EARN-A-BIKE @ SCHOOL:

PROGRAM EVALUATION PROPOSAL

UTHealth Houston School of Public Health

Chisom Ovuegbe

Matthew Usevitch

*Course Directors:*

Dr. Lara Savas

Dr. Nalina Ranjit

Table of Contents

[ABSTRACT 1](#_Toc102052798)

[INTRODUCTION 3](#_Toc102052799)

[SECTION I - Needs assessment, Program Description, and Hypotheses 4](#_Toc102052800)

[Introduction 4](#_Toc102052801)

[Part A - Epidemiology of Program’s Health Problem 4](#_Toc102052802)

[Part B – Program Description 9](#_Toc102052803)

[Part C – Logic Model and Hypotheses: 11](#_Toc102052804)

[Section I Conclusion 15](#_Toc102052805)

[SECTION II – Process Evaluation 16](#_Toc102052806)

[Introduction: 16](#_Toc102052807)

[Coverage and Delivery 16](#_Toc102052808)

[Stakeholder Interests 17](#_Toc102052809)

[Process Evaluation Questions and Data Sources: 18](#_Toc102052810)

[Commentary 23](#_Toc102052811)

[Section II Conclusion 24](#_Toc102052812)

[SECTION III – Outcome Evaluation 25](#_Toc102052813)

[Introduction 25](#_Toc102052814)

[Definitions 25](#_Toc102052815)

[Possible Outcome Evaluation Study Designs 27](#_Toc102052816)

[Recommendation 31](#_Toc102052817)

[Section III Conclusion 31](#_Toc102052818)

[SECTION IV – Measurement 33](#_Toc102052819)

[Introduction 33](#_Toc102052820)

[Measurement of Outcomes 33](#_Toc102052821)

[Additional Comments 38](#_Toc102052822)

[Section IV Conclusion 39](#_Toc102052823)

[CONCLUSION 40](#_Toc102052824)

[REFERENCES 41](#_Toc102052825)

# ABSTRACT

*Background*

Children attending Title I Schools in San Antonio often demonstrate academic struggles. These children who come from lower-income families also are at risk for lower physical activity.

*Program*

The Earn-a-Bike @ School (EAB) program seeks to promote physical activity and improved academic achievement through a semester-long after-school program. The program consists of weekly after-school sessions which teach children about cycling, followed by a community event where children receive a brand-new bicycle. Attendance, good behavior, and coursework completion are used as incentives for children to receive their bicycle, promoting academic achievement while promoting children’s physical activity.

*Purpose of the Evaluation Plan*

The purpose of the evaluation plan is to present a model to evaluate the EAB program on the effectiveness of the program’s process as well as whether the outcome of increased physical activity in children is achieved.

*Methods*

A two-group quasi-experimental study alongside process evaluation measures is proposed to understand how the EAB program influences children to be physically active by cycling.

*Discussion*

Strengths of the proposed plan are that it presents a theory-based approach to evaluate coverage and delivery of the EAB program, it suggests new measurement techniques, and it presents an appropriate outcome evaluation design. Limitations are that it is not comprehensive in planning an evaluation for all EAB outcomes and does not include cost estimates. Future work should solicit the input of all stakeholders in the EAB program, expand the plan to cover all EAB outcomes, and prioritize which evaluation activities will be most important.

# INTRODUCTION

The worldwide pandemic of physical activity (Kohl et al 2012) extends to children and youth within local Texan communities. Children have distinct health risks from inactivity, and those risks will only compound as they continue patterns of inactivity into adulthood (Poitras 2016, Warburton & Bredin 2017). A coordinated approach has been suggested to promote physical activity among children and youth, focusing on multiple settings, times, and methods (Szeszulski, et al 2021). The Earn-a-Bike nonprofit in San Antonio, Texas, has focused on promoting cycling to improve the amount of physical activity children get outside of school by creating the after-school program Earn-a-Bike @ School (EAB hereafter). EAB seeks to empower students to cycle while also improving academic performance.

For any public health intervention, evaluation is a key process that informs continual improvement. With a properly designed evaluation, a program can better understand the population receiving the intervention, the process by which the intervention is being carried out, the outcomes of the intervention, and the theoretical basis for why those outcomes are being brought about. For Earn-a-Bike, evaluation can inform future areas of investment and ultimately allow its programs to better serve its target population.

The following document is a basic plan to evaluate the EAB program created by graduate students of the UTHealth Houston School of Public Health. The evaluation plan consists of three major sections. Section I will describe the health problem of physical inactivity, give an overview of the EAB program, and construct a theoretical framework through which the program can be understood. Section II will describe a plan for a process evaluation of the EAB program, which can be used to understand the reach of the program and how it is delivered. Section III will detail an outcome evaluation plan to evaluate the effect of the EAB program on children’s physical activity. The final section will provide evidence-based measures for each of the proposed outcomes to help track the effectiveness of the EAB program.

# SECTION I - Needs assessment, Program Description and Hypotheses

## Introduction

The first step to any evaluation is to assemble a foundation of knowledge. Evaluators need to understand the program they will be evaluating, and the associated health behaviors and health issues targeted. In this section of our proposal, we lay that foundation as a basis for the later plan we will build. The first section will define one of the major health problems targeted by EAB and describe the scope of the issue globally as well as its importance. The second section will provide a brief description of the EAB program. The final section will introduce a logic model that serves as a theoretical framework to understand how EAB’s programs effect change within their target population.

## Part A - Epidemiology of Program’s Health Problem

The Earn-a-Bike at School (EAB) program focuses on the main health outcome of helping children become more physically active by cycling. Physical activity has been defined as “behavior that involves human movement, resulting in physiological attributes including increased energy expenditure and improved physical fitness” (Welk, 2017). Therefore, for the purposes of this evaluation, we use the term physical *in*activity to refer to the health problem of having insufficient movement and energy expenditure, placing one at higher risk of deleterious health outcomes. Guidelines have been created by the US Department of Health and Human Services to identify people who have lower levels of activity and are at higher risk (US Department of HHS, 2018). These guidelines aid clinical and public health professionals in their efforts to promote physical activity. Physical activity guidelines for children have been defined as having 60 minutes or more of moderate-to-vigorous physical activity daily, broken into aerobic, muscle-strengthening, and bone-strengthening activities (Piercy, 2018). The EAB program focuses on helping children reach that physical activity goal.

EAB contracts with various elementary schools throughout San Antonio, Texas to implement their program. They focus on children in grades 3 to 5 who attend schools in economically disadvantaged areas of San Antonio, though they work with other schools as well if requested. Their target schools tend to have a higher proportion of students from minority racial and ethnic backgrounds and students living in families of low socioeconomic status. Schools identify specific grades they would like to implement the program in. The only inclusion criteria are that participants must be enrolled in a contracted school in the specified grade, and there are no exclusion criteria.

The prevalence of physical inactivity in children is high within the San Antonio area in which EAB’s target population resides. US national data estimates that 73.5% of boys and 80.6% of girls do not meet physical activity guidelines (Friel, 2020). Table 1 shows evidence from Friel, et al (2020) describing physical activity levels across various demographic categories. In general, boys are more active than girls and younger children than older children. There are few racial and ethnic differences in activity levels, but there is a trend showing greater activity levels among children of lower socioeconomic status. Children who are overweight and obese have lower activity levels than normal weight children. Data available for Texas states that 80.4% of youth overall do not meet guidelines (SPAN, 2021[)](https://sph.uth.edu/research/centers/dell/texas-child-health-status-report/physical%20activity%202021.pdf). The percentage of Texan children not meeting guidelines is relatively stable across ages (2nd grade: 77%, 4th grade: 91%, 8th grade: 24%, 11th grade: 79%; SPAN, 2020) The evidence for meeting guidelines is fair, as it does come from rigorous cross-sectional surveys but relies on parent proxy-report with one or two survey questions (Friel, 2020). Overall, the data available do correlate in informing that prevalence of inactivity within children is high.

