textsuperscript{th} graders
met the recommended physical activity levels compared to only 24.3% of 12th graders. The percentage of youth meeting the physical activity recommendation was higher in boys (36.6%) than in girls (17.7%) regardless of racial/ethnic group; furthermore, white boys were the most likely to meet recommendations (37.5%) while black girls were the least likely to meet recommendations (16.0%).

Physical activity behaviors begin to be established in childhood and tend to track across the lifespan. Inactive children are more likely to become inactive adolescents; and inactive adolescents are more likely to become inactive adults. Hence, it is important to develop effective strategies to integrate into intervention programs that promote healthy physical activity habits among children.

**Youth physical activity interventions**

Over the years, studies have recognized that youth participation in physical activity is influenced by factors at the individual, social, environment, and policy level. Interventions targeting factors at multiple settings, especially where children live, learn, and play, are more likely to be successful.

Healthy People 2020, 10-year national objectives for improving the health of all Americans, has set the following objectives that aim to increase youth physical activity participation in various settings:

**School settings**

PA-4 Increase the proportion of the Nation’s public and private schools that require daily physical education for all students
PA-5 Increase the proportion of adolescents who participate in daily school physical education

PA-6 Increase regularly scheduled elementary school recess in the United States

PA-7 Increase the proportion of school districts that require or recommend elementary school recess for an appropriate period of time

*Active commuting*

PA-13.2 (Developmental) Increase the proportion of trips of 1 mile or less made to school by walking by children and adolescents aged 5 to 15 years

PA-14.2 (Developmental) Increase the proportion of trips of 2 miles or less made to school by bicycling by children and adolescents aged 5 to 15 years

*Preschool and childcare center settings*

PA-9 Increase the number of States with licensing regulations for physical activity provided in child care

*Community settings*

PA-10 Increase the proportion of the Nation’s public and private schools that provide access to their physical activity spaces and facilities for all persons outside of normal school hours (that is, before and after the school day, on weekends, and during summer and other vacations)
PA-15 (Developmental) Increase legislative policies for the built environment that enhance access to and availability of physical activity opportunities

PA-15.1 (Developmental) Increase community-scale policies for the built environment that enhance access to and availability of physical activity opportunities

PA-15.2 (Developmental) Increase street-scale policies for the built environment that enhance access to and availability of physical activity opportunities

PA-15.3 (Developmental) Increase transportation and travel policies for the built environment that enhance access to and availability of physical activity opportunities

The following section summarizes the effectiveness of physical activity interventions implemented in schools, preschool, and childcare centers, and community settings.

School settings

Schools are an ideal setting for promoting physical activity among children and adolescents as it can reach a large percentage of the population. In the U.S., over 95% of 5- to 17-year-old children attend school. Additionally, there are other reasons for promoting physical activity in school settings, including the number of hours children spend in school each day, and the availability of personnel and infrastructure resources in educational settings.25

Physical education
School physical education (PE) is an important venue for equipping children with the knowledge and skills to engage in lifelong physical activity. The National Association for Sport and Physical Education (NASPE), and the Guidelines for School and Community Programs to Promote Lifelong Physical Activity Among Young People from the Centers for Disease Control and Prevention (CDC) recommend 30 minutes of daily PE for elementary students and 45 minutes for secondary school students. Additionally, teachers are also recommended to maximize physical activity during PE lessons and keep students moderately to vigorously active for at least 50% of class time. However, one study found that daily physical education is offered in only 3.8% and 7.9% of elementary and middle schools, respectively. Other studies have reported that the proportion of time students engage in MVPA during PE class was only 34.7% in elementary schools, 27% in middle schools, and 47% in high schools.

To increase students' time spent in MVPA during PE, several intervention strategies have been used to modify existing PE, including increasing activity choices during PE, providing PE teachers with training on quality instructional time, incorporating motivation components into the PE curriculum, and providing resources (i.e., additional PE specialists and equipment). Kahn and colleagues reviewed 17 studies published between 1980 and 2000. They found that, on average, the PE interventions increased students' time in MVPA during PE by 50.3%. Slingerland and Borghouts reviewed 19 articles published
between 1989 and 2009 and they also found that modified PE increased students’ MVPA during PE classes by 4% to 21%.

Although school PE is effective in increasing the amount of time students engage in MVPA, its contribution to students’ total physical activity was small. For example, the Sports, Play and Active Recreation for Kids (SPARK) program provided 40 minutes of MVPA per week, which is only 13.3% of the total amount of weekly physical activity recommended for children. Slingerland and Borghouts\textsuperscript{31} examined the effects of 13 PE interventions on youth’s physical activity levels outside of class, but the findings are mixed. Seven studies found a null effect; two studies\textsuperscript{32,33} reported significant increments in students’ physical activity levels; two studies\textsuperscript{34,35} found significantly less reduction in physical activity levels; and another two studies\textsuperscript{36,37} found significant increments in physical activity levels among boys only.

**Recess**

As the contribution of PE to overall physical activity levels is limited, researchers have begun to explore other non-PE approaches to promote physical activity during school hours. Providing opportunities for students to participate in physical activities outside PE, such as recess and lunch time is a potential alternative.\textsuperscript{38} Ridgers and colleagues\textsuperscript{39} observed that, under non-intervention condition, activities performed during recess breaks could contribute to a significant proportion of the recommended MVPA of 60 minutes a day in children (4.7% to 40% in boys and 4.5% to 30.7% in girls).
Regarding the effectiveness, two review studies\textsuperscript{39,40} have found that providing activity breaks during recess or lunch time increased children’s physical activity by 17% to 60%; however, negative effects were observed in a few studies. Erwin and colleagues\textsuperscript{41} conducted a meta-analysis of 28 recess interventions to examine their effects on children’s physical activity levels. The findings demonstrated that post-intervention physical activity levels were significantly higher in children who participated in recess interventions than those in the control group. Although the overall effect size reached a medium range (d=0.56), it varied widely across individual studies (d=0.01 to 2.36). Also, it is important to note that over half of the studies included only attained a small effect and the effect sizes ranged from the lowest of 0.01 to highest of 2.36.

**Active commuting to/from school**

Cross-sectional studies showed that active commuters are more likely than passive commuters to engage in MVPA throughout the day (4.7 to 40 minutes).\textsuperscript{36,37} However, the prevalence of children using these active transportation modes to school has been low. In the U.S. in 2009, 12.7% of elementary and middle school children reported usually walked or biked to school.\textsuperscript{42} In response to the low prevalence, increasing numbers of interventions have been designed to promote active transportation to/from schools in recent years. Chillion and colleagues\textsuperscript{43} reviewed 14 active transport programs published through January 2010. They found that almost all programs have demonstrated an improvement on the percentage of children using active travel modes to school, but the improvement varied from 3% to 64%. The magnitude of effects
also varied widely. Twelve studies reported modest to small effect sizes (d=0.07 to 0.32), but three studies produced a large to very large effect (d=0.86 to 2.90).

Preschool and childcare center settings

Similar to the elementary, middle, and high school settings, the preschool and childcare center setting provides an important venue to reach a large proportion of young children. Over 60% of 3- to 5-year-old American children are enrolling in center-based preschools; millions of American children spend a significant portion of their day in childcare centers. 44 NASPE guidelines for preschoolers recommend at least 120 minutes of physical activity daily. Previous studies 45,46 have found that preschoolers generally accumulate less than 60 minutes of MVPA per day that is much lower than the recommended levels. Hence, there is a need to promote physical activity among this population.

Ward and colleagues 47 reviewed eight interventions that aimed to improve children’s physical activity levels in childcare settings. Three studies used policies or environmental strategies and two of the three found positive effects. The other five studies delivered a physical activity program during the curricular time, but only two of the five studies found positive effects. Kreichauf and colleagues 48 reviewed the effectiveness of seven preschool-based physical activity interventions. Two studies used environmental strategies and both studies found that providing portable play equipment in the preschools increased children’s MVPA. The other five studies employed the curricular approach, but only two studies found positive effects on children’s physical activity levels. Among the three studies with null findings, one study specifically explained the
possible reasons for the null findings. Reilly and colleagues\textsuperscript{49} implemented a 24-week physical activity program which consisted of three components: a 30-minute physical activity session for three days per week; distribution of guidance on physical activity to intervention group’s parents; and display posters focusing on a 6-week physical activity plan in the nursery areas. This study observed a significant intervention effect during the pilot study; while a null effect was found in the main study. The authors noted that the failure in replicating the findings could attribute to the difference in implementers’ competencies. Compared to the nursery head teachers who implemented the pilot program, the nursery staff members that delivered the main study appeared to be less competent and they delivered the program with a lower quality.

**Community settings**

The communities where children live in and spend their leisure-time likely have an effect in shaping their physical activity behaviors. The Healthy People 2020 objectives have devoted much attention on enhancing the community built environment. Observational evidence showed that children’s participation in physical activity is positively associated with higher accessibility to recreational facilities and schools, neighborhood walkability, residential density, and mixed land-use.\textsuperscript{50} However, there is dearth of interventions designed to promote physical activity through changing the built environment. Van Sluijs and colleagues\textsuperscript{51} reviewed four community-based interventions, but none of them targeted the built environment. Salmon and colleagues\textsuperscript{52} identified three community-based physical activity interventions, but only one study focused on
the built environment. The intervention had three components: traffic control; improvements to pedestrian crossings; and sidewalk improvements. The results showed that children who passed by the improved areas are more likely than those did not pass by the areas to walk or bike to school.

Overall, the majority of the intervention studies to increase physical activity in youth have demonstrated positive effects. However, the magnitude of change varied greatly within and across settings. Researchers have started to explore what could have contributed to the variability. Is it due to poorly designed or inappropriate use of theoretical frameworks\textsuperscript{5,53} or other reasons? Unfortunately, many intervention studies have employed a “black box” evaluation approach that only tells us “whether” a program works but nothing on “why and how” a program does or does not work.\textsuperscript{54,55} To open the “black box”, researchers first need to know whether or not a program was implemented. If it did get implemented, to what extent, how well, how receptive were the change agents and participants, and to whom? In response to these questions, increasing research attention has recently been devoted to study program implementation.

\textit{The importance of monitoring and assessing program implementation}

Information on program implementation can have an important impact across all intervention phases. During the implementation phase, information on program implementation is crucial for monitoring progress toward the program goal, detecting potential errors or problems, and providing on-going feedback to maintain and improve program delivery in implementation.\textsuperscript{56-59} Ultimately, the intended program outcomes are more likely to be achieved. During the
evaluation phase, implementation data are essential for drawing valid conclusions about the effectiveness of a program. Without evidence of program implementation, researchers may erroneously conclude that the program is ineffective when, in fact, the insignificant outcomes are a result of poor program delivery (Type III error). The incorporation of implementation data into outcome analyses enables researchers to correctly conclude that any significant or insignificant findings were linked to the program theories but not because of poor program delivery. During the dissemination phase, data on program implementation can facilitate researchers’ understanding of the strengths and weaknesses of an intervention (e.g., fit between the program and local context, barriers for carrying out the program). Such information should also be used to inform program modifications and the development of standard service delivery protocol for large scale dissemination.

**Introduction to program implementation**

**Definitions**

Implementation is the process of putting to use or integrating evidence-based interventions within a setting. It has two dimensions: implementation process and implementation outcome. Implementation process refers to “a specific set of activities designed to put into practice an activity or program of known dimensions.” It could include activities such as training staff members on using the intervention, and forming an implementation team. Implementation outcome or levels of implementation are also termed as implementation integrity or treatment fidelity in other studies. It is defined as “the extent of change that
has occurred at some particular time toward full, appropriate use of the target innovation” 58. The focus of the proposed study is on implementation outcome or levels of implementation.

Aspects of implementation levels: Individually-oriented interventions

According to Liannan and Steckler, 63 the level of implementation is a multidimensional construct which consists of four aspects: dose delivered, fidelity, dose received, and reach.

Dose delivered

Dose delivered is also known as quantity, dose, or completeness. It is defined as the actual amount or proportion of the total prescribed intervention components delivered by the implementers to the target participants. 63 For example, an intervention required teachers to deliver a 10-session health education curriculum. If a teacher implemented eight sessions, the dose delivered for the intervention would be eight out of 10 sessions or 80%. Data on dose delivered can be obtained through direct observation. Trained observers count the proportion of prescribed intervention components that were delivered during any observed sessions. Another approach is to have implementers self-report the amount of intervention elements they delivered by using checklists or activity logs.57

Fidelity

Different from the quantitative emphasis of dose delivered, fidelity focuses on the quality and integrity of the implementation. It is defined as how well the program was delivered according to the pre-specified core elements that
manifest the program theory and philosophy. An essential step for determining fidelity is defining the core elements of the program. An example of core elements could be teachers have delivered the contents of the health education lessons correctly or teachers have provided appropriate examples to facilitate students’ understanding of complex concepts. Fidelity is often difficult to measure because it requires evaluators to have an in-depth understanding of the program theory and philosophy in order to develop appropriate instruments. Also, the standard of quality appears to be a subjective perception, which may vary among individuals. A relatively objective approach to assessing fidelity is through direct observation. Using structured observation guides developed by the program designers, trained observers rate the existence of pre-specified core elements. Another approach is to ask implementers to self-report how well the core elements were carried out through surveys and interviews.

Dose received

Dose received is also called participant exposure, responsiveness, engagement, or satisfaction. It measures how well the program was received or absorbed by the participants. Specifically, dose received is defined as the extent to which participants have actively engaged in the program activities, and their enjoyment of or satisfaction with the program. An example of dose received could be 60% of the participants have used pedometers to self-monitor their daily physical activity or 80% of the participants reported enjoying the intervention sessions. It could be measured by asking participants to self-report their awareness of the program components, or rate their enjoyment of and
satisfaction with the program. Other studies have also assessed participants’
engagement and enjoyment during the intervention sessions using direct
observations.\textsuperscript{57}

Reach

Program reach is how much of a program was attended or participated in
by the participants. It is often expressed as the proportion of the target population
who attended the intervention activities.\textsuperscript{65} For example, 80% of participants
attended the health education lessons or 60% of the parents showed up in the
family events. The most frequently used measure of program reach is to have
implementers complete the registry and attendance logs.

Additional aspects

There are other aspects of levels of implementation in the literature
including, program differentiation (the extent to which the program theory and
practices can be distinguished from the existing program in the setting);
monitoring of comparison condition (describing the nature and monitoring the
amount of services received by the comparison conditions);\textsuperscript{66} and adaptation
(changes made in the original program during implementation to better meet the
needs of the implementing setting).\textsuperscript{3} As the relationship of these additional
aspects with program outcomes are less clear in the literature,\textsuperscript{3,7} the focus of the
proposed study will be on dose delivered, fidelity, dose received, and reach.

Researchers are recommended to measure all four aspects because each
aspect has a unique impact on program effectiveness. An optimal dosage is
necessary but not sufficient for a program to produce the intended program
outcomes. If a program delivered with high dosage but zero fidelity (quality), it is unlikely to produce the desired outcomes. For example, a health education intervention aimed at increasing students’ health literacy. A teacher implemented all of the required health education lessons to the students but the information being delivered was about Mathematics. In this case, the researchers should not expect any significant intervention effects on the intended outcomes. Even if the evaluation did detect significant changes in the program outcome (students’ health literacy) the observed effects are unlikely to be attributed to the intervention.