**Table I.1. Estimated Prevalence of US Children Meeting 2008 Movement Guidelines**

|  |  |  |
| --- | --- | --- |
| **Category** | **Characteristic** | **Percent Meeting Guidelines (95% CI)** |
|  | Overall | 23.0 (22.1,23.9) |
| Sex | Female | 19.4 (18.1, 20.6) |
| Male | 26.5 (25.2, 27.8) |
| Age | Age 6-11 | 27.8 (26.4, 29.3) |
| Age 12-17 | 18.2 (17.0, 19.4) |
| Race/Ethnicity | White | 23.2 (22.2, 24.3) |
| Black | 24.8 (22.1, 27.4) |
| Asian | 16.4 (13.0, 19.7) |
| Other | 22.3 (19.4, 25.3) |
| Ethnicity | Non-Hispanic | 23.7 (22.8, 24.5) |
| Hispanic | 21.1 (18.5, 23.7) |
| Family poverty-to-income ratio level | <200 | 26.3 (24.6, 28.1) |
| 200-399 | 22.7 (20.6, 24.8) |
| 300-399 | 19.0 (16.9, 21.0) |
| >= 400 | 20.1 (19.0, 21.2) |
| BMI | Underweight (BMI <5th percentile) | 17.5 (14.3, 20.8) |
| Normal weight (5th-84th percentile) | 21.4 (20.0, 22.7) |
| Overweight (85th-94th percentile) | 16.5 (13.8, 19.2) |
| Obese (>= 95th percentile) | 14.5 (11.7, 17.2) |

Source: Friel, et al., 2020

Much effort has been put into researching determinants of children’s physical activity. One recent framework that has been applied is the COM-B model, a meta-theory that combines concepts from multiple individual behavioral theories. The COM-B model describes three constructs that combine to influence behavior: capability, opportunity, and motivation. Of the three, capability and motivation have specifically been correlated with physical activity. Capability refers to an individual’s capacity to perform a behavior, and motivation refers to all the processes of the brain which energize behavior; specific definitions of those two constructs can be found in Table 6. Both motivation and capability have been positively associated with physical activity in young adults (Howlett, 2019; Wilmott, 2021; see Table 2). The quality of these studies is low-to-moderate, one being a small prospective cohort trial and the other being a cross-sectional survey. Though sample sizes are smaller, correlation results show a relatively good association for this type of theoretical research. Framing the work of EAB by these two theoretical constructs will help inform the entire evaluation plan.

**Table I.2. Capability and Motivation determinants**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Determinant | Source | Sample Description | Study Design | Study Results: Measure of Effect |
| Motivation | Willmott, T. J., Pang, B., & Rundle-Thiele, S. (2021). Capability, opportunity, and motivation: An across contexts empirical examination of the COM-B model. *BMC Public Health*, *21*(1), 1014. <https://doi.org/10.1186/s12889-021-11019-w> | The study measures the effectiveness of the constructs of COM-B on physical activity and eating behaviors in young Australian adults aged 18-35 | Cross- sectional survey | Motivation has a positive correlation to physical activity  (β=.71, p < .001). |
| Capability | Howlett, N., Schulz, J., Trivedi, D., Troop, N., & Chater, A. (2019). A prospective study exploring the construct and predictive validity of the COM-B model for physical activity. *Journal of Health Psychology*, *24*(10), 1378–1391. <https://doi.org/10.1177/1359105317739098> | Adults 18 and above residing in the United Kingdom | Prospective cohort | Capability shows significant direct effect (β = .27, 95% CI, 0.09–0.50, p = .008) and indirect effect on MVPA (IE = .37, 95 per cent CI, 0.18–0.53, p < .001). |

Physical inactivity has been linked to an increased risk of a host of disease processes (Piercy, 2018). Two relevant consequences of physical inactivity in childhood are an increased risk of obesity and diabetes (See Table 3), each of which carries their own risks and health-compromising consequences. Lower physical activity has been clearly associated with an increased risk of obesity in longitudinal studies of adults (odds of developing obesity among active adults are 0.64 and 0.63 at 5 and 10 years, respectively; Bell, 2014). Among children, current evidence has been slightly less clear, but there are several studies which have found links between lower physical activity with future risk of obesity (Bleich, 2011). Physical inactivity has also been clearly associated with increased risk of developing diabetes among adults (Cleven, 2020). Like obesity research among children, there is at least initial evidence showing links between physical inactivity in childhood and incident diabetes. Children with greater volume of physical activity had decreased insulin resistance among children (b=-0.30 +/- 0.14, p= .044; Krekoukia, 2007). The quality of evidence linking physical activity to outcomes of obesity and diabetes is excellent, with multiple longitudinal cohort studies showing trends. Though there are fewer studies linking that evidence back to childhood, the trend is clear. The deleterious effects of physical inactivity are undeniable, giving precedence to intervening as early as possible.

**Table I.3. Evidence table for consequences physical inactivity**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Consequence | Source | Sample Description | Study Design | Results: Measure of Effect |
| Increased Risk of Obesity | Bell, J. A., Hamer, M., Batty, G. D., Singh-Manoux, A., Sabia, S., & Kivimaki, M. (2014). Combined effect of physical activity and leisure time sitting on long-term risk of incident obesity and metabolic risk factor clustering. *Diabetologia*, *57*(10), 2048–2056. <https://doi.org/10.1007/s00125-014-3323-8> | 3670 adults, 73% male, 55.5 ± 6.0 years | Three longitudinal studies | Lower odds of developing obesity with high levels of PA compared to low after 5 and 10 years (OR 0.64, 0.63, respectively). |
| Increased Risk of Obesity | Bleich, S. N., Ku, R., & Wang, Y. C. (2011). Relative contribution of energy intake and energy expenditure to childhood obesity: A review of the literature and directions for future research. *International Journal of Obesity*, *35*(1), 1–15. <https://doi.org/10.1038/ijo.2010.252> | Children | Review | Variability in evidence of how nutrition and PA affect childhood obesity, but many studies have found links. |
| Increased Risk of Diabetes | Cleven, L., Krell-Roesch, J., Nigg, C. R., & Woll, A. (2020). The association between physical activity with incident obesity, coronary heart disease, diabetes and hypertension in adults: A systematic review of longitudinal studies published after 2012. *BMC Public Health*, *20*(1), 726. <https://doi.org/10.1186/s12889-020-08715-4> | Adults | Review of longitudinal cohort studies | Out of 11 longitudinal studies published since 2012, 9 studies found that higher levels of physical activity were associated with lower risk of developing diabetes. |
| Increased Risk of Diabetes | Krekoukia, M., Nassis, G. P., Psarra, G., Skenderi, K., Chrousos, G. P., & Sidossis, L. S. (2007). Elevated total and central adiposity and low physical activity are associated with insulin resistance in children. *Metabolism*, *56*(2), 206–213. <https://doi.org/10.1016/j.metabol.2006.09.014> | 28 boys and 26 girls (27 obese and 27 lean) ages 9- 11.5 years |  | Childrens’ total volume of physical activity negatively associated with insulin resistance (b=-0.30 +/- 0.14, p= .044). |

Understanding the scope and effects of physical inactivity in childhood provides a strong motivating force for intervention. There is an ever-increasing need for resources to be devoted to helping children learn to become more active. The Earn-a-Bike nonprofit has found one way to help fulfil that need through its Earn-a-Bike @ School program.

## Part B – Program Description

The Earn-A-Bike @ School (EAB) is an incentive-based program designed to focus on attendance, coursework, positive behavior, health, and fostering self-empowerment. The program incorporates nutrition, safety, maintenance, and wellness components laced with the common thread of working towards a goal and earning it. The program is 10 weeks long and comprises nutrition and physical activity education wrapped up by a separate five-hour-long Build-a-Bike event.

The educational component comprises after-school 45–50-minute lectures about the following topics: Intro to Bike Knowledge, Nutrition, Bicycle Safety, Helmet Safety, Road Safety, How to Patch a Tire, How to Build a Bike, Recap. Lectures are delivered in person, within the classroom. During these lectures, materials such as bike parts, road signs, MyPlate sheet, Bicycle repair kits are used to give children guided practice of the concepts they are learning. After the lectures, the students are given tasks to complete at home and report their progress through a journal of the activities they completed.

While the students complete the educational aspect, in-school attendance logs are tracked by teachers and staff to make sure students are participating in school. Other documents such as referrals, write ups, awards/recognitions will also be tracked to ensure good behavior in students. EAB staff also track students’ attendance at the after-school program and completion of EAB-specific coursework.

At the end of the 10-week program is a 5-hour long Build-a-Bike event. This event features students that have successfully completed the program. Students partner with a volunteer to assemble a brand-new bicycle which is then used in a 2-mile group ride. Parents of the participating students engage in educational workshops about health promotion and community resources about physical activity and nutrition while their children assemble the bikes.

The program is estimated to cost about $25,000 – 39,700 for one grade and one semester. The budget is set for 100 bikes, tools to assemble bikes, decorations, bike helmets, food for volunteers, entertainment, shirts, bicycle kits, and transportation of event materials.

Program participants are recruited through outreach to the school directors to agree upon hosting the program at the school. An in person/ online meeting is held with school board members to explain program intent and execution plan. Emails are sent out to teachers and parents to gain their interest. Fliers are posted within schools to attract more students.

The target population is 3rd- 5th graders in title 1 schools that are at risk or economically disadvantaged. The program expects to serve about 100-180 participants based on the previous record and budget size.

Students not in 3rd to 5th grade will be automatically excluded based on eligibility. Even though the program aims to actively reach economically disadvantaged schools, any school that shows interest in the program is accepted into the program.

There are multiple stakeholders involved in the EAB program which may have an interest in the results of a formal evaluation. Table 4 shows a list of stakeholders that can benefit from the evaluation of the Earn a Bike at school program and potential priorities they may focus on.

**Table I.4. Stakeholders and evaluation priority**

|  |  |
| --- | --- |
| Stakeholder | Likely Evaluation Priority (What is their priority?) |
| Earn-A-Bike Nonprofit | * Evaluate efficacy of the program for improving child health and nutrition * Determine efficiency and effectiveness of program process * Gather data for future funding applications |
| Schools | * Evaluate outcomes for increasing attendance and improving behaviors * Gather data for future funding applications |
| Teachers | * Evaluate outcomes for improving student behavior * Determine efficacy of the program in improving classroom participation and grades |
| VIA transportation | * Determine the number of participants to allocate seats to during the event * Collect log of families that need public transportation within the area |
| San Antonio Food Bank | * Evaluate the amount of Myplate sheets are used throughout the program * Determine the state of healthy food access within the community * Estimate the number of parents to reach during the Build-a-Bike event |
| YMCA | * Estimate the number of parents to reach during the Build-a-Bike event |

## Part C – Logic Model and Hypotheses:

A logic model is a visual diagram of how a program should theoretically work to benefit participants. As one source describes, "It is the ‘If-Then’ sequence of changes that the program intends to set in motion through its inputs, activities, and outputs” (United Way of America, 1996, p. 38). The model is informed by social and behavioral theory as well as evidence of benefit based in current scientific literature. Here, we present the logic model of specifically how the EAB program affects behavioral change in influencing children to become more physically active. We use a tabular format to easily visualize program inputs, activities, outputs, and outcomes. As the current evaluation plan focuses only on the physical activity outcome, all components of the actual EAB program are not included in the outcomes and potential influences. The logic model is as follows:

**Table I.5. Logic Model for Earn A Bike at School**

| **Inputs** | **Activities** | **Outputs** | **Short-Term Outcomes** | **Intermediate Outcomes** | **Long-term Outcomes** |
| --- | --- | --- | --- | --- | --- |
| *What is necessary to have for this program to operate* | *Actual events or actions that take place to create the desired results* | *Products of program activities*  *Data on activities of the program* | *Psychosocial impacts of the program* | *Behavioral impacts of the program* | *Quality of life / health outcomes* |
| Time:   * After school sessions * One weekend day for event   Physical space   * Classrooms * Event location   Funding   * Currently under a co-grant with St. Mary’s (survey development), no follow up requirement * $25,000 pre-covid, $39,700 post-covid (100 bikes) * Donations, co-grant   Personnel   * Directors * Volunteers * School Admins * Parents * Hired EAB Teachers / College Interns   Materials   * Bicycle Kits * Shirts * Laminated MyPlate sheet * Built in online nutrition lesson * Surveys * Event: entertainment, transportation of bikes, helmets, decorations, food for volunteers, etc.   Equipment   * Tools to assemble bicycles   Relationships   * With School * With Donors * Community Partners – Food Bank, VIA bus, YMCA, etc. | Weekly classes for 10 weeks at school   * After school, 45-50 minutes * Didactics plus activities to solidify information * Lesson 1: Intro to Bike Knowledge * Lesson 2: Nutrition\* * Lesson 3: Bicycle Safety * Lesson 4: Helmet Safety * Lesson 5: Road Safety * Lesson 6: How to Patch a Tire * Lesson 7: How to Build A Bike * Lesson 8: Recap   Tracking Education ABC’s:   * Attendance: perfect or increasing attendance at school\* * Behavior: tailored to school but can include office referrals, write-ups, etc.\* * Coursework: EAB- specific work\*   Build-a-Bike Event:   * Assign child to volunteer * Assemble bicycles with volunteer * Parent educational workshops on health promotion, community resources, etc. * 2-mile group ride | Delivery outputs:   * # of sessions delivered * Number of delivered sessions that go overtime * Number of Build-a- Bike volunteers signed up * Number of Build-a-Bike volunteers attended * Educational session student to instructor ratio   Coverage outputs:  Education ABC results:   * Attendance Number of absences in school\* * Behavior score (Number of bad behavior episodes- school determined)\* * Coursework completion rate - Number of EAB assignments completed\*   Number of children participating   * # registered to participate in the EAB program * # who attend each session * # who complete coursework * # who attended the final Build-a-Bike event. | Increase capability to:   * Cycle * Practice bicycle safety * Assemble and maintain a bicycle.   Increase Motivation to:   * Cycle * Learn about safe cycling practices | Increase in cycling activity | Decreased risk of obesity |

*\*The activities that target nutrition and educational outcomes were not evaluated in this plan but would be in a more comprehensive evaluation plan*

The theoretical constructs that we use to give context to EAB’s functions center around the COM-B system developed by Michie, et al (2011). The COM-B system describes how three constructs, capability, opportunity, and motivation, work together to influence behavior. The two constructs most relevant to EAB are capability and motivation, defined in the table below:

**Table I.6: Theoretical Constructs and Definitions**

|  |  |  |  |
| --- | --- | --- | --- |
| **Short term outcome** | **Theory** | **Conceptual Definition** | **Citation** |
| Capability | COM-B | The individual’s psychological and physical capacity to engage in the activity concerned. It includes having the necessary knowledge and skills | Michie, et al. (2011) |
| Motivation | COM-B | All those brain processes that energize and direct behavior, not just goals and conscious decision-making. It includes habitual processes, emotional responding, as well as analytical decision-making. | Ibid. |

Based on our logic model, we have identified the following hypotheses that the EAB intervention addresses:

1. Student educational activities that include 8 didactic educational sessions, completion of online modules, and guided practice will result in an **increase in capability** to cycle, correctly and safely ride a bicycle, and assemble and maintain a bicycle.
2. Student activities that include an interactive skill-building session to assemble a bicycle and an enactive mastery experience of a group 2-mile bicycle ride will result in an **increase in capability** to successfully assemble and maintain a bicycle and be physically active (by cycling).
3. Student activities that include an interactive skill-building session to assemble a bicycle and an enactive mastery experience of a group 2-mile bicycle ride will result in an **increase in motivation** to cycle and learn about safe cycling practices.

Causal hypotheses:

1. Increased student capability to cycle, practice bicycle safety, assemble and maintain a bicycle will lead to students engaging in more cycling activities.
2. Increased student motivation to cycle and learn about safe cycling practices will lead to students engaging in more cycling activities.

## Section I Conclusion

The logic model depicts the structure of the program, and provides a logical sequence of events including the program elements. Having this laid out appropriately will provide a solid foundation for setting up a rigorous evaluation of the program. Using this model, we will be able to dive further into evaluating the implementation of the program in a process evaluation.

# SECTION II – Process Evaluation

## Introduction:

Process evaluation involves monitoring the way a program is carried out. It covers every step of the program's implementation and provides a way to analyze the linkage from input to program activities to the program's intended outcomes. By evaluating the process, or implementation of a program, one can determine why a program was successful and why it was not successful and gives information about specific areas within the program that can be altered to make it more effective (Saunders et al 2005). In this section, we will present examples of how to evaluate the coverage and delivery of the EAB program, indicate the stakeholders' interest in the process evaluation, suggest important questions to measure each step in the program process, and clearly state the measurement methods for these steps.

## Coverage and Delivery

An important use of the process evaluation is for its summative benefits. Summative benefits allow the implementers and stakeholders to judge whether the program was carried out properly, to what extent it was carried out, and whether it reached the intended audience. Knowing this information is key for explaining the results of the program outcome (Saunders et al 2005). Coverage and delivery of the program are essential components of the summative aspects of process evaluation, **Coverage** is defined as the extent to which the target population is participating in the program services; **delivery** refers to how well the implementers carried out the program. Errors in the recruitment process such as improper selection of participants, poor recruitment methods, or participants’ negative judgment of the program may cause the program to include fewer participants from the intended audience or additional participants outside the intended audience.