If a well-designed program, with a high degree of dose delivered and fidelity, fails to engage the participants (low dose received), it is unlikely to achieve the intended outcomes. For example, students shown up for a mandated program that aimed at preventing drug abuse in teenagers, but they did not engage in any activities or complete any assigned tasks. Without active engagement, the likelihood that the participants have processed and internalized the intervention materials is low. As a result, the intended behavior changes are less likely to occur. Additionally, many public health programs aim at influencing population health. To produce a significant impact on public health, an effective, well-delivered program has to reach a sufficient number of participants (Impact=Efficacy/Effectiveness x Reach). In an extreme scenario, an effective intervention (100% efficacy) could have zero impact if it failed to attract any participants (0% reach). Thus, program reach should also be considered when evaluating a program.
Aspects of implementation levels: Environmental interventions

Unlike individually-oriented interventions, levels of implementation of environmental interventions are conceptualized to be influenced by only two aspects: dose delivered and fidelity. Environmental intervention emphasizes installations of programs and practices at the organization level that can change the context where the target population spends time. Once installed, every individual who spends time in the intervening setting are exposed to the intervention. As such, the reach of an environmental intervention is considered to be 100%. Dose received can be measured as the extent to which the target population is aware of the environmental change. Since awareness is not necessary for initiating behavior changes, dose received is of less concern to levels of implementation of environmental interventions.

Additionally, dose delivered and fidelity are often united as one aspect in environmental interventions. In individually-oriented interventions, an implementer can deliver a program with high dosages but with zero quality. Thus, it is important to measures them as two distinct aspects. In environmental interventions, the goal is to install a “complete set of appropriate environmental elements based on the conceptual framework”. The completeness (dose delivered) and appropriateness (fidelity) have to go hand in hand. For example, an environmental intervention aimed to increase preschoolers’ physical activity levels by providing more portable physical activity equipment in the preschool settings. In this case, dose delivered is defined as the extent to which preschoolers are provided with the equipment; while fidelity is defined as whether
the preschoolers are provided with the equipment. As such, it may be difficult to distinguish between dose delivered and fidelity. Considering the difference in conceptualization, interventions that consist of both individually-oriented and environmental components should measure levels of implementation of each component separately.

This section overviewed the terminology and measures of different aspects of implementation levels. In the following section, we will first summarize the findings on implementation of preventive interventions for youth because many existing conceptual frameworks were generated from prevention programs targeting youth, such as substance abuse, mental disorder, or smoking cessation. Then, we will review the findings from studies specifically focusing on implementation of youth physical activity interventions.

**Implementation research in preventive interventions for youth**

*Prevalence of assessing levels of implementation*

Although implementation research is much more advanced in preventive interventions for youth, studies that have measured and reported their implementation levels remain low. Moncher and Prinz\(^\text{71}\) reviewed 359 treatment outcome studies published in clinical psychology, psychiatry, behavior therapy, and family therapy journals from 1980 to 1988 and they found only 18.1% assessed levels of program implementation. Gresham and colleagues\(^\text{72}\) evaluated 181 behavioral interventions published between 1980 and 1990 and showed that only 14.9% of the studies have measured levels of implementation. Dane & Schneider\(^\text{66}\) reviewed 162 children’s preventive interventions published
between 1980 and 1994 and found that only 24% of the outcome studies documented levels of implementation.

For studies that have measured levels of implementation, they mostly focused on dose delivered and/or fidelity. Dane & Schneider\textsuperscript{66} reviewed 162 children’s preventive interventions and found that while 39 studies have assessed at least one aspect of levels of implementation, dose delivered and fidelity were measured the most often. Domitrovich and Greenberg\textsuperscript{59} reviewed 34 preventive interventions that were found to be effective in improving specific psychological symptoms or risk factors of mental health disorders in children aged 5 to 18 years. Seventy-six percent of the programs documented levels of implementation but only 21% of them have examined two or more aspects. Among the evaluated aspects, fidelity (59%) and dose delivered (33%) were evaluated most often, but dose received was only assessed in 6% of the studies. Dulark and DuPre\textsuperscript{3} examined the impact of levels of implementation on program outcomes by reviewing the evidence from five relevant meta-analyses and 59 empirical studies. Findings from the 59 empirical studies showed that the majority of studies only assessed one aspect of levels of implementation and 31% evaluated at two or more aspects. Fidelity (62%) was the most frequently measured aspect followed by dose delivered (49%).

\textit{Associations between levels of implementation and program outcomes}

Prevention research has clearly demonstrated that implementation is variable across individual implementers and implementing organizations.\textsuperscript{56} Dulark and DuPre\textsuperscript{3} found that levels of implementation varied by 20% to 40%
across implementers in the same study. Dusenbury and colleagues\textsuperscript{7} reviewed the evidence of levels of implementation of school-based drug abuse preventive interventions from published review studies. They found that the degree of dose delivered ranged from 44% to 75%; and the degree of fidelity varied from 16% to 42% across implementing organizations.

Several review studies\textsuperscript{3,56,59,66} have examined the linkage between levels of implementation and program outcomes. Dane & Schneider\textsuperscript{66} found some evidence in support of a positive association between fidelity and outcomes in youth substance abuse programs. Domitrovich and Greenberg\textsuperscript{59} concluded that dose delivered and fidelity were positively associated with outcomes in mental health preventive interventions for school-age children. In the review of Dulark and DuPre,\textsuperscript{3} evidence from meta-analyses showed that promotion and preventive interventions with higher levels of implementation obtained two to 12 times higher effect size than those with lower levels of implementation. Among 59 intervention studies, 76% showed positive associations between levels of implementation and program outcomes.

Due to heterogeneity in study design, measurement of implementation, and inconsistent use of terminologies, there is a lack of evidence about the level of implementation that is necessary for generating and maximizing program outcomes. \textsuperscript{3,56,59,66} Durlak et al.\textsuperscript{3} was the only review article that has suggested a threshold for obtaining intended program outcomes. Their findings indicate that positive results appear to show in programs that have achieved at around 60% of implementation.
On the other hand, some empirical studies suggested that there may not be a clear cut point for yielding positive program outcomes. For some programs, positive effects may only occur when certain levels of implementation are attained. For other programs, positive effects may still show with a low level of implementation, but the effects become stronger as levels of implementation increase.\textsuperscript{56}

Botvin and colleagues\textsuperscript{73} assessed the effects of a school-based drug preventive intervention. The results showed a significant intervention effect on reducing cigarette smoking in the intervention group, but insignificant effects on alcohol consumption and marijuana use. However, a secondary analysis was conducted on a subsample of students who received a high-fidelity program and the intervention effects on reducing the prevalence of alcohol consumption and marijuana use became significant. James and colleagues\textsuperscript{74} evaluated the effectiveness of an HIV and AIDS life skills program for secondary students. The analyses of the total sample did not detect any significant intervention effects on students’ safe sex practices (i.e., condom use, sexual intercourse); however, when stratifying students according to teachers’ levels of implementation, the authors found that students in the fully-implemented group had more appropriate perceptions of sexual behaviors, were less sexually active in the past six months and used more condoms at last intercourse than those in the partially-implemented group and the control group. These two studies demonstrate that significant intervention effects may only be observed in well-implemented interventions.
Pentz and colleagues\textsuperscript{75} compared the effects of another school-based drug preventive intervention among students in the high-implementation group, low-implementation group and control group. The results showed that students’ cigarette smoking and marijuana use were significantly lower in both high-implementation and low-implementation group than the control group. However, the effects were stronger in the high-implementation group than the low-implementation group. Compared to the control group, students in the high-implementation group reported 43\% less cigarette smoking, 34\% less alcohol consumption, and 33\% less marijuana use. For the low-implementation group the reductions were 18\% for cigarette smoking, 25\% in alcohol consumption, and no significant changes in the level of marijuana use. McGraw and colleagues\textsuperscript{76} examined the effects of a school-based health promotion program and found that every 1\% increases in teachers’ levels of implementation accounted for a 10\% increment in students’ self-efficacy in choosing healthy foods. These two studies show that even low levels of implementation can yield some beneficial effects, but higher levels of implementation may produce stronger effects.

In summary, the prevalence of assessing levels of implementation is low. For studies that have measured it, few of them have included comprehensive measures and the majority of studies focused only on dose delivered and fidelity. There is consistent evidence in support of a positive association between levels of implementation and program outcomes. However, the levels of implementation that are necessary to yield significant beneficial effects remain unclear.

*Factors influencing levels of implementation*
Several reviews\textsuperscript{3,6,7,59,62} of preventive interventions for youth (e.g., mental health, drug abuse, smoking cessation) have generated a list of candidate factors that could influence levels of implementation. There is a general consensus among these reviews that levels of implementation are influenced by factors operating at multiple levels.

\textit{Macro level}

The associations between macro-level factors and levels of implementation are briefly described as follows. A program that is consistent with federal mandates and professional practices will be easier to establish mutual interest among stakeholders, thus leading to better implementation outcomes.\textsuperscript{5,16} Additionally, partnerships and coalitions with other agencies or organizations are found to be associated with better implementation outcomes.\textsuperscript{5,16} For example, partnerships with universities can provide communities with the knowledge to select an appropriate program (e.g., evidence-based program) and equip them with the capacity to implement the program with quality. Because macro-level factors (e.g., policies and community-university partnerships) are considered to be the most distal to levels of implementation and are difficult to modify, the proposed study will specifically focus on factors at the other levels.

\textit{Organizational level}

\textit{Organizational characteristics}

An important component of program implementation is the delivery system—the implementing organizations. Several organizational characteristics have been shown to be associated with levels of implementation, including size,
organizational complexity, 6,77 organizational cultures, organizational functioning, work climate, and absorptive capacity. 3,6

The size of an organization can effect levels of implementation either positively or negatively. 78 Large organizations usually have more resources (both monetary and human) to support the implementation of interventions. However, other studies submit that large organizations may be less flexible and have lower willingness to change which may hinder implementation effectiveness. Organizational complexity refers to the degree in which an organization possesses members with a high level of knowledge and a great variety of expertise. 77 Complex organizations are more likely to have higher levels of implementation as their members are more knowledgeable in understanding of key concepts of the intervention and more skillful to implement the intervention.

Organizational cultures refer to the norms, values, and shared beliefs of the organization. 3,59 An organization with norms that encourage changes and integration of new programming are more likely to support intervention implementation 3,59 A well-functioning organization, which is reflected by effective intra-organizational communication, shared decision-making and clear procedures for role and responsibility assignments, is associated with more effective implementation. 3,6 Work climate is the perception of organization’s members of their workplace environment. 79,80 A positive work environment where organization’s members have trusting relationships, are collegiate and supportive to each other is also positively associated with intervention implementation. 78,81 Absorptive capacity is an organization’s ability to find, interpret, and use new
knowledge. Organizations with higher absorptive capacity are more likely to have better implementation. It is possible that these organizations can better capture the core ideas of the intervention, link it with the existing knowledge base, and integrate it into their existing practices.6

**Administrative leadership and support**

Administrative leadership and support have consistently shown to be the facilitating factors of program implementation.3,6,77 Effective administrative leadership that is manifested by setting priorities, establishing consensus, providing incentives and managing the overall implementation process are associated with higher levels of implementation. Furthermore, specific leadership styles also have implications on levels of implementation. Democratic and coaching styles that forge implementer’s input and value implementer’s development are more effective than other styles (e.g., coercive) in promoting levels of implementation.82 Strong administrative support is reflected by actively engaging in the intervention planning, participating in the training, establishing policies to facilitate program implementation as well as formally committing staff members and administrators to intervention-related activities. These processes can create a supportive environment that encourages the implementers to devote time and effort to deliver the intervention with quality.3,6,77

**Resources**

The investment of appropriate resources is important to ensure the actual operations of the intervention.3,6 Resources can include monetary incentive, dedicated staff time for the intervention-related activities, space, equipment, and manuals. One important resource is the presence of program champions.3,6,77
The program champion is someone who is knowledgeable about the local context, and trusted and respected by organizational members and administrators. They play an imperative role in the implementation processes by ensuring program-context fit, developing cross-function coalitions within the organization, establishing informal systems to monitor and support implementation, and negotiating solutions to any problems that arise.

Provider level

Many preventive interventions for youth were implemented by a change agent (e.g., school teacher) who is responsible for delivering the essential elements of the program to the target audience (e.g., students). Therefore, the characteristics of the implementer are likely to play an imperative role in the success of program implementation. Levels of implementation are found to be associated with implementers’ professional characteristics, psychological characteristics, and perceptions of the intervention. Professional characteristics refer to the education level, skills, and experiences related to implementing the intervention. These professional characteristics may also impact the implementers’ psychological characteristics, (i.e., self-efficacy). Self-efficacy is the implementers’ self-confidence in their ability to deliver the intervention. Implementers who have higher levels of education, skills and more successful experience relevant to implementing the interventions tended to have higher levels of self-efficacy. Self-efficacious implementers are more likely to implement the program with quality as they are more comfortable and less anxious in doing so. Additionally, implementers who perceived the intervention
as relevant to organization’s needs and beneficial to the organization, participants or themselves tended to be more motivated to deliver the program, thus leading to better program implementations.\(^3,^6\)

*Program level*

**Program characteristics**

Program-level factors are the most proximal to levels of implementation. Effective program implementation has been associated with the following program characteristics: adaptability, compatibility, and complexity. Adaptability is the extent to which the intervention can be adapted and modified to fit the needs of the implementers and the local contexts. Compatibility is the degree to which the program is perceived as consistent with the values, experiences, and needs of the adopting organization.\(^7,^7\) An adaptable program can also increase its compatibility because adaptation makes the program more sensitive to the organization’s culture and easier to integrate into an organization’s existing routines. As such, implementers will feel more familiar with and have lower resistance to implement the intervention,\(^6,^7\) thus resulting in higher levels of implementation. Complexity of a program refers to the ease of use.\(^7\) Compared to a simple intervention, a complex intervention consists of many elements requires special skills, and large investments of time and human resources is less likely to achieve adequate levels of implementation. Possible reasons could be that complex interventions are less likely to be perceived as effective by the implementers, thus diminishing their motivation to deliver the intervention. Also,
implementers may be discouraged to implement the intervention as they perceived more barriers to carry out complex interventions.\textsuperscript{7}

\textit{Implementation process}

Implementation process is defined as a specific set of activities designed to put an intervention into practice.\textsuperscript{62} It could include activities such as training staff members on the intervention delivery, forming an implementation team, and developing an implementation plan.\textsuperscript{3,6} The involvement of stakeholders (e.g., organizational administrators and implementers) in the development and planning processes is found to be associated with levels of implementation. Involving the key stakeholders in the planning process can facilitate their understanding of the intervention, increase community ownership, and enhance perceived fit between the program and local contexts. As a result, they would have a higher acceptance to the program, thus increasing the likelihood of the program being implemented effectively. The quantity and quality of the facilitation strategies are also found to be related to program implementation. Facilitation strategies such as provision of manual, guidelines, training, ongoing technical assistance and feedback are important for optimizing and standardizing what is being implemented.\textsuperscript{3,6} Among those strategies, providing trainings and ongoing technical support are viewed as the essential elements for achieving adequate levels of implementation. Trained and well-supported implementers tended to be more confident and competent in delivering the intervention more completely.\textsuperscript{78,84,85} Maintaining clear communication during the implementation phase has also been related to implementation success.\textsuperscript{6}
In summary, this section summarized measures of implementation levels in preventive interventions for youth. This section also identified a list of candidate factors that could influence implementation levels in preventive interventions for youth. However, the interrelationships among these factors and their associations with the four aspects of levels of implementation have not been identified. Also, the numbers of factors that are necessary to predict successful implementation has not been established. Noteworthy, whether these findings taken from the literature of preventive interventions can be generalized to the context of physical activity interventions require further investigation. Therefore, the following section will specifically focus on summarizing the measures and influences of implementation levels in youth physical activity interventions.