Recruiting fewer than expected numbers of the target population is called **under-inclusion**. Based on EABs recruitment format to accept all students in grades 3-5, under-inclusion is not a concern. **Over-inclusion** occurs when participants from outside the target population end up participating in the program. EAB at school has a relaxed inclusion criteria for students in grades 3-5 from title 1 schools. Given the relaxed nature of these criteria, students from schools which are not economically disadvantaged can participate, opening the chance for over inclusion to be a problem. Over-inclusion only poses a challenge if the intended budget and resources do not reach enough of the students from title 1 schools. To solve the issue of over-inclusion, the evaluation can monitor the number of participating schools that are considered title 1 schools, measure the number of students within these title-one school that got through the entire process as well as the number of students from non-title 1 schools that go through the program.

## Stakeholder Interests

Due to the multifaceted nature of the program, multiple stakeholders have an interest in various points in the process evaluation steps.

The EAB organization: These are the proprietors of the program, they may be interested in knowing the number of students who were aware of the program, registered for it, participated in it, and saw it through to completion, the number of title-1 schools that participated in the program to increase their level of reach in future events.

Participating schools: Another major stakeholder is the school in which the program is carried out. Given that they are the ones providing the major program participants, they may also be interested in measuring the participation and completion rates of the students. Data such as how many students signed up, demographics of students that signed up and those that did not, how many completed each program, if there were dropouts, and what are common reasons for dropping out may help the school and EAB decide if students that went through the entirety of the program had better outcomes and if the program is beneficial for their students.

VIA Metropolitan Transit: One of the stakeholders that are partnered with EAB to provide bus rides for students to and from the final event. They may be interested in participation numbers for the event to determine the number of bus allocation slots to make available for the final Build-a-Bike event.

Parents of the participating students: This group may be interested in knowing how effective the program implementation was, i.e., did the volunteers deliver each lecture as planned, were any concepts missed, or were missed concepts important.

Other stakeholders include physical activity and nutrition-related organizations like YMCA and the San Antonio food bank that present to parents during the final Build-a-Bike event. These stakeholders may be interested in the attendance of parents at that event specifically.

## Process Evaluation Questions and Data Sources:

The process evaluation is guided by specific questions to be answered which can help improve coverage and delivery. Table II.1 contains possible questions that the process evaluation can answer regarding **coverage** of the target population, and Table II.2 contains possible questions to answer regarding **delivery**. Each of the tables contains suggested data sources that can answer the questions posed.

**Table II.1: Program Coverage Topics, Process Evaluation Questions, and Data Sources:**

|  |  |  |
| --- | --- | --- |
| Topic | Relevant Question(s) | Data Source(s) |
| a) Awareness of program/policy | What percentage of title-1 schools In San Antonio have already participated in EAB?  What percentage of title-1 school principals are aware of EAB?  What percentage of school parents are aware of the program? | Records of participating schools in the program’s history  Survey of principals |
| b) Program participation | What percentage of students signed up for the program?  How many children attend 50%, 75%, and 100% of the after-school sessions?  How many students attended the final Build-a-Bike event?  What percentage of children who finish all the after-school sessions attend the Build-a-Bike event?  What percentage of children eligible to attend the Build-a-Bike event do not, and why?  What percentage of students are not allowed to participate in after-school activities and therefore are excluded from participation due to school policy restrictions?  What percentage of students who participate in the program are from title 1 schools?  What are the socio-demographic characteristics of participants (% of students by Race/Ethnicity, Household Income, age/grade, primary language) | Registration Forms  EAB Instructor Logs    Build-a-Bike attendance Log  EAB Instructor Logs  Build-a-Bike attendance log  Build-a-Bike attendance  Parent Interview  School records  Registration forms  Registration form |
| c) Non-participation | What percentage of Title I schools do not participate in program activities?  What percentage of grades 3-5 in San Antonio Title 1 schools do not participate in program activities?  What percentage of each grade implementing EAB (3rd, 4th, and 5th) do not participate?  What were the reasons for non-participation among eligible students | Program participation logs  ISD-level data  EAB records      School records  EAB Registration Data  Parent phone interview |
| d) Dose of program/policy received | How many children attend 50%, 75%, and 100% of the after-school sessions?  What percentage of students received a new bicycle at the end of the program?  What percentage of participating students attend the final Build-a-Bike event? | EAB Instructor Logs  Inventory list  Event attendance list |
| e) Drop-outs | What percentage of students that registered did not qualify/meet requirements to receive a bike at the end of the program?  What were their reasons for dropping out? | Attendance logs  Parent phone interview |
| f) Differences across sites, time, location | How do the registration percentages within each grade compare across schools?  How did registration percentages differ pre-COVID vs currently? | School records  EAB Registration records  EAB Registration records |

**Table II.2: Program Delivery Topics, Process Evaluation Questions, and Data Sources:**

|  |  |  |
| --- | --- | --- |
| Topic | Relevant Question(s) | Data Source(s) |
| a) Quality and accuracy of services delivered | Frequency:  What percentage of the sessions planned were delivered?  How often are sessions presented in the order planned?  If the order changes, which sessions are most commonly shuffled/combined, and why?  Accuracy:  What percentage of planned activities are performed in each session?  Which session activities are most likely to be skipped, and why?  Were any activities not planned added into the session, and why? | EAB Instructor Logs  EAB Instructor Logs  EAB Instructor Logs  EAB Instructor Logs  Direct observation by EAB staff |
| b) Duration of program delivery | How many times are there more than 1 week between sessions?  How does the timing of delivery differ between fall and spring semester?  Which major events most commonly alter the scheduling of weekly sessions?  How many sessions go over the allotted time?  Which sessions tend to be the ones which go long? Why? | Calendar of Events  Calendar of Events  Calendar of Events  EAB Instructor Logs |
| c) Staffing and training | What is the student-to-instructor ratio at the sessions?  What is the average length of service of instructors?  How confident do teaching staff feel in:   * Delivering the content of the sessions? * Leading children through activities? * Maintaining order in the sessions?   If they are lacking in confidence in any of these areas, why do they feel that way?  How satisfied are instructors with:   * Their training in the content knowledge they will deliver? * Their training in classroom management? * The instructor-to-student ratio at the sessions?   How many volunteer bicycle assemblers signed up for the Build-a-Bike activity?  How many volunteers attended?  What is the volunteer-to-student ratio at the Build-a-Bike event?  How confident do volunteer bike assemblers feel in:   * Connecting with their assigned child? * Assembling a bicycle? * Coaching the student to assemble their bicycle themself?   If they are lacking in confidence in any of these areas, why do they feel that way?  How satisfied are the volunteer bicycle assemblers with their orientation for the Build-a-Bike event, including:   * Communication prior to the event * Explanation of the schedule * Explanation of their role   What improvements do they suggest? | Attendance Logs  Internal volunteering documentation  EAB Instructor survey  EAB Instructor Survey  EAB Instructor Survey  Volunteer Signup  Volunteer Sign-in  Student Sign-in  Volunteer Survey  Volunteer Survey  Volunteer Survey  Volunteer Survey |
| d) Materials, supplies, facilities, & resources | At the educational sessions, what % of sessions ran out of:   * Handouts? * Supplemental educational materials? (models, kits, etc.) * Teaching materials? (Whiteboard markers, chalk, paper and pencils, etc.)   What % of session did not have enough classroom space to give students an appropriate educational environment?  Is the Build-a-Bike event location adequate for the event, including   * Accessibility by participants? * Amount of space for bicycle assembly? * Availability of outdoor space for group bicycle ride?   At the Build-a-Bike event, are there sufficient:   * Bicycle kits? * T-Shirts? * Swag Bags / handouts? * Food for volunteers?   How many participants are provided VIA passes to attend the event?  How many participants require transportation assistance to and from the location?  How many participants require assistance to transport their bicycles home? | Direct Observation by EAB staff  Instructor Logs  Direct Observation by EAB staff  Direct Observation by EAB staff  Parent survey  Volunteer Survey  Supply logs  Attendance Logs  Build-a-Bike Registration forms  Build-a-Bike Registration forms  Build-a-Bike Registration forms |
| e) Consistency across sites and staff | How does the instructor-to-student ratio compare across sites?  How does the percentage of sessions delivered compare across sites? | Attendance Logs  EAB Instructor Logs |
| f) Participants’ perception of delivery channel | How excited are students to attend the educational sessions?  ~~How much do students feel that they are learning in the sessions?~~  How engaging do students feel that the instructors are?  How much do students relate with the instructors?  ~~How engaging do students feel the educational materials are?~~ | Student interviews  Student interviews  Student interviews  Student interviews |
| g) Participants’ satisfaction with various service components | How satisfied are students with the educational sessions?  How satisfied are students with the Build-a-Bike session?  How satisfied are students with their volunteer at the Build-a-Bike session?  How satisfied are parents with the community resource training at the Build-a-Bike event?  How satisfied are parents with the program as a whole? | Student Interviews  Parent survey |
| h) Coordination with other programs or agencies | Schools:  How satisfied are the schools with the EAB program?  Is coordination between schools and EAB adequate?  VIA  Is there appropriate coordination of transportation needs for the Build-a-Bike event?  How soon prior to the event are the number of needed bus passes requested and received?  Food Bank  How soon prior to the nutrition session are materials requested and received? | School Principal Interviews or Survey  Interview with VIA representative  Email / phone records  Email / phone records |
| i) Protocols and procedures in place | How are protocols and procedures documented?  How are changes in protocols and procedures requested and approved?  How are changes in protocols and procedures communicated to staff and volunteers? | Review of EAB documentation |