**Implementation research in youth physical activity interventions**

Program implementation is an emerging topic in youth physical activity interventions. However, there has not yet been a systematic summary of evidence on implementation of youth physical activity interventions and factors influencing their levels of implementation. Therefore, a review was conducted to summarize current evidence to help inform future intervention studies about issues related to implementation and the finding is presented in this section.

This review specifically focused on the following research questions:

1) What is the prevalence of measuring levels of implementation?

2) How were levels of implementation measured?

3) What are the observed levels of implementation and there relations to program outcomes?
4) What are the factors that influence levels of implementation?

_Literature search and inclusion criteria_

Relevant intervention studies were identified from three sources. First, 451 youths’ physical activity intervention studies were obtained from 14 reviews published in the past 5 years\textsuperscript{23,31,40,86-96} and scanned for eligibility. Second, computer searches were conducted of MEDLINE, ERIC, PsychInfo, and Web of Science to identify articles published between January 1, 1990 and May, 2014. The search period was set to begin from 1990 because this is the time when published public health intervention studies have begun to include extensive process evaluation components\textsuperscript{97}. The following search strings were used:

\texttt{(physical activity OR physical education OR motor activity OR exercise OR physical fitness) AND (intervention stud* OR randomized controlled trial OR cluster randomized trial OR group randomized trial OR quasi experimental stud*) AND (school OR after-school OR preschool OR elementary school OR middle school OR high school) AND (preschooler OR children OR adolescents OR youth) AND (implementation OR monitoring OR dose delivered OR dose received OR fidelity OR reach OR process evaluation OR train the trainer OR lesson learned) NOT (review OR meta analyses) NOT (feasibility studies OR acceptability OR pilot).}

Third, reference lists of studies identified from source 1 and source 2 were inspected.

To be included, studies have to meet the following inclusion criteria:

- Published between 1990 and May 2014 in a refereed English language journal
• Be an outcome evaluation that included physical activity levels as one of the outcomes.
• Targeted to children and adolescents ages 2.0 to 18.9 years
• Be an intervention focusing on the initial implementation stage (i.e., a newly adopted intervention is being used for the first time; efficacy trials) and full implementation stage (i.e., an intervention is being integrated into the organizations and put into full operation with full staffing complements and full client loads; effectiveness trials).
• Measured and reported at least one aspect of implementation levels (i.e., dose delivered, fidelity, dose received, and reach).

Interventions limited exclusively to mass-media campaigns, video instructions and Internet-delivered programs with limited involvements of human change agents were excluded. Interventions limited exclusively to policy implementation were excluded because their implementations are influenced by factors that are very different from implementing structured programs or practices. Lastly, studies focused on program exploration (e.g., formative studies), program installation (e.g., pilot studies) and sustainability were also excluded.

Coding

Key information of eligible articles was extracted using a structured form, including authors, publication year, area of focus, study settings, types of intervention (i.e., individually-oriented or environmental), types of implementer (i.e., site personnel or external personnel), aspects of implementation levels (i.e., dose delivered, dose received, fidelity and reach), data collection methods,
observed levels of implementation, changes in program outcomes (i.e., physical activity levels) and factors that influence levels of implementation.

*Calculation of observed levels of implementation*

The observed levels of implementation were calculated based on the information provided in the articles. Observed degrees of each implementation aspect were summarized for all implementing organizations combined.

- Dose delivered was calculated as the percentage of the total intervention components that were delivered by the implementers across all implementing organizations.

- Fidelity was calculated as the percentage of total required core elements that have been adopted across all implementing organizations. To calculate the observed degree of fidelity, the first step was to define the core elements that reflected the spirit or philosophy of the interventions. If a study did not explicitly state its core elements, the elements that were measured and presented in the article would be treated as the core elements. The second step was to set a criterion for determining existence a core element across the implementing organizations. An element would be classified as presence if it was adopted by at least 75% of the participating organizations. This arbitrary cut point is consistent with the findings from Durlak\(^3\) that an implementation of at least 60% is needed to produce positive program outcomes, but greater than 80% is rare. A cut point of 75% (between 60% and 80%) is considered reasonable.
• Dose received was calculated as the percentage of the total targeted participants across all implementing organizations that have reported themselves enjoying, being engaged with or being satisfied with the intervention.

• Reach was calculated as the percentage of the total participants across all implementing organizations who have attended or shown up for the intervention activities.

Results

How prevalent level of implementation was measured

Article selection is illustrated in Figure 2. Among the 179 identified articles, 43 articles\textsuperscript{32,34,64,70,98-136} met the inclusion criteria. The 43 articles described findings of 35 unique physical activity interventions for youth. In other words, only 23% of the identified intervention studies have verified levels of implementation.

Among the four implementation aspects, the most frequently measured aspect was reach (71%), followed by dose delivered (64%), dose received (58%), and fidelity (26%). Only 13% of the studies measured all four aspects; 26% evaluated three aspects; 35% of the studies measured two aspects (mainly dose received and reach or dose delivered and reach), and 26% of the studies measured only one aspect (mainly dose delivered).

How level of implementation were measured in individually-oriented intervention

Dose delivered

There are five process evaluation questions related to dose delivered (Appendix A).
• What is the actual number of intervention components that were delivered by the implementer? For example, the total number of intervention activities organized,\textsuperscript{100,113} or the total number of health lessons taught.\textsuperscript{137}

• What is the proportion of the total intervention components that were completed by the implementers? It is often calculated by dividing the number of completed components by the total required components.\textsuperscript{119,126}

• What is the duration of the intervention sessions? Sample measures could be, total minutes of structured physical activity sessions delivered by the implementers to children in the childcare centers\textsuperscript{135}, or number of minutes per day that the prescribed intervention was used by the teachers\textsuperscript{114,127}.

• What is the frequency with which the intervention components/activities were implemented? For example, number of days per week of teacher-led physical activity breaks,\textsuperscript{137} or number of days per week of the physical activity component being implemented in a classroom.\textsuperscript{133}

• Did the implementers comply with the prescribed dosages? For example, whether or not teacher complied with the recommended activity break once per day over the intervention period.\textsuperscript{99}

Among the 20 studies that have assessed dose delivered,\textsuperscript{34,99,102,103,106,109,111-115,117-119,123,126,128,129,133,134} twelve studies (60%) were based on self-reports (i.e., checklist, evaluation form, survey, interview, and focus group),\textsuperscript{99,102,106,109,113-115,117,128,129,133,134} four studies (25%) used direct observation;\textsuperscript{34,103,112,126} and two studies used program records.\textsuperscript{111,123} Another two studies (10%) employed more
than one measure to assess dose delivered (e.g., observation and survey or observation and program documentation).\textsuperscript{118,119}

Fidelity

There are two process evaluation questions related to fidelity (Appendix A).

- To what extent were the program objectives implemented as intended in general? For example, the Switch Play study\textsuperscript{108} determined the program by asking the implementers, in general, whether or not the planned objectives of the lessons were achieved.

- To what extent was the program implemented in accordance with its intervention philosophy as defined by a set of core element? For example, the Trial of Activity for Adolescent Girls (TAAG)\textsuperscript{34} assessed the fidelity of the TAAG PE component based on a set of core elements. For example, group sizes were appropriate for activity, and teachers used strategies to minimize management time.

Among the eight studies that have assessed fidelity,\textsuperscript{34,106,108,112,114,118,119,126} 50\% of the studies used direct observation\textsuperscript{34,112,119,126} and the remaining studies asked implementers to self-report their compliance to those core elements.\textsuperscript{106,108,114,118}

Dose received

Three process evaluation questions were related to dose received (Appendix A). Were children satisfied with the program? For example, the Ready. Set ACTION! study\textsuperscript{121} asked participants to rate their overall satisfaction with the program.
• Did children enjoying the program? For example, the Child and Adolescent Trial for Cardiovascular Health (CATCH)\textsuperscript{138} determined participants’ enjoyment levels based on their facial expression (e.g., smiles and laughers) during the CATCH PE sessions.

• Were children actively engaged in the program? For example, the Switch Play study\textsuperscript{108} determined participants’ engagement based on students’ compliance with in-class tasks.

Among the 18 studies\textsuperscript{34,101,103,105,108,109,116,118-122,126-128,132,133,139} that have measured dose received, 15 studies (88\%) asked the participant to self-report their satisfaction, enjoyment and/or engagement in the intervention.\textsuperscript{34,101,103,108,109,113,115,116,118,120,122,127,128,132,133,139} Four studies (25\%) employed a proxy measure was used in four studies\textsuperscript{108,118,128,132} and three studies used direct observation.\textsuperscript{34,119,126}

Reach

The percentage of the total participants who attended or showed up for the intervention activities was the main process evaluation question for program reach (Appendix A).

Among the 22 studies that have measured program reach, most of the studies did not specify the approach for measuring or recording program attendance.\textsuperscript{34,98,101-103,105,106,108,111-113,115,118-122,125,126,128,131,134} However, it is reasonable to assume that most studies determined participants’ attendance based on program records (e.g., a registry or an attendance log). Five studies
consisted of a family component and their reaches were assessed by parent surveys (e.g., whether or not parents received the newsletter.)\textsuperscript{108,113,119,121,126}

Aspects of implementation levels and their association with program outcomes

Dose delivered

Seventeen out of the 20 (85\%) studies have provided sufficient information for calculating the observed degrees of dose delivered.\textsuperscript{34,99,102,103,109,112-115,117-119,123,126,128,129,133} Four studies delivered >80\% of the prescribed dosages;\textsuperscript{112,118,119,126} 11 studies implemented 60\% to 80\%,\textsuperscript{34,99,102,103,109,113-115,123,128,129} and two studies completed 22\% to 39\% of the prescribed dosages\textsuperscript{23,40} (Appendix B).

In terms of the association pattern between dose delivered and program outcomes, a positive effect was observed in studies that achieved around 70\% of the prescribed dosages. Two studies\textsuperscript{99,123} have statistically tested the association between dose delivered and changes in children’s physical activity levels. Patrick and colleagues\textsuperscript{123} found a positive association between dose delivered and children’s physical activity levels, but the effects were gender-specific. For boys, the likelihood of meeting the physical activity guideline was higher for children in the high dose group (≥80\%) than those in the low dose group (<80\%). No significant association was found in girls. Erwin and colleagues\textsuperscript{99} also found a positive association between degree of dose delivered and children’s step count per day. Children of teachers who complied with the recommended one activity break per day over the intervention period accumulated significantly more daily steps than those in the non-compliant and control groups. Besides, changes in
children’s daily step did not differ significantly between the non-compliant and control groups.

Fidelity

Six out of the eight studies (75%) have provided sufficient information for calculating the observed degrees of fidelity.\textsuperscript{34,112,114,118,119,126} Four studies\textsuperscript{112,118,119,126} have reported degrees of fidelity at ≥80% and two studies\textsuperscript{34,114} reported a low degree of fidelity (33% to 43%). A specific association pattern between degree of fidelity and changes in children’s physical activity levels was not observed between all eight studies. No studies have statistically tested the association between fidelity and program outcomes (Appendix B).

Dose received

Fifteen out of the 18 studies (83%) have provided sufficient information for calculating the degree of dose received.\textsuperscript{34,101,103,105,108,109,114,118,119,121,122,126,132,133,139} The majority of studies have achieved a moderate to high degree of satisfaction and enjoyment (60% to 99.5%. Also, a moderate to high degree of engagement (50% to 90%) were attained in most studies; and a low degree of engagement was observed in one study\textsuperscript{140} (11%).

No specific association patterns were observed between dose received and changes in children’s physical activity levels. Of the four studies\textsuperscript{103,105,116,131} that statistically tested the effect of dose received on children’s physical activity levels, all of them found a null association. However, three studies showed a positive association between dose received and other program outcomes.
Marcoux and colleagues\textsuperscript{103} found that participants’ overall engagement in the intervention was significantly associated with the improvement in children’s attitude and intention towards physical activity. Gentile and colleagues\textsuperscript{116} showed that participants’ and parents’ engagement in the family component was positively associated with children’s levels of fruit and vegetable consumption. In a father-child dyad study,\textsuperscript{105} the fathers’ engagement in the exercise self-monitoring component has a significant positive correlation with child’s weight loss (Appendix B).

Reach

Twenty out of the 22 studies have provided sufficient information for calculating the observed degree of reach\textsuperscript{34,98,101-103,105,106,108,111-113,115,118,119,121,122,125,126,131,134}. The reach of the overall program or the key program component (e.g., curriculum component, PE lessons) was >60% in 12 studies, 40% to 60% in six studies,\textsuperscript{102,111,112,121,125,134} and 5% to 20% in two studies.\textsuperscript{106,115} Additionally, several studies have measured the reach of a sub-component, such as family components and refresher sessions. Compare to the key components, the reach of these sub-components were lower and had a greater variability (16 % to 82%). No studies have statistically tested the association between program reach and program outcomes (Appendix B).

How level of implementation were measured in environmental intervention

As mentioned earlier in this proposal, levels of implementation of environmental interventions mainly consist of two aspects: dose delivered and fidelity. The present review identified four environmental interventions\textsuperscript{110,124,130,135}. 
All of them have measured dose delivered and fidelity (100%); one study\textsuperscript{130} has additionally measured dose received (25%). Three out of the four environmental interventions (75%) have provided information to calculate the degree of dose delivered and fidelity and it ranged from 43% to 80%. The study\textsuperscript{130} has assessed dose received reported that 100% of the parents were satisfied with the program (Appendix C).

Due to a small number of studies, the pattern of association between dose delivered, fidelity and program outcomes cannot be determined. A study by Saunders and colleagues\textsuperscript{124} was the only one that has statistically tested the associations of dose delivered and fidelity with program outcomes. They found that the prevalence of girls participating in vigorous physical activity was higher in the high-implementation group than the control group. They also found a significant dose response relationship. Girls’ vigorous physical activity increased proportionally to the levels of implementation, with the control group showed the smallest improvement in vigorous physical activity at post-intervention and the high-implementation group showed the highest improvement (Appendix C).

**Factors influencing levels of implementation of youth physical activity interventions**

Few physical activity interventions targeting youth have measured levels of implementation. Studies that have examined factors that influenced levels of implementation are even fewer. In the present review, only four out of the 179 included studies (2%) have examined this issue.\textsuperscript{112,113,127,129}
Zarrett and colleagues\textsuperscript{141} studied the factors influencing levels of implementation of the Active by Choice Trial (ACT). The ACT was a 17-week randomized controlled trial (RCT) which aimed to use a community-based after-school program to promote physical activity in underserved adolescents. The intervention staff (external personnel) implemented three program components: homework/snack, a physical activity component that included activities which the students selected each week of moderate and vigorous intensity (60 minutes), and a curriculum component in which intervention staff taught participants behavioral skills and motivational strategies to increase their physical activities with friends and at home. To promote implementation, the intervention staff provided training, technical assistance and supervision. Factors associated with levels of implementation were obtained by interviewing the intervention staff members. Data were summarized into themes. A key organizational factor that could improve levels of implementation was availability of space. At the implementer level, interactions with program leaders and perceived benefits of the program were associated with more effective implementation. At the program level, increased novelty and challenge levels of intervention games, more training for implementers to manage participants’ resistance and disruptive behaviors were suggested to promote levels of implementation.

DeMeij and colleagues conducted a comprehensive process evaluation to understand factors that influence levels of implementation of the JUMP-in study.\textsuperscript{113} The JUMP-in was group RCT that aimed to promote sports participation and daily physical activity among primary school children in Amsterdam by
intervening with children’s points-of decision as well as changing school social and physical environments. School teachers implemented six intervention components, including increasing school sports activities, monitoring children’s physical activity, providing in-class exercise sessions, providing lessons aimed at increasing participants’ awareness, organizing a parental information session and organizing activity weeks. Process data were collected from multiple sources (i.e., participants, implementers, school directors and the intervention team) using multiple methods (i.e., observations, questionnaires, structured in-depth interviews and review of documents). Data were summarized into themes. At the organizational level, the following organizational characteristics were suggested to be associated with more effective implementation: strong organizational commitment and motivation to comply with shared goals, clear organizational hierarchical structures, effective communication, clear protocols within and between organizations, and willingness to integrate the innovation into current practices. At the implementer level, implementers’ commitment and motivation to achieve goals, perceived compatibility between intervention tasks and their regular tasks, and the perceived benefit and perceived importance of the program were found to be associated with higher levels of implementation. At the program level, well-defined program components, a good fit between the program and local context, ease of integrations to the current routine, and involvement and support from experts were suggested to promote levels of implementation.
Naylor and colleagues examined the factors influencing levels of implementation of the Action Schools! BC model (AS! BC). AS! BC was a 16-month RCT that aimed to use a multi-component school-based intervention to increase students’ participation in moderate intensity physical activities to 150 minutes per week. To achieve the study goal, school teachers were trained to incorporate more physical activity opportunities across six action zones: school environment, physical education, classroom, family, and community, extracurricular and school spirit. Specially, the goal of 150 minutes moderate intensity physical activity was supposed to be achieved by providing students with at least 15 minutes of additional physical activities each school day in the classroom (75 minutes per week), and the remaining 75 minutes were made up by activities in other action zone throughout the day. To support implementation, intervention team provided training, technical assistance, and access to further training on professional development. Semi-structured interviews and focus groups were conducted with school administrators, implementers (i.e., teachers) and school facilitators to identify factors that influenced levels of implementation. Data were summarized in themes. At the organizational level, administrative support (i.e., permission to devote class time to physical activity), provision of resources, and availability of space were important factors for effective implementation. At the implementer level, implementers’ priority of physical activity, perceived benefits to the children, and workload were related to levels of implementation. At the program level, flexibility of the program, and starting from a small scale were the key facilitators.
Gibson and colleagues\textsuperscript{127} investigated the barriers negatively influencing levels of implementation of the Physical Activity Across Curriculum (PAAC) intervention. PAAC was a cluster RCT that aimed to prevent excessive weight in elementary school children. To achieve the study goal, school teachers were asked to integrate 90 minutes of moderate intensity physical activities into the classroom curriculum. To promote implementation, intervention team provided one training session and ongoing technical assistance. Focus interviews were conducted at the end of the intervention, in which school teachers were asked to discuss the barriers of their implementation. At the program level, the misfit between designed intervention activities and the local classroom context, and insufficient demonstration on how to integrate physical activities into the classroom were suggested to be related to less effective implementation.

\textit{Summary and discussion of the review findings}

Implementation research in youth physical activity interventions is still developing. This review found that only 23\% of the studies have measured levels of implementation, which was similar to several other studies (ranged from 15\% to 24\%)\textsuperscript{66,71,72} Among the four implementation aspects, the most frequently measured aspect related to levels of implementation was dose delivered, while fidelity was assessed the least often. These results are consistent with the findings reported in Dane and Schneider’s review\textsuperscript{66} that dose delivered (54\%) was the most frequently measured aspect. However, the present findings are inconsistent with those reported by Durlak et al.\textsuperscript{3} and Domitrovich et al.\textsuperscript{59} who found fidelity was assessed the most often. The inconsistency could be due to
variations in terminology. In the current review, the term fidelity specifically refers
to quality of the implementation as defined by a set of core elements (e.g.,
implementer enthusiasm). In the literature of preventive interventions for youth,
the fidelity has been used as an umbrella term (e.g., implementation fidelity)
which generally refers to the overall levels of implementation. Very often, studies
that measured implementation fidelity were only measuring non-fidelity aspects
(e.g., dose delivered). It is, therefore, important for the field to standardize the
use of terminologies to enable meaningful comparisons on levels of
implementation across studies. Similar to the findings in preventive interventions
for youth, the current review also observed a high variability in implementation
levels across studies. Among the eligible studies, the observed degrees ranged
from 22% to 97% for dose delivered; 33% to 98% for fidelity; 11% to 98% for
program engagement; 42% to 96% for participants’ enjoyment or satisfaction;
and 16% to 91% for reach. Additionally, variations in implementation levels were
also observed among implementers in the same study. For example, the
observed degree of dose delivered of the physical activity component among the
intervention schools varied from 29% to 100% in the JUMP-in study, and 32% to
100% in the ACT trial.

Although an insufficient number of eligible studies (n=4) have examined the
factors that influenced implementation levels of youth physical activity
interventions, the present finding illustrated that the influences of implementation
levels in youth physical activity interventions are not entirely the same as
preventive interventions for youth. For example, availability of spaces is a key
factor in facilitating the implementation of youth physical activity interventions. However, it is less of a concern in preventive interventions for youth because they are mostly curriculum-based and require fewer physical spaces. Teachers’ perceived priority of the target behavior (i.e., physical activity) may be more influential to physical activity interventions than for prevention programs. Physical activity has typically been perceived as less important than academic achievements or other risk behaviors (e.g., delinquent behaviors, smoking, substance abuse). Under the constraints of competing recourses, implementers who have less positive perceptions may choose to devote less effort to implementing physical activity interventions.

A notable gap in this body of literature is a lack of quantitative studies. In this review, all of the included studies used a qualitative method to examine factors associated with levels of implementation. Although these qualitative studies have provided valuable information, they did not provide information for determining the salience, interaction, or the direction of the relationships. Currently, only one relevant quantitative study is identified in the literature. Cardon and colleagues conducted a cross-sectional study to examine the factors associated with implementation levels of the Physical Activity Promotion Framework (PAPF) (i.e., providing sports during after-school and lunch break, developing active school yards or playgrounds, promoting active school commuting, developing health education policy, and organizing after-school sports and physical activities). The study conducted a survey in 226 primary and secondary schools in Finland regarding factors affecting the implementation success of PAPF. Regression
modeling was used to analyze the data. Among elementary schools, organizational factors including school size, traffic safety around school and administrative support for the framework significantly were found to be associated with levels of implementation. With regards to implementer characteristics, better knowledge of community schools, and higher awareness of current physical activity promotion projects were also found to be associated with higher levels of implementation. In terms of program characteristics, in-service training on school-community partnerships and whole-school physical activity promotions were the significant predictors of implementation levels. Among secondary schools, the significant predictors at the organizational level were urbanity and size, interests from school board in school-community partnership, administrative support for the PAPF and school priority of physical activity. At the implementer level, better knowledge of community schools was significantly associated with higher levels of implementation. At the program level, in-service training on school-community partnership and whole-school physical activity promotion were the significant predictors of implementation levels.

The lack of quantitative studies in this area could be because many studies did not collect data related to program implementation processes. As shown in this the present review, only 23% of the identified children’s physical activity interventions have measured and reported levels of implementation. Among them, only 2% of the studies have examined factors influencing levels of implementation. For studies that have examined the factors, analytical challenges may have hindered their ability to use a quantitative method. A
major analytical challenge is to obtain sufficient statistical power. To do so, intervention studies involving a large number of participating organizations with varying levels of implementation will be required. Unfortunately, collecting and analyzing process data in large-scale intervention studies is expensive and time consuming. Thus, an analysis of available data set by using alternative statistical methods that are not limited to the large sample assumption, for instance a Bayesian approach, can provide timely and valuable information to inform the design of future studies.

**Overview of Bayesian modeling**

In this dissertation, the analysis of the three proposed studies will be using the Bayesian approach. Therefore, this section provides an overview of Bayesian modeling. There are two dominant statistical modeling approaches: frequentist and Bayesian. The key difference between the two approaches is the interpretation on probability. The frequentist approach estimates the probability of the event (the unknown parameter (\( \theta \))) by observing its relative frequency in a hypothetical infinite replications of the study. Frequentists treat the data as random, which consists of a sampling distribution; while the parameters have fixed population values. For statistical inference, for instance null hypothesis statistical test is used. Assuming the null hypothesis that \( \theta = 0 \) is true, how improbablt is the observed \( \theta \) different from zero. When the parameter is assumed as a constant, it does not have a sampling distribution, and it can only be estimated as a single number. However, for a continuous variable, we know that the probability of \( \theta \) equals to any single number is always zero. In other words,
the estimated $\theta$ is always incorrect. To indicate the uncertainty of $\theta$, frequentists calculate the 95% confidence intervals (CI) to provide a possible interval that may contain the true parameter. The 95% CI is interpreted as 95% of the time from the sampling that the CI will capture the true parameter under the null hypothesis is true. From this perspective, the 95% CI does not give any information about the probability of the parameter falling in the interval. It only indicates the probability of that the parameter is in the interval is either zero or one.\textsuperscript{144-146}

The Bayesian approach estimates the probability of $\theta$ by observing the probability of the $\theta$ given the data $P(\theta | \text{data})$ which is called a posterior probability.\textsuperscript{147} Bayesians derive the posterior probabilities by combining observed data with prior information using Bayes' rule. Bayes’ rule is an equation: $P(\theta | \text{data}) \propto P(\text{data} | \theta)P(\theta)$. The equation states that the posterior probabilities of parameters given the observed data are proportional to (i.e., $\propto$) the probability of the observed data as informed by the parameters, which is the likelihood $P(\text{data} | \theta)$, multiplied by the prior probability of the parameter $P(\theta)$. The priors $P(\theta)$ is one’s belief on the probability of observing the parameter without giving any information about the data.\textsuperscript{146,147} There are two types of priors: non-informative or informative. The non-informative priors are referred to as diffused priors. This type of prior is appropriate for estimating parameters that we may not have enough knowledge about its shape and scale of the distribution. It is specified as $\theta$ follows an uniformed distribution Unif(a,b), where a and b are the boundaries of the possible range of $\theta$. With this prior, $P(\theta) \propto 1$. From a Bayesian
perspective, it is important to consider and incorporate our ignorance of the estimating parameter into statistical specification and it has impact on the precision of the estimation.\textsuperscript{146,147} Gustafson and colleagues\textsuperscript{3} have found that models that used a non-informative prior can be as accurate in predictive ability as models that incorporated a prior knowledge. The informative priors are typically used when we have enough prior information about the estimating parameter, where the information may come from expert opinion or empirical studies. For example, based on previous studies, we can specify that $\theta$ follows a normal distribution with a mean of 0.35 and standard deviation of 0.04

$$p(\theta) \sim \text{N}(0.35, 0.04).$$

In additional to treating the data as random, Bayesians also treat the unknown parameter $\theta$ as random variables to indicate their uncertainty about the parameter. As both parameters and data are assumed random, they can model the joint probability as a function of the conditional distribution of the data given the parameter, and the prior distribution of the parameters.\textsuperscript{146} As Bayesians assume the parameter has a probability distribution, when they sample from the posterior distribution of the model parameters, they can obtain its quantiles. From the quantiles, they can directly obtain the probability that a parameter lies within a particular interval (95\% posterior probability distribution (PPI)).\textsuperscript{146} The 95\% PPI tell us the probability that the parameter lies in the interval is 0.95, which is entire different interpretation from the frequentist.

The Bayesian approach is more appropriate than the frequentist approach for this dissertation for several reasons. The frequentist approach estimates the
probability of the parameters based on the assumption that the event is repeated a large number of times. In some situations, it is difficult or perhaps impossible to repeat the experiment many times or even conceive of repeating it. For example, the implementation of a newly adopted physical activity intervention is conceptualized as a one-time event. The Bayesian approach overcomes this issue by using a subjective probability estimated as an individual’s degrees of belief that an (one-time) event will occur. Additionally, the Bayesian approach is more suitable than the frequentist approach for constructing models in a developing field like implementation research in youth physical activity interventions. There is a lack of empirical studies that have examined influences of implementation levels of youth physical activity interventions. While there is existing empirical evidence; the studies often have small sample sizes when the organization serves as the unit of analysis. When using the frequentist approach, the unknown parameter is bounded by a fixed probability distribution (i.e., normal distribution under the central limited theorem). Extensive primary data are required to obtain sufficient statistical power in order to construct a stable, reliable predictive model. Alternatively, the Bayesian approach treats the unknown parameters as a random variable. Without a constraint on probability distribution, it is possible to construct statistical models without it being conditional on the sample size. Importantly, previous studies have indicated that mathematical models developed using a subjective Bayesian model can be as accurate in prediction as those developed using more laborious, time-
consuming, and costly approaches that use extensive empirical records and primary data collection.\textsuperscript{150}

**Study One Methods**

*Purpose*

The purpose of this study is to develop a model for predicting successful implementation of physical activity interventions in schools.

*Aim 1: To develop a Bayesian model for predicting successful implementation of physical activity interventions in schools.*

Objective 1a: Convene a panel of experts to identify factors influencing the successful implementation of youth physical activity interventions. Panel members will have expertise in physical activity interventions in children and adolescents, and implementation research.

Objective 1b: The information obtained in Objective 1a will inform the development of an initial Bayesian model to predict the probability that a school will successfully implement a physical activity intervention.

Objective 1c: To examine the internal validity of the model developed in Objective 1b.

*Study Design*

This study will use a cross-sectional design. Data will be collected using in-depth interviews and discussions with a panel of experts.

*Methods*

*Participants*
The expert panel will consist of five to six experts. To ensure consistency of expertise levels within the panel, we will target senior researchers who have substantial experience in implementing youth physical activity interventions. Eligibility of the panelists are as follows: 1) academic appointment at the rank of associate professor or higher, 2) a track record of leading implementation of youth physical activity interventions, and 3) a demonstrated record of publications on process evaluation and implementation of youth physical activity interventions. Based on the eligibility criteria, members of the study team will generate a list of panelists through reviewing journal articles and faculties’ biographical descriptions on university websites, and consulting with senior researchers. Since the objective of this study was to identify a set of core factors that influence implementation of physical activity interventions across youth-serving organizations, we will attempt to obtain a balance of individuals with expertise in various settings, such as schools and communities.

A Bayesian model of the current study

In this study, a Bayesian model will be constructed to predict the successful implementation of physical activity interventions in schools based on the factors identified and defined by an expert panel. There are two assumptions of the model. First, there are only two competing and mutually exclusive outcomes to be predicted (i.e., successful implementation or unsuccessful implementation). Second, the probability of the outcomes must sum to 1. The model will comprise three components: Prior odds of success, product of likelihood ratios for each level of each factor, and posterior odds of success.
Prior odds of success

The prior odds are the initial estimates of how much more likely it is that a change will be a success than a failure, before one has any information about the characteristics of the school. It is the ratio of prior probability of successful to probability of failure: \( \frac{P(S)}{P(F)} \). This prior information can be obtained from empirical evidence or expert opinions. If such information is not available which is not uncommon in a still-developing field, a non-informative (1/1) prior can be used. Gustafson and colleagues\(^3\) have found that models that used a non-informative prior can be as accurate in predictive ability as models that incorporated a prior knowledge. In this study, a non-informative (1/1) prior will be used.

Data likelihood

The data likelihood is the ratio of the conditional probability of observing the level of a particular factor (a datum) given a successful implementation, to the conditional probability of that same datum given a failed implementation: \( \frac{P(D|S)}{P(D|F)} \). Due to insufficient empirical data, this study will estimate the likelihood ratios based on subjective opinions elicited from an expert panel. This is explained further in the procedures section.