## Commentary

The above process evaluation plan trends more towards being comprehensive than feasible. Completing the entire plan would take a significant amount of time, personnel, and funding to develop instruments, collect data, and analyze data. Multiple data sources would be required, many of them new. The following list summarizes the possible data sources from the above tables:

During Program Delivery:

* Registration forms for students, including school attended and basic demographics
* An instructor log during each session tracking attendance and delivery of session components
* A direct observation component where EAB staff would occasionally monitor the program, including quality of delivery, adequacy of space, and adequacy of materials.
* Student interviews about satisfaction with sessions and instructors
* Attendance logs for both participants and volunteers at the Build-a-Bike event
* Calendar of events detailing when each session was held

After Program Delivery:

* EAB instructor survey to determine confidence in abilities and satisfaction with training
* Build-a-Bike volunteer survey to examine confidence in abilities and satisfaction
* Parent survey at the end of the program to determine satisfaction
* Survey or interview with school principals
* Phone interviews with some parents, including parents of participants, non-participants, and those who did not complete the program

Other:

* School records, including enrollment in each grade that participates in the EAB program
* District-level data on the number of Title I schools in the target region
* Internal documents including protocols and procedures
* Internal records of length of service for volunteers and materials purchased

EAB will likely wish to prioritize components of the process evaluation which are most feasible and applicable to their current situation. One topic that may be a priority is the delivery of the educational sessions to ensure that the program is given as accurately and successfully as possible. Prioritizing the topic of delivery would likely have the greatest effect on the outcomes that are expected from the program. EAB may also prioritize evaluating the process of the Build-a-Bike event since it is a highly visible community event that can develop a large amount of support to help with the continuation of the program. By focusing on key topics, EAB can make the most of its limited resources available for evaluation.

## Section II Conclusion

Diving into the process of the EAB program will help identify both strengths and opportunities for growth. Performing the process evaluation will benefit stakeholders by developing a better understanding of the protocols and extent of the program. Exploring the coverage of the program will help identify how many participants are involved, who is being involved, and who is not being included. Evaluating the delivery of the program will identify whether protocols are being followed properly and whether changes in protocol are needed. It will also inform EAB about the participants’ satisfaction with program delivery. The information gleaned from this portion of the evaluation will help determine the efficiency of the EAB process so that there is a greater likelihood that the expected outcomes will be realized.

# SECTION III – Outcome Evaluation

## Introduction

In an outcome evaluation, a program is studied to understand whether it is having the effect that it intends within the target population. Section III will discuss how to evaluate the outcomes of the EAB program which relate to physical activity and cycling. We will first define key terms to give context to the possible study designs that we will present. We will give a recommendation on the type of outcome evaluation study design that will be most appropriate for EAB, and we will conclude with suggestions on possible ways to measure the program’s outcomes.

## Definitions

We will first start with definitions of concepts surrounding study design which will help explain rationale for recommendations we will give for outcome evaluation design.

*Internal Validity* is the degree to which a study controls outside /confounding variables and allows causal inferences to be made between the independent and dependent variables (Flanelly 2018). For an evaluation study of the EAB program, internal validity would exist if there was a proper study design and all biases were minimized. *External Validity* is the degree of generalizability of a research result to external populations that are similar to the sample population in terms of context, individuals, time and setting. It could also be compared to the population that the sample was taken from (Lavrakas 2008). The population targeted by the EAB program is kids in grades 3-5 from title 1 schools. Based on this criterion, we would expect a decent level of generalizability to economically disadvantaged children in elementary school. *Random assignment* is defined as a type of assignment in a study where each member in the study sample has an equal chance of participating in a specific group or condition (Salkind 2010). In an experimental design where a control group and an intervention group exist, each participant will have the same probability of entering either one of the groups through random assignment. In the EAB project, a random assignment would apply if certain title 1 schools were randomly selected from the pool of title 1 schools to serve as control and intervention groups. *Random selection*occurs in a situation where the sample group is chosen from the general population such that every member of the general population has an equal chance of participating in the study (Lavrakas 2008). Random sampling would not apply in the EAB at school program because the inclusion criterion that is title 1 schools, excludes other schools in San Antonio.

*Experimental study vs quasi-experimental study:* In order to fulfill the criteria for an experimental study, participants within the design have to be randomly assigned into groups within the study (Salkind 2010). In a situation where the participants within the study are assigned to groups using a non-random method, it is known as a quasi-experimental study or a non-randomized pre-post intervention study (Harris 2006).

Typically, random sampling is used to increase the internal validity of a study, however, some threats to internal validity may exist. These threats include,

*Maturation:* Maturation refers to bodily/ biological changes that occur with the passage of time

*History:* History refers to the things a participant experiences that are not part of the experiment but can affect the results. When historical events affect participants during the course of the experiment that change outcomes in any way, internal validity is threatened.

*Testing:* Testing refers to taking measures of study outcomes in participants. Testing threatens internal validity because in a pre-post test design, testing participants before may affect the way they respond to the question if posed after the study.

*Instrumentation:* In an instance where an instrument is used to take outcome measurements, decay in the instrument over time might pose a risk to internal validity

*General Selection bias:* When there is a potential bias in assigning members to control or experimental groups, the groups will differ in important ways, and this poses a threat to the internal validity of the study.

*Attrition:* Attrition refers to the loss of study participants in experimental or control groups. In most cases, higher dropout rates exist among the experimental group due to increased demands and commitment requirement. This skewed attrition will affect the internal validity of the study. (Flanelly 2018)

*Secular trends*: Secular trends refer to the occurrences in society not produced by the intervention that may lead away from or towards the desired outcome (Fallin 2011).

*Differential threats:* In a situation where these threats to internal validity occur differentially i.e, in only one of the study groups, a dramatic difference will be noticed, and this difference may be measured as an outcome. An example is with attrition, if attrition happens non-differentially, there is an equal distribution in control and experimental group. Differential attrition in the experimental group will end up reducing the effect of the study in that group making it seem like the program has little to no effect.

## Possible Outcome Evaluation Study Designs

*Experimental Design*

In an experimental setting, the EAB program would first have to randomly select intervention and control schools out of the pool of title 1 schools in San Antonio. Using a pre-post format, we would assign a pre-participation survey assessing the students’ physical activity measures. After implementing the program in the intervention schools, we will take another physical activity measurement for both groups and compare the outcome of the program using statistical analysis. The format for the intervention group has a design notation , R representing the random selection, O representing the assessment points and X representing the intervention. For the control group the notation is as follows:. Because of the nature of the PAQ-C questionnaire which was recommended for the pre and post-test evaluation, the post-test evaluation will be conducted 7 days after the intervention ends.

The *advantage* of using an experimental study is that it allows the EAB program to determine the relationship between the program activities and desired outcomes. Given that the program is run with limited resources, all title 1 schools cannot receive the intervention at the same time and it will be convenient to take measurements from non-participating schools through an experimental study. Additionally, through random assignment, most of the threats to internal validity are drastically reduced, including;

Selection bias: Picking schools at random eliminates the possibility of preferentially selecting schools for participation.

Maturation: Students in grades 3-5 would generally undergo similar biological processes as they age, having a similar control group where the intervention is not introduced allows EAB to measure if maturation makes a difference in physical activity during the intervention period.

History: Given that the target schools are relatively from a similar demographic, changes in the community over time will apply similarly across the board. If these changes show up in the control group, they are accounted for and separated from the program’s outcome.