Posterior odds of success

Posterior are the products of priors and data likelihood. In this study, it refers to the final probability score of successful implementation in a school. The mathematical form of the Bayesian model is presented below. The example assumes two factors and each factor has three levels (a=high, b=medium,
The example illustrated that implementation success is predicted by two factors, in which factor 1 is at medium level and factor 2 is at low level.

\[
\frac{P(\text{success} | F1b, F2a)}{P(\text{Failure} | F1b, F2a)} = \frac{L(\text{success} | F1b)}{L(\text{Failure} | F1b)} \times \frac{L(\text{success} | F2a)}{L(\text{Failure} | F2a)} \times \frac{P(\text{success})}{P(\text{Failure})}
\]

**Procedures**

A four-round, modified Delphi process will be used to elicit the necessary information. In the first round, the panelists will complete an online survey to independently define successful implementation and suggest factors that influence successful implementation. The second round will provide a group setting for the panelists to elaborate and discuss their views on the definition and the suggested influencing factors of successful implementation through a video conference. In the third round, the panelists will complete another online survey to independently rate the importance of the suggested factors. In the fourth round, a final online survey will be used to collect data for assessing test-retest reliability of panelists’ ratings. The design of the surveys and video conference were guided by previous studies.\(^{150,151}\)

**First round**

The online survey will consist of seven open-ended questions that require the panelists to 1) operationalize successful implementation of a physical activity intervention carried out in youth-serving organizations based on their own experience, 2) suggest six factors that are most important in predicting successful implementation, and 3) describe the suggested factors at the three
factor-levels: high, moderate, and low. Responses will be aggregated and summarized into a straw model containing all the factors suggested by the panelists. The straw model will be circulated among panelists for review before the video conference.

Second round

All panelists will participate in a 90-minute video conference one month after the first survey. Section one of the video conference will provide opportunities for the panelists to elaborate and discuss their thoughts about 1) the definition of successful implementation, 2) importance of including the suggested factors, and 3) ways to improve descriptions for the suggested factors. Consensus on these three items will be achieved through an iterative process of voting and discussion. In section two of the video conference, the panelists will evaluate conditional independence of the potential factors. The panelists will be told to assume that an organization had a successful implementation and that the organization was rated as having a high-level on a specific factor, such as leadership support. They will be asked to discuss whether knowing this piece of information tells them a lot about how the organization might have responded to any of the other factors. If a factor violated the conditional independence, it will either be rewritten or eliminated. This process will be repeated for every potential factor.

After the video conference, straw model will be refined in light of the discussion. The revised straw model will be distributed among the panelists for
Third round

This survey will consist of two sections. In section one of this survey, the panelists will estimate likelihood ratios of the final list of factors. The likelihood ratios are the weights of each identified factors in contributing to a successful implementation. Panelists will be asked to assume that there are 100 hypothetical youth-serving organizations that had a successful implementation, and another 100 organizations had an unsuccessful implementation. They will be told to distribute the 100 successful cases and the 100 unsuccessful cases among the three factor-level for each of the identified factors.

In section two, the panelists will estimates for testing internal validity of the predictive model. The predictive model would ideally be applied to predict a successful implementation in real cases, which is external validity. In the absence of a suitable empirical data base, however, we will use experts’ opinions to generate a hypothetical data set for testing internal validity of the model. The panelists will be asked to assume that a physical activity intervention was carried out in a sample of 60 youth-serving organizations. Then, they will be provided with a set of computer-generated, hypothetical profiles reflecting how the 60 organizations rated on the factors identified in the second round. Every profile will include all of the identified factors but with varying factor-level for each factor. Each panelist will be asked to estimate how likely organizations with a specific
profile would have a successful implementation, when taking all the identified factors into account at the same time. These estimates are called “holistic ratings.” The holistic ratings will be estimated by using a 0-100 scale, where zero indicates absolutely no chance and 100 indicates 100% chance of successful implementation.

**Fourth round**

Since the holistic ratings will be used as a “criterion” for testing internal validity of the Bayesian predictive model, it is important to establish reliability of these ratings. Two weeks after the third round, panelists will complete a final online survey to re-rate 40 hypothetical profiles randomly selected from the original 60 profiles.

**Analysis**

**Objective 1b: test-retest reliability of holistic rating**

Pearson Product Moment Correlation will be performed to examine the test-retest reliability of the subjective holistic ratings. It will correlate the subjective holistic ratings obtained in step 6 with those obtained in step 5. A higher correlation value suggests a higher agreement between the scores measured in two time points.

**Objective 1b: refinement of the Bayesian model**

First, a diagnostic power score will be calculated for each factor to serve as a criterion for factor selection. The diagnostic power score refers to the range between the largest and the smallest likelihood ratio for that factor. If the highest and lowest likelihood ratios for a factor are 2.5/1 and 1/10, its diagnostic power
would be $2.5 + 10 = 12.5$. This score provides a crude measure regarding the amount of information that a certain factor can provide compared to other factors, with a larger value indicating a factor as more informative.

A backward factor selection procedure will be used in attempt to reduce the final list of factors to those that are most important in predicting successful implementation. We will start with a full model consisting of all factors identified in the third round and dropped one factor that had the lowest diagnostic power score at a time. A factor will be removed from the model if dropping it led to an increased or unchanged internal validity. The procedures will be repeated for every identified factor until internal validity had no further improvements.

Objective 1c: interval validity of the Bayesian model

Pearson Product Moment Correlation will also be used to assess internal validity of the Bayesian model resulting in step 7. It will correlate the panelists’ subjective holistic ratings for each profile to the corresponding score generated for that profile resulting from the application of the Bayesian model. A higher correlation value indicates better capability of the model in capturing the panelists’ judgment. The strength of the relationship between the two scores will be classified as low ($r < 0.5$), moderate ($r = 0.50$ to $0.69$), strong ($r = 0.70$ to $0.89$) and very strong ($r = 0.9$ to 1.0).
Study Two Methods

*Purpose*

The purpose of this study is to examine the influence of hypothesized factors on levels of implementation of a preschool-based physical activity intervention.

*Aim 2: To examine the effects of specific characteristics of the organization, implementer, and program on levels of implementation of a physical activity intervention delivered in a preschool setting.*

Objective 2: To assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in influencing levels of implementation of a physical activity intervention delivered in a preschool setting.

Hypotheses 2a: The specific characteristics of the organization, implementer, and program will have significant direct effects on levels of implementation.

Hypotheses 2b: The specific characteristics of the organization and program will have significant indirect effects on levels of implementation mediated through the characteristics of the implementers.

*Study Design*

This study will use a prospective observational study design. A subset of data from the Study of Health and Activity in Preschool Environments intervention (SHAPES) will be used for secondary data analysis. The SHAPES intervention was a group randomized trial (nested cohort design) conducted in 16 preschools. The preschools were recruited in Columbia, South Carolina.
(Lexington and Richland school districts). All preschools identified in the two districts were stratified into publicly-supported and privately-supported sub-groups. After stratification, eight preschools in each stratum (16 schools total) were randomly selected and invited to participate in the study. After random selection, preschools were pair-matched and each preschool in the pair was assigned to either the intervention (n=8) or a waiting list control condition (n=8).

The one-year intervention was implemented in 16 four-year-old preschool classrooms (nested in eight intervention preschools) for three consecutive academic years (from September 2008 through May 2011). Preschool teachers were the primary implementers. Outcome evaluation data were collected at pre-, mid- and post-intervention in each academic year. During the first intervention year, a sample of students (first cohort) was assessed at pre-intervention and post-intervention. During the second and third intervention years, new students in each participating teacher’s classroom were assessed at pre-intervention (the second and third cohort). To monitor the implementation progress, process evaluation data were collected at baseline and throughout the intervention period. The current study will only use process evaluation data collected from the 16 classrooms that have implemented the SHAPES intervention. The preschool classroom will serve as the unit of analysis.

**Methods**

**Participants**

The participants will be 16 four-year-old preschool classrooms that have implemented the SHAPES intervention. The preschool classrooms were from
eight intervention preschools. The intervention preschools varied on organizational characteristics. The number of enrolled students varied from 199 to 870. Fifty percent of the preschools were public schools. Thirty-eight percent of the preschools predominantly served Caucasian children, 38% served predominately African American children, and 24% have an equal distribution in race/ethnicity. Sixty-three percent of them offered full time programs; and 75% of them provided physical education. Among the 16 participating classrooms, the class size ranged from 12 to 21 students and all classrooms had a female teacher as the implementer.

*SHAPES intervention description*

The development of the intervention and process evaluation was guided by the social ecological model of health behavior. The social ecological model posits that behavior is influenced by factors operating at multiple levels including intrapersonal, interpersonal, institutional, community, and public policy. In applying this model to preschool settings, the SHAPES intervention theorized that changes in preschool social environments and institutional policies and practices would create a physical activity promoting environment for preschoolers. These changes include modifications in classroom resources, curriculum, teacher instructional practices, and allotment of time for specific activities such as preschool physical education and recess. Ultimately, a more supportive instructional and social environment would lead to improvement in children’s physical activity levels during their hours of preschool attendance.
The SHAPES intervention was not conceptualized as a curriculum, but a framework for increasing physical activity during the preschool day. This framework included an intervention protocol that provided the components for increasing physical activity; and an implementation protocol that assisted teachers to deliver those components.

**Intervention protocol**

The SHAPES intervention aimed to increase children’s physical activity by modifying the instructional and social environments within the preschools. It was designed to be flexible and adaptive to the preschool settings. Interventionists provided a framework with intervention components, examples and goals for overall physical activity, whereas preschool teachers adapted those intervention strategies to fit their own classrooms. There were four major intervention components: 1) Move Inside, 2) Move Outside, 3) Move to Learn and 4) enhanced social environment. To better fit the intervention into the preschool setting, the study investigators modified and enhanced the four intervention components over the 3-Year period based on on-going feedback provided by the interventionists and teachers.153

**Move Inside (MI)**

The Move Inside component was designed to provide daily opportunities for children to engage in structured, fundamental movement skill-based physical activities that were similar to formal physical education classes. During Year 1, the goal for teachers was to provide 60 minutes of skill-based physical activity per week. At the end of the first year, teachers expressed difficulties in leading
skill-based activities due to limited competences. Therefore, the goal for Year 2 and 3 was changed to providing at least 10 minutes per day of indoor activities (e.g., obstacle courses, dancing, and calisthenics) that was not part of recess and academic lessons. Teachers could break the 10 minutes into two 5-minute bouts of activities.

**Move Outside (MO)**

The Move Outside component was designed to increase children's outdoor time. Teachers were encouraged to provide outdoor recesses when possible. The goal for Year 1 was 60 minutes of recess daily. The feedback suggested that 60 minutes was an unrealistic time goal. Therefore, the goal for Year 2 and 3 was modified to providing two 20-minute recesses per day, in which each session should include at least one 5-minute teacher-led physical activity.

**Move to Learn (MTL)**

Move to Learn component was designed to integrate physical activity into academic lessons. Teachers were encouraged to incorporate physical activity into their typical daily lessons. The goal for Year 1 was 20-minutes of activity-based lessons. The feedback indicated that this component was highly valued by teachers but the time goal appeared to be unrealistic. Thus, the goal for Year 2 and 3 was modified to providing two 5-minute physical activity lessons per day for a total of 10 minutes daily.

**Enhanced social environment**

The interventionists first addressed the social environment by encouraging teachers to participate in physical activity with the children during physical
education, recess, and other physical activity times. Process data and initial teacher feedback showed that many teachers were less likely to participate actively with the children. Thus, the second strategy was recommending teachers to verbally encourage children’s participation during physical activity times. The verbal encouragement included acknowledging physical activity behaviors and promoting additional physical activities and not discouraging safe, appropriate physical activities.

Implementation protocol

The SHAPES intervention used a facilitative approach. The intervention staff developed a collaborative partnership with the preschool personnel and the change agents (preschool teachers in the 4-Year-old classroom). Interventionists provided training and on-going assistance (i.e., site visits, self-assessment, and newsletters) to increase teachers’ implementation capacity. In brief, initial training and group workshops included activities such as discussions and demonstrations to facilitate teachers to adapt the intervention to their own classrooms. Site visits mainly focused on problem solving. Teachers were encouraged to self-monitor their implementation progress and children’s participation. The newsletters mainly served as a platform for teachers to learn and share examples on how to integrate SHAPES in their preschools. This assistance was expected to provide preschool teachers with the knowledge, skills and confidence to implement the SHAPES intervention and, thus result in subsequent intervention delivery. The intervention delivery was expected to create an environment that is supportive for
physical activity and, ultimately resulted in increased physical activity in preschool children.

**Process evaluation**

An extensive process evaluation was conducted to document and monitor implementation of the SHAPES intervention. The logic model was used to guide the development of the process evaluation questions. Process data were collected throughout the 3-Year intervention period from multiple sources.

1. **Direct observation.** Two observational instruments were used. The Process observation checklist was developed for recoding the number of minutes of physical activity opportunities provided via each intervention component and the context in which the opportunity observed. The Observation System for Recording Physical Activity in Children-Preschool (OSRAC-P)\(^{154}\) was employed to assess preschool children’s physical activity and associated contextual condition. As components could be provided flexibly throughout the school day, observations were conducted over the entire school day. The administration procedures for the observations were slightly different across intervention years. In Year 1, observations were conducted on four fall days and four spring days and observers sampled children from multiple classrooms in each observation day. For Year 2 and 3, due to resources constraints, observations were conducted on one fall day and one spring day and observers sampled children from a single classroom in each day.

2. **Teacher survey.** The survey assessed teachers’ responses to the intervention. It was completed by the preschool teacher twice a year.
3. SHAPES staff rating form. The form was developed to evaluate the implementation progress. Interventionists completed rating forms once a year.

4. Field notes. The field note was designed to document observations and interactions with teachers during school visits. It was completed the interventionists after each school visit.

Measures related to the current study are described in more detail in the next section.

**Measures**

**Outcome**

Level of implementation is the outcome variable of this study. Level of implementation is a composite score that indicates the overall quality of implementation at the classroom level (program delivered from teachers to children). According to Liannan and Stecklers, this score is calculated by summing up the observed scores of four process evaluation elements. In this study, the maximum scores for each process evaluation elements equal to 100, which results in a maximum score of 400 for the intended level of implementation. The level of implementation is expressed as a percentage (observed levels of implementation dived by intended levels of implementation:

\[
Percent \text{ levels of implementation} = \frac{Observed \text{ levels of implementation}}{Intended \text{ levels of implementation}} \times 100
\]

**Dose delivered**

Dose delivered is defined as opportunities to engage in physical activity through the intervention components (i.e., MI, MO and MTL). During all years of the intervention, an evaluator used the process observation checklist to record
minutes of physical activity opportunities provided through the intervention components. For each intervention year, a percent of goal met for each component will be calculated by averaging the minutes of physical activity opportunities provided through MI, MO and MTL. A dose delivered score for each intervention year will be calculated by averaging the percent of goal met across the three components. The maximum dose delivered score equals to 100. A higher score indicates a higher percent of the total intended intervention components were delivered.

**Fidelity**

As the intervention evolved, the definition of fidelity was slightly different for Year 1 versus Year 2 and 3. For Year 1, fidelity is defined as children being physically active throughout a school day. For Year 2 and 3, fidelity is defined as children being physically active as reflected by MVPA, during physical activity opportunities. For all Years, a trained process evaluator used the OSRAC-P to observe children’s physical activity. The OSRAC-P uses a focal child, momentary time sampling observational system to record children’s physical activity and associated contextual condition.\(^{154}\) The observational system measures physical activity intensity, type (e.g., running, sitting, walking, and riding), and context (e.g., social environment such as group composition and child location). Physical activity intensity was rated on a 1 to 5 scale: (1) stationary, (2) stationary with limb movement, (3) slow, easy activity, (4) moderate intensity, and (5) vigorous activity. Intervals coded as level 1 and 2 were considered sedentary, level 3 were considered light intensity, level 4 and 5 were considered moderate and vigorous
intensity, respectively. These codes were modified from the Children's Activity Rating Scale (CARS)\textsuperscript{155}. For this study, moderate-to-vigorous physical activity (MVPA) included all intervals coded as 4 or 5.

During each 30-minute observation session, a sub-set of six children was sampled. Observers watched a child for a five-second observation interval followed by a 25-seconds record interval. Five-second observation intervals were repeated every 25 seconds across a five minute period of time to create an individual session for each focal child. Files from each child were merged and summarized by calculating the frequency of each activity code across all observation sessions. On average, each classroom was observed for four to seven 30-minute observation sessions (each session with a different subset of six children). A total of 217 hours of direct observation were conducted. A good inter-observer agreement for every observation category (Kappa=0.97) and for physical activity level (Kappa=0.93) were demonstrated in SHAPES. Data were collected using INTMAN software with handheld Dell Axim X5 computers (Dell World Trade LP, Round Rock, TX). For Year 1, the fidelity score is the mean daily percent of intervals in physical activity. For Year 2 and Year 3, the fidelity score is the mean percent of intervals in MVPA during the opportunities through the intervention components. The maximum fidelity score is 100. A higher score indicates a higher percent of the total intended intervention components were delivered.
Dose received

Dose received is defined as children enjoyment in intervention activities. During all years of the intervention, the evaluator assessed children’s enjoyment with intervention components on a four-point scale (1= none of the time; 4=all of the time). The score on this item will be used as a dose received score. The maximum fidelity score equals to 100. A higher score indicates a higher percent of the total intended intervention components were delivered.

Reach

Since SHAPES was implemented school-wide, reach could be considered to be all the children in the four-year-old participating preschool classrooms (100%). In this study, a score of 100 will be assigned to all classrooms.

Exposures

Organizational characteristics

Several constructs of the organizational characteristics were assessed at baseline, including preschools’ physical activity policies and practices, organizational complexity, organizational climates for physical activity, physical activity resources, and organizational functioning and administrative support. The SHAPES staff conducted structured interviews with the preschool directors to collect information on organizational characteristics. Preschool director responded to six items related to the preschools’ physical activity policies and practices on four-point Likert scales. Based on the Nutrition And Physical Activity Self Assessment for Child Care (NAPSACC) guideline, these items will be coded as 0=not meeting the NAPSACC guideline and 1= meeting the NAPSACC
guideline. Two items were about the organizational structure, including teacher’s education levels and teachers’ training on physical activity- or exercise-related aspects in the past year. Additionally, the SHAPES staff reviewed records to obtain information related to the structural characteristics (i.e., size, full- or half-day, public or private) of the preschool programs.

Preschools’ organizational climates and physical activity resources were measured by using the sub-scales of the Early Childhood Environment Rating Scale, Revised Edition (ECERS-R). The ECERS-R is a standardized rating scale that has been widely used to evaluate the resources and quality of early childhood education programs. The full ECERS-R scale consisted of seven sub-scales with 43 seven-point Likert-type items. The scale has been demonstrated to be reliable at the individual item and total scale score levels in analyses with 45 preschools. The reported correlations between observers were .921 for a Pearson product moment correlation and .865 for a Spearman rank order correlation. Internal consistency measures ranged from intraclass correlations of .71 (Parent and Staff Subscale) to .88 (Activities Subscale) on the seven subscales and was .92 for the total scale. The proposed study used four sub-scales with 21 seven-point Likert-type items that assessed the play environment (e.g., activities related to fine motor skills), social environment (e.g., supervision of gross motor activities), adult work environment (e.g., opportunities for professional growth), and physical activity resources (e.g., room arrangement for play).
During the end of each intervention year (spring only), the SHAPES intervention staff completed a rating form to evaluate organizational functioning and administrative support of the preschools. Two external factors that are suggested to have indirect influence on organizational characteristics were also measured at baseline. The observer used a sub-scale of the ECERS (1 items) to rate parental involvement in school activities on a seven-point Likert scale. Additionally, the connection between school and community was measured via the administrator interview. Administrators responded to an item related to the type of community program/activity provided on preschool campus. The mean scores of each sub-scale will be calculated. A composite score for organizational characteristics will be calculated for each preschool by summing up the scores across all seven constructs. A higher score indicates a more supportive organizational environment for implementing physical activity interventions.

**Program characteristics**

Two constructs of program characteristics were measured, including ease of use and implementation processes. At the end of each semester (fall and spring), teachers completed a teacher survey with items related to the program characteristics. Three items assessed their perceived difficulties in implementing the SHAPES intervention components. The scores of the three items were averaged. Three yes/no items measured their perceived barriers to implementation for each intervention components. A barrier score was calculated by counting the number of reported barriers. In regards of the implementation processes, teachers responded to one item regarding their perceived support
from the SHAPES staff. Additionally, the interventionists also completed a rating form to evaluate their working relationship with the classroom teachers. A composite score for program characteristic will be calculated by averaging the scores across all items for each intervention year.

Implementer characteristics

Two constructs of implementer characteristics were assessed, including self-efficacy in implementing the SHAPES intervention and perceived importance of the SHAPES intervention. At the end of each intervention Year (spring only), teachers responded to two items on the teacher survey indicating their perceived self-efficacy and perceived importance of the intervention. A composite score for implementer characteristics will be calculated by averaging the scores of the two items for each intervention year.

Analysis

Objective 2

Bayesian multilevel path analysis will be used to assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in influencing levels of implementation. Path analysis is an extended form of multiple regressions, but it allows a simultaneous modeling of several related regression relationships. Path analysis allows researchers to decompose the effects into direct and indirect component. The current study selects Bayesian estimation rather than the maximum likelihood estimation because it is more appropriate for modeling data based on a small sample and with non-normal distribution. These are often the case in process evaluation data.
Bayesian estimation does not rely on asymptotic (large-sample) theory and provides the whole distribution not assuming that it is normal. It makes inferences about a generic parameter $\theta$ by combining prior distributions for parameters with the data likelihood to form posterior distributions for the parameter estimates (posterior distribution = data likelihood $\times$ prior distribution), which is expressed as $P(\theta|\text{data}) \propto L(\text{data}|\theta)^*p(\theta)$. In most cases, posterior distribution is done by simulation using the Marko chain Monte Carol (MCMC) method that is an iterative procedure of generating samples estimates the parameters. By observing the simulated outcomes, we can estimate the population mean, variance and 95% posterior probability distribution (PPI) of the distribution for this samples. The 95% PPI is the 95% probability that in the population parameter lies between the two values.

**Bayesian path models of the proposed study**

In the proposed study, a prior hypothesized path model will be tested. Figure 3 presents a path diagram of the hypothesized relationships between organizational characteristics (ORG), program characteristics (PROG), implementer characteristics (IMT) and levels of implementation (LEVELIMT). The inner box includes two level-1 variables (PROG and IMT) measured at the teacher level and the outer box includes one level-2 variable (ORG) measured at the school level.

The level-1 model is given as:

$$M_{ij} = d_{Mj} + a_j X_{2ij} + e_{mij}$$

$$y_{ij} = d_{yi} + r_j X_{2ij} + b_{wj} M_{ij} + g_1 X_{2ij} + g_2 M_j + e_{yij}$$
The level-2 model is expressed as:

$$d_{Mj} = d_M + zX_{1j} + \mu_{mij}$$

$$d_{Yj} = d_Y + cX_{1j} + b^bd_{mj} + \mu_{Ymj}$$

$$b^w_j = b + \mu_{bj}$$

$$\alpha_j = \alpha + \mu_{\alpha j}$$

$$\tau'_j = \tau + \mu_{\tau j}$$

In level-1 model, \(i\) represents the classroom and \(j\) represents the school. The term \(y_{ij}\) denotes the outcome variable (IMTLEVEL); \(M\) is the mediating variable (IMT); and \(X_2\) is a classroom-level independent variable (PROG). The term \(e_{mj}\) and \(e_{yij}\) are the residuals of \(M\) and \(Y\); \(g_1\) and \(g_2\) are the function for centering the value of \(X_2\) and \(M\) to the group mean; the parameter \(d_{Mj}\) and \(d_{Yj}\) are random intercepts, and \(\alpha_j, b^w_j, \tau'_j\) are random slopes. The specification of random intercept and random slopes allows variations for level-2, schools. In particular, for the \(j^{th}\) school, \(\alpha_j\) quantifies the relationship between the mediating variable and independent variable, and \(b^w_j\) measures the relationship between the dependent variable and mediating variable after adjusting for the effects of the independent variable. In level-2 model, \(z\) denotes a school-level independent variable (ORG). The terms \(\alpha\) and \(b\) are population (or average) slopes, which specify the average effect of the independent variable on the mediating variable, and the average effect of the mediating variable on the dependent variable after controlling the independent variable, respectively. The parameter \(d_{Mj}\) and \(d_{Yj}\) are population (or average) intercepts.
In multilevel modeling, the level-1 residuals are assumed to be independent and follow normal distributions,

\[ e_{mi} \sim N(0, \sigma^2_e) \]

\[ e_{yi} \sim N(0, \sigma^2_{\varepsilon}) \]

And the level-2 residuals \( \mu_j = (\mu_{mij}, \mu_{yij}, \mu_{bij}, \mu_{\alpha j}, \mu_{\tau j})^T \) follow a multivariate normal distribution

\[ \mu_j \sim N(0, \Sigma) \]

where 0 is a vector of 0, and \( \Sigma \) is a 5x5 covariance matrix.

To conduct Bayesian multilevel modeling, priors are assigned to all unknown parameters in the model, including regression parameters (i.e., \( z, b^b, b, \alpha, \tau \)), level-1 variance (i.e., \( \sigma^2_e, \sigma^2_{\varepsilon} \)) and level-2 variance parameters (i.e., \( \Sigma \)).

For the regression parameters, an independent non-informative uniform prior will be assigned as follow,

\[ P(z, b^b, b, \alpha, \tau) \propto 1 \]

The level-1 variance parameters \( \sigma^2_e, \sigma^2_{\varepsilon} \) will be assumed to independently follow an inverse-gamma distribution

\[ P(\sigma^2_e, \sigma^2_{\varepsilon}) \propto IG(e_2, f_2) IG(e_3, f_3) \]

where \( e_2=f_2=0.01, e_3=f_3=0.01 \).

For the level-2 covariance matrix \( \Sigma \), the inverse Wishart distribution will be chosen. The inverse Wishart distribution is a multivariate generalization of the inverse gamma distribution, which is indexed by a degree of freedom parameter \( v \) and a scale matrix parameter \( S \). To represent vague prior knowledge, a small degree of freedom and a diagonal matrix with small values at the diagonal will be
assigned. In this study, a degree of freedom of 2 and a scale matrix parameter
\[
\begin{pmatrix}
0.0001 & 0 \\
0 & 0.0001
\end{pmatrix}
\]
will be used.

For variances of \(d_{Mj}, d_{Yj}\) and \(\tau_j'\) a noninformative prior will be chosen
\[
\sigma_{d_{Mj}}, \sigma_{d_{Yj}}, \sigma_{\tau_j'} \sim 1
\]
where \(\sigma_{d_{Mj}}, \sigma_{d_{Yj}}, \sigma_{\tau_j'}\) denote standard deviations of \(d_{Mj}, d_{Yj}\) and \(\tau_j'\), respectively.

The average indirect effects will be calculated as follow
\[
\text{ORG} \rightarrow \text{IMT} \rightarrow \text{IMTLEVEL} = zb^w_j = zb + \sigma_{zb_j}
\]
\[
\text{PROG} \rightarrow \text{IMT} \rightarrow \text{IMTLEVEL} = \alpha_j b^w_j = \alpha b + \sigma_{\alpha_j b_j}
\]
where \(\sigma_{zb_j}\) and \(\sigma_{\alpha_j b_j}\) denote the covariance between \(zb\) and \(\alpha b\), respectively.

The total effect (c) will be calculated as:
\[
c = c' + \tau_j' + zb + \sigma_{zb_j} + \alpha_j b + \sigma_{\alpha_j b_j}
\]

All modeling will be conducted using M-plus software, version 6.11 (Los Angeles, California)\(^{156}\) with the option of two-level and Bayes estimator. Bayesian estimates of the posterior mean, posterior standard error, and 95% PPI of the average direct and indirect effects will be obtained with 1000 iterations burn-in and 10,000 MCMC posterior draws. Initial path models will be conducted to identify the most important factors to construct the latent variable organizational characteristics. It will be done by entering one factor of the organizational characteristics into the hypothesized path model. Constructs that showed significant effect will be used to calculate the composite score for organizational characteristics for the final model.
The convergence of the final model will be determined by multiple criteria. The Potential Scale Reduction (PSR) factor will be used as the first convergence criterion. In brief, Bayesian analysis uses MCMC algorithms to iteratively obtain an approximation to the posterior distributions of the parameters. Such iterations are referred to as a chain. The PSR approach to determining convergence is to form an overestimate (between-chain) and an underestimate (within-chain) of the variance of the target distribution. The PSR criterion essentially requires the between-chain variation to be small relative to the total of between- and within-chain variation. A PSR value equal to 1 indicates convergence. The second criterion is the stability of trace plots for the posterior samples of the parameters. A tight and horizontal shape is desired as it suggests reliable estimations of the parameters. The third criterion is the autocorrelation plot. The autocorrelation plot shows the degree of correlatedness of parameter values across iterations for different intervals in the chain. A small value (≤0.1) is desirable to obtain approximately independent draws from the posterior. It is important to note that model fit and model comparison indices are not available for multilevel models and are thus not presented here. This is one of the limitations of using the MCMC estimation.

Hypothesis 2a and 2b

The hypotheses will be tested based on the 95% PPI. If the 95% PPI for the average direct and indirect effect includes zero, then we will reject the alternative hypothesis and conclude that the null hypothesis is true.
Study three methods

Purpose

The purpose of this study is to examine the influences of hypothesized factors on levels of implementation of a physical activity intervention in residential children’s homes (RCHs)

Aim 3: To examine the effects of specific characteristics of the organization, implementer, and program on levels of implementation of a physical activity intervention delivered in a children’s group home setting.

Objective 3: To assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in influencing levels of implementation of a physical activity intervention delivered in a children’s group home setting.

Hypotheses 3a: The specific characteristics of the organization, implementer, and program will have significant direct effects on levels of implementation.

Hypotheses 3b: The specific characteristics of the organization and program will have significant indirect effects on levels of implementation mediated through the characteristics of implementers.

Study Design

This study will use a prospective observational study design. A subset of data from the Environmental Intervention in Children’s Homes (ENRICH) will be used for secondary data analysis. The ENRICH intervention was a group randomized crossover trial which aimed to promote healthy eating and physical
activity among children living in residential children’s homes (RCH). In ENRICH, 69 RCHs in North Carolina (NC) and South Carolina (SC) that were affiliated with The Duke Endowment were invited to participate the study. The inclusion criteria were 1) having a relatively stable population of children, 2) requiring low-to-moderate management and 3) no restrictions on physical activity. Eligible RCHs were pair-matched based on the following organizational characteristics: location (SC or NC), complex versus simple organizational structure based on number of locations and services provided, participation in National Breakfast and Lunch Program, state accreditation, and existing physical activity programs. After the matching, each pair was assigned to the Early (n=17) or Delayed (n=12) intervention group.