*Disadvantages* of conducting this program using an experimental design are: Given the previous success of the program, it will be difficult to have a school agree to be a control group that deliberately sits out of participating in the program. Providing an incentive to the schools that participate in the control group could be a work around to this potential limitation. Despite the reduction in threats to internal validity that an experimental study provides, other threats might persist, including:

Attrition: Differential attrition might be a problem with experimental studies since most of the task for participation are carried out by the schools in the intervention group, hence if they feel that the work is overbearing, they might choose to drop out.

History: A potential issue with this program might arise because of the mode of recruitment. Although each school has a similar demographic of students, they are also enclosed communities that might experience historical changes that are different from other schools in the same demographic. If changes happen within the community of one school that affect the outcome of the results but not in other schools, a differential effect is noticed which affects the internal validity of the program e.g some schools have a policy that only students signed up for after school programs within the school can participate in any after school program including EAB at school, this might potentially affect the number of students and type of students that participate in the program.

*One-group Quasi-Experimental Design*

Another design possibility for testing the outcomes of the EAB program is a one-group quasi-experimental design. In this design, there is one group of participants who receive the intervention with no comparison group and no randomization. These participants would undergo an assessment prior to the intervention and after the intervention to determine whether a meaningful change had occurred. A diagram of the design, again with O representing observations and X representing intervention, would be the following: . This design is most similar to what EAB currently has in place. Participants would be given a baseline observation, such as a survey, the intervention would be carried out, and then a follow-up observation would be done at the end of the program.

The main *advantage*to using a one-group quasi-experimental design is that it is less expensive to and simpler to carry out, as there is only one group and no randomization protocol. Data collected can help develop a rough idea of the effect of the program. In addition, some types of studies cannot adequately fit a randomized model, for example, if there were only two schools involved in the study, you would not be able to randomize them to two groups.

A major *disadvantage*to the one-group quasi-experimental design is that the internal validity, or confidence we have that the intervention itself is causing the desired effect, is low. Without a control group, we cannot be certain if changes in cycling activity are meaningfully different than if the program was not implemented because of possible internal threats of history, maturation, secular trends, or others. As examples, a change of season could influence physical activity outcomes. Even asking participants the pre-test questions themselves could create a change in behavior. To minimize sources of measurement bias, the same staff members can administer the same test before and after the intervention.

*Two-Group Quasi-Experimental Design*

A last possibility for a design to measure outcomes is a two-group quasi-experimental design. In this design, two groups are created without random assignment: an intervention group and a control group. The control group is chosen to match as closely as possible with the intervention to minimize confounding effects from selection bias. Each group is given a pre-test to measure baseline characteristics. Only the intervention group is then given the intervention, and both groups are then given a post-test to measure outcomes. A diagram of this design is as follows:

. NR represents the non-randomized selection of groups.

To carry out this design, EAB would recruit a typical intervention group but then carefully select a control group for comparison. One simple but flawed idea to create a control group would be to split a grade in half, give one half the intervention, and have the other half be the control. Not only would this design be counter to how EAB runs its program, but there would be potential for contamination, or for students who are receiving the intervention to share pieces of the intervention with students who are acting as controls. Instead, EAB could consider finding a separate school whose students are demographically similar to the intervention school and recruit a group of students in the same grade as the intervention group to create a control group. This control group would be given the pre- and post-assessment but not the intervention.

*Advantages* to the two-group quasi-experimental design is that it has increased internal validity compared to a one-group design and it is more feasible to conduct than a randomized experiment. Pre-post differences could be compared between intervention and control groups to give a sense of the effect of the intervention. Having a control group reduces the effects of internal validity threats of maturation, history, and instrumentation, which is an improvement on the one-group design.

There are still *limitations* to a two-group quasi-experimental design, the main one being selection bias. Since the evaluator would choose the control group or participants may select themselves into intervention or control groups, there may be hidden differences between groups that could affect the results of the study. Choosing a control group that is similar in characteristics such as age, grade, and other characteristics can help minimize the selection bias created. Though it is virtually impossible to create a perfectly matched control group with this design, there are still benefits to creating a group that is as similar as possible. Testing could also still affect outcomes, as the pretest could influence participants to change their behavior.

## Recommendation

EAB currently employs a one-group quasi-experimental design in its program to evaluate outcomes. This design is most cost-effective and gives good data for program effect. Continuing this design for an evaluation would give a fair assessment of outcomes for the program, however the data would be subject to several threats to validity, as mentioned above. An experimental design would not be appropriate because of an inability to adequately randomize as well as the expense involved. To improve validity of outcome data while considering expense, we recommend utilizing a two-group quasi-experimental design to give an evaluation of whether the EAB program has its intended effects. To perform the two-group design, the main logistical challenge would be to create a matched control group. EAB would need to identify a demographically similar school and then create a control group of students in the same grade as the students receiving the EAB program. Two observations would then need to be collected from the students in the control group. In some evaluations of similar school-based programs, monetary incentives are used to recruit a school to create a control group. The intervention school could then undergo a typical EAB program with the evaluation activities throughout.

## Section III Conclusion

Understanding the possibilities for outcome evaluation design helps the evaluator and program to come to an informed decision on what will be the most appropriate course to take. The final evaluation design must strike a balance between quality and feasibility. In reviewing possibilities of evaluation designs, we suggest carrying out a two-group quasi-experimental design to evaluate the outcomes of how the EAB program impacts children’s cycling activity. Such a design would be the most cost-effective at producing valid outcome data.

# SECTION IV – Measurement

## Introduction

In this section, we will identify appropriate outcome measurements for the physical activity component of the Earn a Bike at school program. Measures will be defined for both short and intermediate-term outcomes and the validity and reliability of these measures will be assessed as part of the outcome evaluation process. **Reliability** of a measure is defined as how prone the measurement is to error. A reliable measure gives equivalent results after repeated measurements are taken under several conditions (Scholtes 2011). The **validity** of a measure refers to the degree to which an instrument measures the construct that it claims to measure. Validity consists of three properties, content validity which refers to how accurately the content of the measure reflects the construct, construct validity which estimates how consistent the measurements results are with the hypothesis given the tool measures the construct and criterion validity estimates how representative the result of a measurement is of a gold standard measurement.

The current evaluation plan will include suggestions for tools that measure psychological concepts such as capability and motivation. To measure these concepts, multiple survey questions are used in combination to attempt to represent the true state of the respondent. Specific statistical methods can be used to assess the reliability and validity of these tools. For example, Cronbach’s alpha is a reliability statistical method that computes the level of internal consistency between the related questions, meaning the level to which respondents answered similarly on each question measuring a single concept. As another example, confirmatory factor analysis is a validity statistical method that computes how well a group of questions measures the concept of interest. We will present information on the reliability and validity of each measurement tool we suggest.

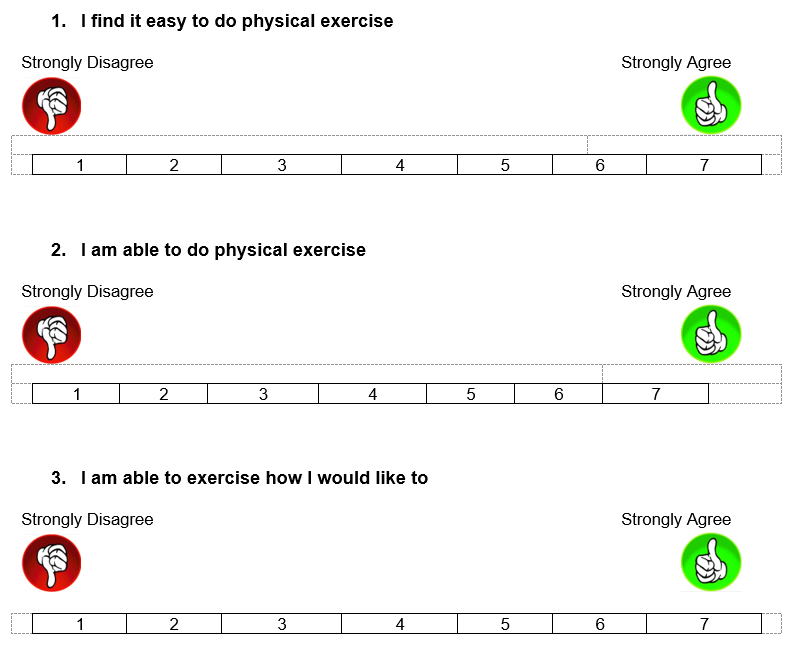
## Measurement of Outcomes

Choosing methods to measure outcomes is crucial to the success of an outcome evaluation. In this proposed evaluation plan, we will present methods which could be used by EAB to help measure proposed outcomes. In this partial evaluation plan, we will propose measurement methods for one short-term outcome, capability, and the main intermediate outcome of [physical activity / cycling]. These measurement methods are examples of what could be used within a complete evaluation plan to track changes in participants’ outcomes.