From 2004 to 2006, the Early group received the intervention and the Delayed group served as a waiting control. From 2006 to 2008, the Delay group received the intervention and the Early group served as the control. Wellness teams (WT) from by adult staff working at the RCHs were the primary implementers. Outcome evaluation data were collected at pre- (Early: 2004; Delayed: 2006) and post-intervention (Early: 2006; Delayed: 2008). To monitor the implementation progress, process evaluation data were collected at baseline and throughout the intervention period. The ENRICH intervention included a nutrition and a physical activity component, but the current study will only focus on the physical activity component. The study will only use process evaluation data related to implementation of the ENRICH physical activity component collected from the 29 RCHs. The RCH will serve as the unit of analysis.
Methods

Participants

The participants will be 29 RCHs that have implemented the physical activity component of the ENRICH intervention. The RCHs varied on organizational characteristics. The number of children served varied from 199 to 870. Among the 29 RCHs, 31% were located in SC and 69% were in NC; 67% had a complex structure; 30% were participating in National Breakfast and Lunch Program; and 67% were accredited.

ENRICH intervention description

The development of the intervention was guided by the Structural Ecological Model of health behavior. The model posits that individual health behaviors are typically influenced by individual-level attributes as well as the settings under which people live. Structural interventions target influencing factors in social and physical environments that are beyond individual control can create a supportive setting to reinforce health behaviors in individuals. The model identifies four structural factors of environmental influences: 1) availability of protective and harmful products, 2) characteristics of available opportunities, 3) social structures and policies, and 4) media and cultural messages. In applying this model to RCH settings, ENRICH theorized that changes in RCH organizational social and physical environment would have a positive influence on physical activity behavior in youth residing in RCHs. These changes included increasing quantity and enhancing quality of physical activity opportunities, modifying organizational physical activity policies and practice, and creating a
supportive social and media environment. Ultimately, a more supportive RCH environment would lead to improvement in physical activity levels of 8-11-year-old children living in RCHs.

*Intervention protocol*

The ENRICH intervention aimed to create and sustain RCH environments to support and promote physical activity among RCH residents. It was designed as a flexible and adaptive intervention to suit the RCH settings. Initial planning meetings were scheduled with participating RCH directors and staff members, community partners and stakeholders to incorporate their inputs to the intervention and implementation protocol. The interventionists developed principles to facilitate the RCHs to create an environment that supported physical activity. The wellness team (WT) formed by adult staff of RCHs adopted those principles and developed strategic plans based on local needs and resources. There were six ENRICH principles that guide the WT in the development of specific environmental features for increasing physical activity: (1) providing more physical activity opportunities which could be achieved through scheduling and provision of equipment; (2) ensuring that physical activity opportunities are appealing; (3) strengthening social support and adult modeling for physical activity; (4) developing, strengthening, and/or enforcing policies; (5) increasing positive media messages; and (6) developing organizational structures to support these change. These strategies also served as the essential elements for determining levels of implementation.

*Implementation protocol*
The investigation team scheduled meetings with participating RCH directors and staff members to develop working relationships and obtain support from the administrators. Then, the investigation team provided trainings to the RCH adult staff to facilitate them in accessing local resources, and development and implementation of strategic plans. In brief, a six-hour initial training was provided in the first summer (2004 for early group; 2006 for delayed group). It was designed to provide skills for assessing the RCH environment and policies/practices with regards to physical activity; to develop the first year plan to support and promote physical activity; and to carry out, monitor, and adjust the plan. In the second summer (2005 for early group and 2007 for delayed group), a follow-up training (4-6 hours) was provided to facilitate the WT in developing the second year strategic plan with objectives that were not addressed in the first plan. After the first training, interested RCH adult staff members in each RCH formed a WT. The WT developed written plans within one month of training and the ENRICH staff reviewed and approved the plan. After the approval, the WTs had one year to implement the plan. Technical assistance was provided throughout the intervention period via consultations, site visits and telephone contacts. If the plan is implemented as intended, the intervention is expected to improve the RCH environment that supported physical activity and, ultimately result in increased proportion of RCH residents aged 11-18 years that meet the physical activity recommendation (60 minutes of MVPA on 5 or more of previous 7 days).
**Process evaluation**

An extensive process evaluation was conducted to document and monitor implementation of the ENRICH intervention. Process data were collected throughout the 2-year intervention period from multiple sources.

1. **Media observation checklist.** The checklist was designed to assess presence of media promoting nutrition and physical activity and opportunities in pre-designated common areas and recreational areas in RCHs. It was administered annually by the evaluator.

2. **End-of-year survey (EOY).** It was a 40-item rating scale designed to assess WT planning and implementation. The evaluator used this survey to conduct an interview with WT contacts at the end of each intervention year.

3. **Post-visit survey.** It was a 9-item rating scale developed to document evaluator impressions on key elements of the RCH environment. The scale was completed by the evaluator following each site visit (once per year).

4. **End-of-intervention (EOI) assessment.** It was a 12-item rating scale developed to document interventionist impressions on the progress of the RCH and WT. The scale was completed by interventionist at the end of the 2-year intervention.

5. **Staff rating scale.** It was a 5-item rating scale that was used to document impression of all ENRICH staff on RCH progress. It was completed by all ENRICH staff once per year.
Items that will be used in the current study are described in more detail in the next section.

**Measures**

**Outcome**

Level of implementation will be the outcome variable of this study. Level of implementation is a composite score that indicates the overall quality of implementation at the RCH level (program delivered from wellness teams to residents). As proposed by Saunders, environmental interventions like ENRICH, levels of implementation only consisted of dose delivered and fidelity and these two aspects are often united as one aspect. In this study, the score on implementation level will be the observed scores of dose delivered and fidelity. We expressed level of implementation as a percentage (observed levels of implementation divided by intended levels of implementation):

\[
\text{Percent levels of implementation} = \frac{\text{Observed levels of implementation}}{\text{Intended levels of implementation}} \times 100
\]

**Dose delivered and fidelity**

This study defines dose delivered and fidelity as to what extent the wellness team implemented the plans to enhance the RCH environment. It was measured by multiple methods and all measures were implemented once in each intervention year. WT contact completed an EOY (1 item) to assess the overall implementation of the strategic plan. Intervention staff conducted observations of the RCH physical activity environment (opportunities and structure, opportunity characteristics, and policies and practices) using the post-visit survey (3 items).
Three items from the post-visit survey assessed RCH opportunities and structure for physical activity, characteristics of the physical activity opportunity and organizational policies and practices related to physical activity. Intervention staff conducted observations on the RCH media environment using the ENRICH media observation checklist (9 items). A dose delivered and fidelity score in each year will be the sum of the 14 items. The score for observed levels of implementation will be calculated by averaging the total dose delivered and fidelity score across the 2 years.

**Exposures**

**Organizational characteristics**

Five aspects of organizational characteristics will be assessed: organizational policies and practices, organizational structure, physical activity resources, organizational functioning and administrative support. At baseline, the assistant chief executive officer (CEO) or designated representative of the RCH reported organizational characteristics by completing two surveys: 1) the Physical Activity and Dietary Environmental Assessment questionnaire (PADEA) and 2) the Organizational Assessment Survey (OA). Both surveys were developed specifically for ENRICH. The PADEA is a 69-item scale developed based on structural ecological model. It was designed to assess the physical activity and food environments of RCHs. The current study will use 17 items from the PADEA that assess three aspects of organizational characteristics: physical activity policies and practices (10 items), organizational structure (3 items), and physical activity resources (4 items). The scale has been shown to have
acceptable to good 2-week test-retest reliability (ICC= 0.38 to 0.98) and internal consistency (Cronbach’s alpha=0.62 to 0.90). The OA is a 30-item scale which was developed to collect descriptive information of RCHs. The current study will use four items measuring physical activity resources (2 items) and organizational structure (2 items). Additionally, ENRICH staff also completed a rating form to assess organizational functioning (1 item) and administrative support (1 item) during the first intervention year. An index score for each aspect will be calculated by averaging scores of the included item. A composite score for organizational characteristics will be calculated by summing up all the index scores. A higher score indicates a more conducive organizational environment for implementing physical activity interventions.

**Program characteristics**

Two constructs of program characteristics were measured, including ease of use and implementation processes. At the end of each intervention year, WT contact persons completed the EOY survey. The study will use three items from the EYO to measure WT contacts’ perceptions of the program, including perceived difficulties in implementing the ENCRICH intervention components and perceived quality of the support provided by the ENRICH staff. In regards of the implementation processes, one item from the staff rating scale will be used to evaluate intervention staff working relationships with the WT. An index score for program characteristic will be calculated by averaging the scores of all items across two years.
Implementer characteristics

Four constructs of implementer characteristics were assessed: preparedness, perceived intervention effectiveness, competence, and WT functioning. Two items from the EYO will be used to measure WT preparedness of implementing the intervention and WT perceived effectiveness of the intervention. At the end of the 2-year intervention, intervention staff completed the EOI to document their impressions of RCH and WT progress. This study will use two items on the EOI survey to assess WT competence and WT functioning. An index score for implementer characteristics will be calculated by averaging the scores of the four items across the two intervention years.

Analysis

Objective 3

Bayesian path analysis will be used to assess the direct and indirect effects of specific characteristics of the organization, implementer, and program in influencing levels of implementation. The rationale for selecting Bayesian path analysis has been described in Study 2. In this study, a prior hypothesized path model will be tested. Figure 4 is a path diagram of the hypothesized relationships between organizational characteristics (ORG), program characteristics (PROG), implementer characteristics (IMT) and levels of implementation (LEVELIMT). The inner box includes three level-1 variables (ORG, PROG and IMT) measured at the RCH level and the outer box represents that all level-1 variables were clustered within states (North Carolina and South Carolina).
The level-1 model is given as:

\[
M_{ij} = d_{Mj} + z_j X_{1ij} + \alpha_j X_{2ij} + e_{mi j}
\]

\[
y_{ij} = d_{Yj} + c'_j X_{1ij} + \tau'_j X_{2ij} + b^{w_j} M_{ij} + g_1 X_{1j} + g_2 X_{2j} + g_j + \varepsilon_{Yij}
\]

The level-2 model is expressed as:

\[
d_{Mj} = d_M + \mu_{mi j}
\]

\[
d_{Yj} = d_Y + b^b d_{mij} + \mu_{ymj}
\]

\[
b^{w_j} = b + \mu_{b}.
\]

\[
z_j = z + \mu_{zj}
\]

\[
\alpha_j = \alpha + \mu_{\alpha j}
\]

\[
c'_j = c' + \mu_{c'_j}
\]

\[
\tau'_j = \tau + \mu_{\tau'_j}
\]

In level-1 model, \(i\) represents the RCH and \(j\) represents the state. The term \(y_{ij}\) denotes the outcome variable (IMTLEVEL); \(M\) is the mediating variable (IMT); and \(X_1\) and \(X_2\) are the independent variable (ORG and PROG). The term \(e_{mi j}\) and \(e_{Yij}\) are the residuals of \(M\) and \(Y\); \(g_1, g_2\) and \(g_3\) are the function for centering the value of \(X_1, X_2\) and \(M\) to the group mean; the parameter \(d_{Mj}\) and \(d_{Yj}\) are random intercepts, and \(\alpha_j, b^{w_j}, z_j, c'_j\) and \(\tau'_j\) are random slopes. The specification of random intercept and random slopes allows variations for level-2, states. In particular, for the \(j^{th}\) state, \(\alpha_j\) quantifies the relationship between the mediating variable and independent variable, and \(b^{w_j}\) measures the relationship between the dependent variable and mediating variable after adjusting for the
effects of the independent variable. In level-2 model, the terms $\alpha$ and $b$ are population (or average) slopes, which specify the average effect of the independent variable on the mediating variable, and the average effect of the mediating variable on the dependent variable after controlling the independent variable, respectively. The parameter $d_{Mj}$ and $d_{Yj}$ are population (or average) intercepts.

In multilevel modeling, the level-1 residuals are assumed to be independent and follow normal distributions,

$$e_{mi} \sim N(0, \sigma^2_2)$$

$$e_{Yi} \sim N(0, \sigma^2_3)$$

And the level-2 residuals $\mu_j = (\mu_{mij}, \mu_{Ymj}, \mu_{bij}, \mu_{zij}, \mu_{aj}, \mu_{c'j}, \mu_{tj})^T$ follow a multivariate normal distribution

$$\mu_j \sim N(0, \Sigma)$$

where $\mathbf{0}$ is a vector of 0, and $\Sigma$ is a 5x5 covariance matrix.

To conduct Bayesian multilevel modeling, priors are assigned to all unknown parameters in the model, including regression parameters (i.e., $z$, $b^b$, $b$, $\alpha$, $c$, $\tau$), level-1 variance (i.e., $\sigma^2_2$, $\sigma^2_3$) and level-2 variance parameters (i.e., $\Sigma$). For the regression parameters, an independent non-informative uniform prior will be assigned as follow,

$$P(z, b^b, b, \alpha, c, \tau) \propto 1$$
The level-1 variance parameters $\sigma_2^2$, $\sigma_3^2$ will be assumed to independently follow an inverse-gamma distribution

$$P(\sigma_2^2, \sigma_3^2) \sim \text{IG}(e_2,f_2) \text{ IG}(e_3,f_3)$$

where $e_2=f_2=0.01$, $e_3=f_3=0.01$.

For the level-2 covariance matrix $\Sigma$, the inverse Wishart distribution will be chosen. The inverse Wishart distribution is a multivariate generalization of the inverse gamma distribution, which is indexed by a degree of freedom parameter $v$ and a scale matrix parameter $S$. To represent vague prior knowledge, a small degree of freedom and a diagonal matrix with small values at the diagonal will be assigned. In this study, a degree of freedom of 2 and a scale matrix parameter $\begin{pmatrix} 0.0001 & 0 \\ 0 & 0.0001 \end{pmatrix}$ will be used.

For variances of $d_{Mj}$, $d_{Yj}$ and $c'_j$, $\tau'_j$ a noninformative prior will be chosen

$$\sigma_{d_{Mj}}, \sigma_{d_{Yj}}, \sigma_{c'_j}, \sigma_{\tau'_j} \sim 1$$

where $\sigma_{d_{Mj}}$, $\sigma_{d_{Yj}}$, $\sigma_{c'_j}$, $\sigma_{\tau'_j}$ denote standard deviations of $d_{Mj}, d_{Yj}, c'_j$, and $\tau'_j$, respectively. The average indirect effects will be calculated as follow

$$\text{ORG} \rightarrow \text{IMT} \rightarrow \text{IMTLEVEL} = zb_{wj} = zb + \sigma_{zbj}$$

$$\text{PROG} \rightarrow \text{IMT} \rightarrow \text{IMTLEVEL} = \alpha_j b_{wj} = \alpha b + \sigma_{a_{jbj}}$$

where $\sigma_{zbj}$ and $\sigma_{a_{jbj}}$ denote the covariance between $zb$ and $\alpha b$, respectively.

The total effect ($c$) will be calculated as:

$$c = c' + \tau'_j + zb + \sigma_{zbj} + \alpha b + \sigma_{a_{jbj}}$$

All modeling will be conducted using M-plus software (version 6.11) with the Bayes estimator option. Bayesian estimates of all parameters and variance
components in the framework will be obtained with 1000 iterations burn-in and 10,000 Markov chain Monte Carlo (MCMC) posterior draws. Posterior mean, posterior standard error and 95% posterior probability interval (PPI) of the average direct and indirect effects will be obtained.