*Short-Term Outcomes*

One example of a validated measure that could be used to track outcomes of capability and motivation for physical activity is the UK Attitudes to Exercise Scale (UKAES) created by Taylor, et al (2016) measuring each component of the COM-B model in a group of obese children. The scale is designed for children as young as seven years old and consists of 20 items: four measuring capability, ten measuring motivation, and six measuring opportunity (another component of the COM-B model). For each item, children rate their agreement toward a presented statement using a 7-point Likert rating scale. The statements assess children’s general attitudes without focusing on a specific time frame. For example, one statement is “I find it easy to do physical exercise” and responses range from 1 (Strongly Disagree) to 7 (Strongly Agree). (See Figure IV.1)

**Figure IV.1: Items 1-3 from the UK Attitudes to Exercise Scale (Taylor, et al. 2016)**



To score the UKAES, responses for the questions corresponding to each component of the COM-B model (capability, opportunity, motivation) are averaged to give a numeric mean response score for each respondent.

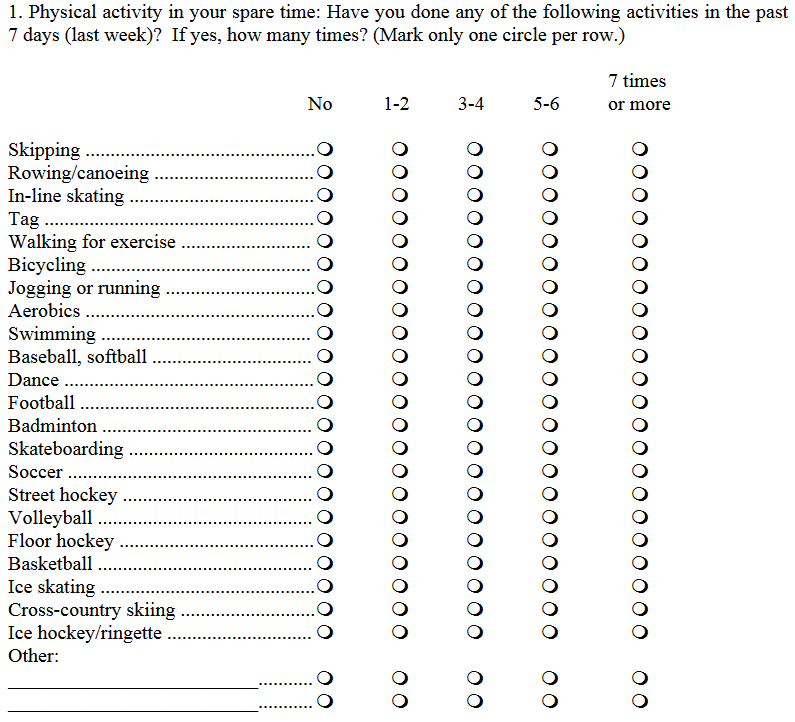
Taylor, et al (2016) performed a study to test the reliability of the scale in 116 children. 71 children (34 female, ages 9-17) were attendees of a weight-management camp in northern England, and 45 children (22 female, ages 9-13) were schoolchildren in northern Spain. To assess reliability, Cronbach's alpha was computed for each section of the scale, resulting in an alpha of 0.786 for capability and 0.796 for motivation. An alpha greater than 0.70 indicates there is adequate internal consistency in the scale items.

The same study also examined construct validity of the scale by using confirmatory factor analysis. The authors examined absolute fit, parsimony of fit, and comparative fit. Acceptable absolute fit was chosen as a χ2/degrees of freedom ratio <2.0, acceptable parsimony of fit was chosen as a root mean square error < 0.08, and acceptable comparative fit was chosen as a comparative fit index > 0.9. The author’s study resulted in a χ2/degrees of freedom ratio of 1.55, a root-mean-square error of 0.72, and a comparative fit index of 0.92. These results indicate that the scale is acceptable for measuring the construct of capability for exercise.

*Intermediate Outcome*

An example of a measure that could be adopted for the intermediate outcome ‘Physical Activity’ could be the Physical Activity Questionnaire for Children (PAQ-C). PAQ-C is a self-administered questionnaire designed to assess physical activity in older children ages 8-14 years (Kowalski et al 2004). The scale consists of 10 items. Each item is scored using a five-point Likert scale, 1 describing a low frequency and 5 describing a high frequency of occurrence. Each item measures physical activity with a seven-day recall, Item 1 measures spare time activity, 2- 8 cover PE, recess, lunch, after school, evening, weekends, and describes you best. Item 9 asks about a general physical activity rating for every day of the week, Item 10 covers anomalies that may hinder physical activity. An example of a question assessing item one is shown in figure IV.2 below.

**Figure IV.2: Item 1 from the PAQ-C (Kowalski, et al. 2004)**



To calculate an overall activity summary score, the mean of the values in items 1-9 is taken. A score of 1 denotes low physical activity and 5 denotes high physical activity (Kowalski et al 2004).

The reliability of the PAQ-C questionnaire was assessed by Benitez et al in a study conducted in a Spanish academic setting. The sample population consisted of children in 4th through 6th grade in primary schools in Malaga and Orense, Spain. Out of the 83 total participants, there were 46 boys and 37 girls. The PAQ-C survey was administered to participants twice during the same day and the Intraclass Correlation Coefficient (ICC) and the Cronbach alpha scores were taken to measure both reliability and reproducibility of the survey. The total ICC score for all the items was 0.96 and the total Cronbach alpha score was 0.76 both scores supporting the reliability of the questionnaire (Benitez et al 2016). Kowalski et al in a separate assessment examined the validity of the PAQ-C questionnaire using a sample of children aged 8-13. The study assessed convergent validity and construct validity of the questionnaire Convergent validity was determined by comparing PAQ-C responses to a physical activity rating, moderate to vigorous activity rating, and a teacher rating of physical activity. These comparisons produced correlation scores r= 0.63, r = 0.53, and r= 0.45 respectively providing support for convergent validity. Construct validity was done by comparing PAQ-C responses to an athletic competence scale. This second comparison produced a correlation score of r= 0.48 supporting the construct validity of the PAQ-C questionnaire (Kowalski et al 2004)

## Additional Comments

Each measurement tool chosen comes with advantages and limitations. For the short-term outcome of capability to cycle, a latent variable that cannot be directly measured, data collection is limited to using self-report questionnaires such as the one presented earlier. The advantages of the UKAES are that it is short, free, and easily scored. There is a need for personnel to administer, score, and interpret the results, but the administrative load is low. The main limitation of the UKAES is that it has currently not been used in many other scenarios besides the study described. In addition, some adaptations would be needed to make it more specific to bicycling to meet the needs of EAB. For the intermediate outcome of cycling, we suggested a general physical activity measurement tool (PAQ-C) which includes one option to designate cycling activity. This tool could be adapted to be more specific to EAB’s needs to measure cycling activity specifically. The benefits of using a self-report tool such as the PAQ-C are that it requires fewer resources to implement and score. One limitation is that a self-report is not as accurate at measuring true physical activity behavior compared to other methods. In addition, the PAQ-C also does not include estimates of the intensity of physical activity, nor does it allow for the calculation of true energy expenditure. However, the PAQ-C may match EAB’s needs because of how similar it is to the current follow-up survey measures, how it can be integrated into existing surveys given to participants, and how it has been shown to be reliable and valid in measuring physical activity of children enrolled in school.

## Section IV Conclusion

Measuring outcomes is a vital part of the program evaluation process. To determine if the activities within the Earn a Bike at school program have an impact on its short-term and intermediate outcomes, reliable and valid measurement tools are needed. Tools like the UKAES and the PAQ-C have proven useful in measuring key constructs like capability and motivation and overall change in physical activity which may be useful in determining the overall impact of the EAB program.

# CONCLUSION

Physical inactivity is high in children in the San Antonio area and has been linked to a multitude of adverse health consequences including obesity and diabetes (Friel et al 2020) (Piercy, 2018). The Earn a Bike @ school program aims to address these consequences. To ensure the effectiveness of this program, we developed an evaluation plan that can be adopted by EAB. In this plan, we provided a background, using evidence to show the severity of the health problem of physical inactivity. We defined the program in detail using a logic model to outline the program’s inputs, activities, outputs, and outcomes. We provided questions and methods that EAB can use to analyze the program’s coverage, delivery, and total process. We described possible study designs that could be adopted by the EAB for an effective outcome evaluation and recommended a two-group quasi-experimental study design based on cost-effectiveness for the EAB. Finally, we recommended that the EAB use the UK Attitudes to Exercise Scale (UKAES) and Physical Activity Questionnaire for Children (PAQ-C) as reliable measures for the short-term and intermediate outcomes proposed in our logic model.

Our evaluation plan has some limitations. One limitation to the plan is that it does not cover all the potential outcomes that can be targeted by the EAB at school program. Due to the nature of the assignment, we decided to leave out the part of the program that targeted increasing attendance and good behavior. In a future evaluation plan, this area can be further explored. Another limitation is that we were not able to get the input of most of the stakeholders and their separate goals might sway the content of the evaluation plan.

Despite the limitations, the detailed nature of the evaluation plan and its heavy scientific backing increases its likelihood of improving the overall effect of the program. Additionally, there are options that allow EAB to adapt if changes to the program occur. Incorporating methods such as those suggested can inform an effective evaluation of the Earn-a-Bike @ School program.

NOTE: We will copy our grader, Dr. Lara Savas, on the email we send with this evaluation plan to Earn-a-Bike.

# REFERENCES

Barlow, S. E., & and the, E. C. (2007). Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary report. *Pediatrics, 120*, S164-S192. doi:10.1542/peds.2007-2329C

Bell, J. A., Hamer, M., Batty, G. D., Singh-Manoux, A., Sabia, S., & Kivimaki, M. (2014). Combined effect of physical activity and leisure time sitting on long-term risk of incident obesity and metabolic risk factor clustering. *Diabetologia, 57*(10), 2048-2056. doi:10.1007/s00125-014-3323-8

Benítez-Porres, López-Fernández, I., Raya, J. F., Álvarez Carnero, S., Alvero-Cruz, J. R., & Álvarez Carnero, E. (2016). Reliability and Validity of the PAQ-C Questionnaire to Assess Physical Activity in Children. *The Journal of School Health*, *86*(9), 677–685. <https://doi.org/10.1111/josh.12418>

Bleich, S. N., Ku, R., & Wang, Y. C. (2011). Relative contribution of energy intake and energy expenditure to childhood obesity: A review of the literature and directions for future research. *International Journal of Obesity, 35*(1), 1-15. doi:10.1038/ijo.2010.252

Clark, M. H., & Middleton, S. C. (2010). Internal validity. In P. Peterson, E. Baker & B. McGaw (Eds.), *International encyclopedia of education (third edition)* (pp. 90-96). Oxford: Elsevier. doi: <https://doi.org/10.1016/B978-0-08-044894-7.00292-X>

Cleven, L., Krell-Roesch, J., Nigg, C. R., & Woll, A. (2020). The association between physical activity with incident obesity, coronary heart disease, diabetes and hypertension in adults: A systematic review of longitudinal studies published after 2012. *BMC Public Health, 20*(1), 726. doi:10.1186/s12889-020-08715-4

Di Cesare, M., Sorić, M., Bovet, P., Miranda, J. J., Bhutta, Z., Stevens, G. A., . . . Bentham, J. (2019). The epidemiological burden of obesity in childhood: A worldwide epidemic requiring urgent action. *BMC Medicine, 17*(1), 212. doi:10.1186/s12916-019-1449-8

Fallin, A., Parker, L., Lindgreen, J., Riker, C., Kercsmar, S., & Hahn, E. J. (2011). Secular trends and smoke-free policy development in rural Kentucky. *Health education research*, *26*(4), 711–721. <https://doi.org/10.1093/her/cyr032>

Flannelly, K.J., Laura T. Flannelly & Katherine R. B. Jankowski (2018) Threats to the Internal Validity of Experimental and Quasi-Experimental Research in Healthcare, Journal of Health Care Chaplaincy, 24:3, 107-130, DOI: [10.1080/08854726.2017.1421019](https://doi.org/10.1080/08854726.2017.1421019)

Foster, B. A., Maness, T. M., & Aquino, C. A. (2017). Trends and disparities in the prevalence of childhood obesity in south Texas between 2009 and 2015. *Journal of Obesity, 2017*, 1-7. doi:10.1155/2017/1424968

Friel, C. P., Duran, A. T., Shechter, A., & Diaz, K. M. (2020). U.S. children meeting physical activity, screen time, and sleep guidelines. *American Journal of Preventive Medicine, 59*(4), 513-521. doi:10.1016/j.amepre.2020.05.007

Gallagher, D. (2020). A guide to methods for assessing childhood obesity. *Childhood Obesity,* , 75.

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association : JAMIA*, *13*(1), 16–23. <https://doi.org/10.1197/jamia.M1749>

Howlett, N., Schulz, J., Trivedi, D., Troop, N., & Chater, A. (2019). A prospective study exploring the construct and predictive validity of the COM-B model for physical activity. *Journal of Health Psychology, 24*(10), 1378-1391. doi:10.1177/1359105317739098

Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: Global action for public health. *The Lancet*, *380*(9838), 294–305. <https://doi.org/10.1016/S0140-6736(12)60898-8>

Kowalski Kent, Crocker Peter R.E, Donen Rachel M. (2004). The Physical Activity Questionnaire for Older Children (PAQ-C) and Adolescents (PAQ-A) Manual. College of Kinesiology, University of Saskatchan. [Donen - CHAPTER 1 INTRODUCTION[5494].pdf](file:///C:\Users\chiso\AppData\Local\Packages\microsoft.windowscommunicationsapps_8wekyb3d8bbwe\LocalState\Files\S0\253\Attachments\Donen%20-%20CHAPTER%201%20INTRODUCTION%5b5494%5d.pdf)

Krekoukia, M., Nassis, G. P., Psarra, G., Skenderi, K., Chrousos, G. P., & Sidossis, L. S. (2007). Elevated total and central adiposity and low physical activity are associated with insulin resistance in children. *Metabolism, 56*(2), 206-213. doi:10.1016/j.metabol.2006.09.014

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

Piercy, K. L., Troiano, R. P., Ballard, R. M., Carlson, S. A., Fulton, J. E., Galuska, D. A., . . . Olson, R. D. (2018). The physical activity guidelines for Americans. *Jama, 320*(19), 2020. doi:10.1001/jama.2018.14854

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J.-P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6 (Suppl. 3)), S197–S239. <https://doi.org/10.1139/apnm-2015-0663>

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412961288

Saunders, Evans, M. H., & Joshi, P. (2005). Developing a Process-Evaluation Plan for Assessing Health Promotion Program Implementation: A How-To Guide. *Health Promotion Practice*, *6*(2), 134–147. <https://doi.org/10.1177/1524839904273387>

Scholtes, Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury*, *42*(3), 236–240. <https://doi.org/10.1016/j.injury.2010.11.042>

School Physical Activity and Nutrition (SPAN) Project. (2020). Child Obesity in Texas. Accessed Feb 26, 2022. <https://sph.uth.edu/research/centers/dell/resources/span%20update%20final.pdf>

School Physical Activity and Nutrition (SPAN) Project. (2021). Healthy Children, Healthy State: Physical Activity in Texas. Accessed Feb 26, 2022. <https://sph.uth.edu/research/centers/dell/texas-child-health-status-report/physical%20activity%202021.pdf>

Szeszulski, J., Lanza, K., Dooley, E. E., Johnson, A. M., Knell, G., Walker, T. J., Craig, D. W., Robertson, M. C., Salvo, D., & Kohl, H. W. (2021). Y-PATHS: A Conceptual Framework for Classifying the Timing, How, and Setting of Youth Physical Activity. *Journal of Physical Activity and Health*, *18*(3), 310–317. <https://doi.org/10.1123/jpah.2020-0603>

Taylor, M. J. (2016). Measuring perceived exercise capability and investigating its relationship with childhood obesity: A feasibility study. *International Journal of Obesity*, 5.

US Department of Health and Human Services. (2018). Physical activity guidelines for Americans, 2nd edition., 118.

Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: A systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556. <https://doi.org/10.1097/HCO.0000000000000437>

Welk, G., Morrow, J., & Saint-Maurice, P. (2017). Measures registry user guide: Individual physical activity. *National Collaborative on Childhood Obesity Research: Washington, DC, USA,*

Willmott, T. J., Pang, B., & Rundle-Thiele, S. (2021). Capability, opportunity, and motivation: An across contexts empirical examination of the COM-B model. *BMC Public Health, 21*(1), 1014. doi:10.1186/s12889-021-11019-w