The convergence of the final model will be determined by multiple criteria. The Proportional Scale Reduction (PSR) factor will be used as the first convergence criterion. In brief, Bayesian analysis uses MCMC algorithms to iteratively obtain an approximation to the posterior distributions of the parameters. Such iterations are referred to as a chain. The PSR approach to determining convergence is to form an overestimate (between-chain) and an underestimate (within-chain) of the variance of the target distribution. The PSR criterion essentially requires the between-chain variation to be small relative to the total of between- and within-chain variation. A PSR value equal to 1 indicates convergence. The second criterion is the stability of trace plots for the posterior samples of the parameters. A tight and horizontal shape is desired as it suggests reliable estimations of the parameters. The third criterion is the autocorrelation plot. The autocorrelation plot shows the degree of correlatedness of parameter values across iterations for different intervals in the chain. A small value (≤0.1) is desirable to obtain approximately independent draws from the posterior.

_Hypothesis 3a and 3b_

The hypotheses will be tested based on the 95% PPI. If the 95% PPI for the average direct and indirect effect includes zero, then we will reject the alternative hypothesis and conclude that the null hypothesis is true.
References


38. Physical Activity Guidelines for Americans Midcourse Report


Brown W, Potter E. Second year report of the evaluation of the four-year-old child development programs funded through the South Carolina Education Improvement Act (EIA). Columbia: SC: College of Education, University of South Carolina; 2003.

Bush PL, Laberge S, Laforest S. Physical activity promotion among underserved adolescents: 'Make it fun, easy, and popular.' *Health Promot Pract* 2010;11(3):79S-87S.


Digest of Education Statistics Web site.


Hortz B, Petosa R. Impact of the “Planning to be Active” leisure time physical exercise program on rural high school students. *J Adolesc Health* 2006;39(4):530-535.


Lubans DR, Morgan PJ, Callister R, Collins CE. Effects of integrating pedometers, parental materials, and E-mail support within an extracurricular school sport intervention. *J Adolesc Health* 2009;44(2):176-183.


Muthén B. Bayesian analysis in Mplus: A brief introduction. 2010.


Panzano PC, Seffrin B, Chaney-Jones S. The innovation diffusion and adoption research project (IDARP): Moving from the diffusion of research results to promoting the adoption of evidence-based innovations in the Ohio mental health system. *New Res Ment Hlth*. In press


Steele JS, Buchi KF. Medical and Mental Health of Children Entering the Utah Foster Care System. *Pediatrics* 2008;122(3):e703-e709.


# Appendix A

## Measures of Implementation Aspects

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Key process evaluation question</th>
<th>Other labels used in the studies</th>
</tr>
</thead>
</table>
| **Dose delivered** | 1. What is the actual number of intervention components that were delivered by the implementer? 33,97,98,105,107,108,110,111,113-118,122,125,127,128,136,143,144 | • Completeness<sup>125</sup>  
• Dose/dosage<sup>33,112,113,118,122,127</sup>  
• Exposure<sup>101</sup>  
• Quantity<sup>107</sup> |
|                  | 2. What is the proportion of the total intervention components that were completed by the implementers? 118,125 |                                  |
|                  | 3. What is the duration of the intervention activities? 33,101,110,113,114,118,125,127,128,132,136 |                                  |
|                  | 4. What is the frequency with which the intervention components/activities were implemented? 112,128,132,133 |                                  |
|                  | 5. Did the implementers comply with the prescribed dosages? 98 |                                  |
| **Fidelity**     | 1. To what extent were the program objectives were implemented as intended in general? 105,107-109 | • Accuracy<sup>118,125</sup>  
• Consistency<sup>33</sup> |
|                  | 2. To what extent was the program was implemented in accordance to the intervention philosophy as defines by a set of core element<sup>33,102,111,113,118,125</sup> |                                  |
| **Dose received**| 1. Were children satisfied with the program? 104,120,121                                                                                     | • Awareness and perception of the program<sup>108</sup>  
• Dose<sup>127</sup>  
• Quality<sup>118</sup>  
• Responsiveness<sup>125</sup>  
Satisfaction<sup>104,120,121,124</sup> |
|                  | 3. Were children actively engaged in the program activities (e.g., read the material, completed and returned prescribed tasks)? 97,102,114,120,125,139 |                                  |
| **Reach**        | 1. What percentage of the total participants who attended or showed up for the intervention activities? 33,102,104,105,107,110-112,114,117-121,124,125,127,130,139 | • Attendance<sup>101,104,118,121,124</sup>  
• Exposure<sup>33,105</sup>  
• Participation<sup>99,120</sup> |
### APPENDIX B

**Observed Degrees of Implementation Aspects of Individually-Oriented Interventions**

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose delivered</th>
<th>Fidelity</th>
<th>Dose received</th>
<th>Reach</th>
<th>Change in PA levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of total intervention component delivered across all participating organizations</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Engagement: % of total participants who actively engaged in intervention activities</td>
<td>Enjoyment &amp; Satisfaction: % of total participants who reported enjoyed or satisfied with the program</td>
<td>% of total participants that have attended or participated in the program activities</td>
</tr>
<tr>
<td>Bush et al, 2009</td>
<td></td>
<td></td>
<td></td>
<td>NA: 71%</td>
<td>NA</td>
</tr>
<tr>
<td>Cardon et al, 2009</td>
<td></td>
<td></td>
<td></td>
<td>Curriculum: Teacher-reported: 55%</td>
<td>Curriculum: Child-reported: 80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Parent-reported: 56%</td>
<td>Family: Child-reported: 67%</td>
</tr>
<tr>
<td>Chomitz et al, 2003</td>
<td></td>
<td></td>
<td></td>
<td>Family: Parent-reported: 89%</td>
<td>Overall: Parent-reported: 63%</td>
</tr>
<tr>
<td>Donnelly et al, 2009; Gibson et al, 2008 (PAAC)</td>
<td>Overall: Implementer-reported: 72%</td>
<td>Overall: Implementer-reported: 33%</td>
<td>Overall: Child-reported: 93%</td>
<td></td>
<td>Parent component: Program record: 16%</td>
</tr>
<tr>
<td>Study</td>
<td>Dose delivered</td>
<td>Fidelity</td>
<td>Dose received</td>
<td>Reach</td>
<td>Change in PA levels</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Dzewaltowski et al, 2009 (Healthy Youth Places)</td>
<td>Curriculum: Implementer-reported: 82% (7th grade); 71% (8th grade)</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Engagement: % of total participant who actively engaged in intervention activities</td>
<td>% of total participants who have attended or participated in the program</td>
<td>+</td>
</tr>
<tr>
<td>Erwin et al, 2011</td>
<td>Overall: Implementer-reported: component delivered in 55% teachers#</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Enjoyment &amp; Satisfaction: % of total participant who reported enjoyed and/or satisfied with the program</td>
<td>% of total participants that have attended or participated in the program activities</td>
<td>+</td>
</tr>
<tr>
<td>Gentile et al, 2009 (Switch)</td>
<td>Overall: Child- &amp; Parent-reported: NA</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Overall: Site director-reported: NA</td>
<td>% of total participants that have attended or participated in the program activities</td>
<td>O</td>
</tr>
<tr>
<td>Gortmaker et al, 1999 (Planet Health)</td>
<td>Curriculum: Implementer-reported: 22%</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Overall: Site director-reported: NA</td>
<td>% of total participants that have attended or participated in the program activities</td>
<td>O</td>
</tr>
<tr>
<td>Gortmaker et al, 2012</td>
<td>Overall: Implementer-reported: component delivered in 64% sites#</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Overall: Site director-reported: NA</td>
<td>% of total participants that have attended or participated in the program activities</td>
<td>+</td>
</tr>
<tr>
<td>Study</td>
<td>Dose delivered</td>
<td>Fidelity</td>
<td>Dose received</td>
<td>Reach</td>
<td>Change in PA levels</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------</td>
<td>----------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Herbert 2013 (TigerKids)</td>
<td>Overall: Implementer-reported: 39%</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Hoelscher et al, 2010 (CATCH)</td>
<td>Overall: Implementer-reported: NA</td>
<td></td>
<td></td>
<td></td>
<td>Curriculum: Child-reported: O</td>
</tr>
<tr>
<td>Jones et al, 2010 (HIKCUPS)</td>
<td>Overall: Implementer-reported: 97%; observer: 96%</td>
<td>Overall: Implementer-reported: 98% across all centers</td>
<td>FMS: Parent-reported: 42% to 82%</td>
<td></td>
<td>Curriculum: Program-record: 72%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Refresher session: Program record: 51%</td>
</tr>
<tr>
<td>Lubans et al 2008 (Program X)</td>
<td></td>
<td></td>
<td>Self-monitoring: Child-reported: 11%</td>
<td></td>
<td>Overall: Program-record: +</td>
</tr>
</tbody>
</table>
## APPENDIX B (CONTINUED)

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose delivered</th>
<th>Fidelity</th>
<th>Dose received</th>
<th>Reach</th>
<th>Change in PA levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of total intervention component delivered across all participating organizations</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Engagement: % of total participant who actively engaged in intervention activities</td>
<td>Enjoyment &amp; Satisfaction: % of total participant who reported enjoyed and/or satisfied with the program</td>
<td>% of total participants that have attended or participated in the program activities</td>
</tr>
<tr>
<td>Madsen et al, 2013 (SCORES)</td>
<td>Overall: Implementer-reported: 73%</td>
<td>Overall: School staff: 48%</td>
<td>Overall:</td>
<td>Overall:</td>
<td>+ OW children</td>
</tr>
<tr>
<td>Marcoux 1999 (SPARK)</td>
<td>Overall: observer: 68%</td>
<td>Overall: Child-reported: NA</td>
<td>Family: Parent-reported: 68%</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>McKenzie et al, 2004 (M-SPAN PE)</td>
<td>Overall: Child-reported: NA</td>
<td>Overall: Child-reported: NA</td>
<td>Overall: Child-reported: NA</td>
<td>+ boys</td>
<td></td>
</tr>
<tr>
<td>Morgan 2011 (HDHK)</td>
<td>Self-monitoring: Father-reported: 74%</td>
<td>Overall: Father-reported: 96%</td>
<td>Overall: Program record: 81%</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Naylor et al, 2006 (AS! BC)</td>
<td>Overall: Implementer-reported: 85% (INT1), 84% (INT2)</td>
<td>PA component: Implementer-reported: 67%</td>
<td>+ boys</td>
<td>+ boys</td>
<td></td>
</tr>
</tbody>
</table>

261
## APPENDIX B (CONTINUED)

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose delivered</th>
<th>Fidelity</th>
<th>Dose received</th>
<th>Reach</th>
<th>Change in PA levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neumark-Sztainer et al, 2009, (Ready Set ACTION)</td>
<td>% of total intervention component delivered across all participating organizations</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Engagement: % of total participant who actively engaged in intervention activities</td>
<td>Enjoyment &amp; Satisfaction: % of total participant who reported enjoyed and/or satisfied with the program</td>
<td>% of total participants that have attended or participated in the program activities</td>
</tr>
<tr>
<td>Neumark-Sztainer et al 2010, (New Moves)</td>
<td><strong>Family:</strong> Parent-reported: 90%</td>
<td><strong>Theatre session:</strong> Child-reported: 70%</td>
<td><strong>Theatre session:</strong> Program record: ≥75% in 59% participants <strong>Family:</strong> Child- and parent-reported: 44%-67%</td>
<td><strong>Overall:</strong> Program record: 91% children attended 81% lessons</td>
<td>O</td>
</tr>
<tr>
<td>Pate et al, 2003 (Active Winner)</td>
<td><strong>Overall:</strong> Implementer- and Child-reported: NA</td>
<td><strong>Overall:</strong> Implementer- and Child-reported: NA</td>
<td><strong>Curriculum:</strong> Program record: 5% children attended 50% of lessons</td>
<td><strong>Overall:</strong> Program record: &gt;80% in 64% participants*</td>
<td>O</td>
</tr>
<tr>
<td>Patrick et al, 2006, (PACE+)</td>
<td><strong>Overall:</strong> program record: &gt;80% in 64% participants*</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>
## APPENDIX B (CONTINUED)

<table>
<thead>
<tr>
<th>Study</th>
<th><strong>Dose delivered</strong></th>
<th><strong>Fidelity</strong></th>
<th><strong>Dose received</strong></th>
<th><strong>Reach</strong></th>
<th>Change in PA levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of total intervention component delivered across all participating organizations</td>
<td>% of total core elements adopted across all participating organizations</td>
<td>Engagement: % of total participant who actively engaged in intervention activities</td>
<td>Enjoyment &amp; Satisfaction: % of total participant who reported enjoyed and/or satisfied with the program</td>
<td>% of total participants that have attended or participated in the program activities</td>
</tr>
<tr>
<td>Salmon et al, 2005 (Switch-Play)</td>
<td>Overall: Implementer-reported: NA</td>
<td>Curriculum: Child-reported: 92%</td>
<td>Curriculum: Child-reported: 81% (INT1), 84% (INT2)</td>
<td>Curriculum: Child-reported: 88%</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Family: Child-reported: 57%</td>
<td>Parent-reported: 78%</td>
<td>Family: Parent-reported: 70%</td>
<td></td>
</tr>
<tr>
<td>Simon et al, 2004 (ICAPS)</td>
<td>Overall: Program record: NA</td>
<td></td>
<td></td>
<td>Overall: Program record: 50%</td>
<td>+</td>
</tr>
<tr>
<td>Webber et al, 2008; Young et al, 2008 (TAAG)</td>
<td><strong>Curriculum:</strong> Observer: 91%</td>
<td><strong>PE:</strong> Observer: 43%</td>
<td><strong>Activities Challenges:</strong> Observer: 61.5%</td>
<td><strong>Curriculum:</strong> Staffs: 83%</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Activity challenges:</strong> Observer: 38%</td>
<td></td>
</tr>
<tr>
<td>Weintraub 2008, (SPORT)</td>
<td>Overall: Program record: 42%</td>
<td></td>
<td></td>
<td>Overall: Program record: 3-month</td>
<td>+</td>
</tr>
<tr>
<td>Werch et al, 2003 (SPORT plus)</td>
<td><strong>Consultation:</strong> Observer: 97%</td>
<td><strong>Consultation:</strong> Observer: 98%</td>
<td><strong>Consultation:</strong> Observer: 89%</td>
<td><strong>Family:</strong> Program record: 82%</td>
<td>+ current drinker</td>
</tr>
</tbody>
</table>
## APPENDIX B (CONTINUED)

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose delivered</th>
<th>Fidelity</th>
<th>Dose received</th>
<th>Reach</th>
<th>Change in PA levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson et al, 2009 (ACT)</td>
<td>Overall: Observer: 86%</td>
<td>Overall: Observer: 83%</td>
<td>Overall: Program record: 56%</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Notes: # Information directly extracted from study. CH=children group home, NA= not available, PA=physical activity, PE=physical education. “+”: significant improvement in children’s physical activity level at p<0.05, “O”: no significance improvement in children’s physical activity levels at p<0.05, “-”: significant decline in children’s physical activity levels significant at p<0.05
## APPENDIX C

### OBSERVED DEGREES OF IMPLEMENTATION ASPECTS OF ENVIRONMENTAL INTERVENTIONS

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose delivered and fidelity</th>
<th>Dose received</th>
<th>Change in PA outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonvin et al, 2013 (Switzerland Youla Bouge)</td>
<td>Overall: observer ratings: 43%</td>
<td>Overall: parent-reported: 100%</td>
<td>O</td>
</tr>
<tr>
<td>Finch et al, 2014</td>
<td>Overall: Observer: NA</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Saunders et al, 2013, (ENRICH)</td>
<td>Overall: Implementer and observer: 100% of the PA component adopted by 53% CH#</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Saunders et al 2006; Ward et al, 2006 (LEAP)</td>
<td>Overall: Implementer and observer: 80%</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

Notes: # Information directly extracted from study. CH=Children group home, NA= not available