Measures Registry User Guide: Physical Activity Environment

Jordan Carlson, PhD • Kelsey Dean, MS, RD, LD, CCRP • James Sallis, PhD
Suggested citation:

Measures Registry
User Guide:
Physical Activity
Environment
# Table of Contents

Acknowledgments .......................................................................................................................... ii  

**Section 1. Introduction** .................................................................................................................. 1  
Overview of the Physical Activity Environment ............................................................................. 1  
Measures Registry User Guide ......................................................................................................... 2  
Organization of this User Guide ...................................................................................................... 2  

**Section 2. Assessing and Defining Physical Activity Environments** .................................................. 4  
Why Study the Physical Activity Environment? .............................................................................. 5  
Physical Activity Can Be Supported in Diverse Environments ....................................................... 6  
Evidence on Environments and Physical Activity .............................................................................. 7  
Multiple Users Have Multiple Uses for Physical Activity Environment Measures ........................... 7  

**Section 3. Key Concepts in Physical Activity** .................................................................................. 8  
Environment Assessment ............................................................................................................... 8  
- Levels of Built Environment Attributes ......................................................................................... 9  
  - Macro-scale Attributes .................................................................................................................. 9  
  - Micro-scale Attributes .................................................................................................................. 9  
  - Social Attributes ......................................................................................................................... 9  
- Methods of Environmental Assessment ......................................................................................... 9  
  - GIS-based Measures ................................................................................................................... 10  
  - Observational Measures ........................................................................................................... 11  
- Questionnaires and Environment Perceptions ............................................................................... 11  
Proximity, Accessibility, and Quality of Environmental Attributes ................................................... 12  
Measurements Using Specific Environments and Physical Activity in Specific Environments .......... 12  
Considerations for Youth ................................................................................................................ 12  
  - Environmental Correlates Unique to Youth ............................................................................... 12  
  - Self-reporting in Youth ................................................................................................................ 12  

**Section 4. Evaluating Existing Measures** ....................................................................................... 14  
Reliability and Validity .................................................................................................................... 14  
  - GIS-based Measures ................................................................................................................... 15  
  - Observational Measures ........................................................................................................... 16  
  - Questionnaires .......................................................................................................................... 16  
- Single Items, Scales, and Indices .................................................................................................... 16  
  - GIS-based Measures ................................................................................................................... 16  
  - Observational Measures ........................................................................................................... 17  
  - Questionnaires .......................................................................................................................... 17  
- Other Measurement Issues ............................................................................................................ 17  
  - Response Scales ......................................................................................................................... 17  
  - Stability vs. Sensitivity to Changes ............................................................................................ 17  

**Section 5. Examples of Measures with Reliability and/or Validity Evidence** .................................... 18  
GIS-based Measures ....................................................................................................................... 19  
  - Community Design and Transportation System ....................................................................... 19  
  - Observational Measures ........................................................................................................... 19  
  - Streetscapes ............................................................................................................................... 19  

**Section 6. Selecting Measures** ...................................................................................................... 22  
Defining Project Purpose ................................................................................................................ 22  
Considerations in Selecting Measures ............................................................................................ 23  
  - Setting of interest ....................................................................................................................... 25  
  - Comprehensiveness ................................................................................................................... 25  
- Reliability and Validity ................................................................................................................ 25  
  - Relevance to the Population Being Studied ............................................................................. 25  
  - Resources and Expertise Required .............................................................................................. 26  
- Flexibility and Adaptability ......................................................................................................... 26  
Suggested Process for Using the Measures Registry ......................................................................... 27  
  - Searching and Filtering Results ................................................................................................. 27  
  - Navigating the Information Tabs Within Each Publication ....................................................... 27  

**Section 7. Collecting and Reporting Data** ..................................................................................... 28  
GIS-based Measurement Considerations .......................................................................................... 29  
  - Needed Expertise ....................................................................................................................... 29  
  - Access to Databases .................................................................................................................. 29  
  - Variable Computations ............................................................................................................. 29  
  - Buffer Size, Type, and Origin .................................................................................................... 29  
Observation Measurement Considerations ....................................................................................... 30  
  - Training ....................................................................................................................................... 30  
  - Selection of Environment Samples for Auditing ..................................................................... 30  
  - Use of Google Street View for Observational Audits ................................................................. 30  
  - Smartphone and Tablet Applications ........................................................................................... 31  
- Photovoice ..................................................................................................................................... 31  
  - Statistical Considerations .......................................................................................................... 31  
  - Nested Data ............................................................................................................................... 31  
  - Standardized and Unstandardized Regression Coefficients ...................................................... 31  
  - Multicollinearity ........................................................................................................................ 31  

**Section 8. Case Studies of Selecting Measures** .............................................................................. 32  
Case Study 1: Improving Streetscapes and Parks Around Schools ............................................... 34
NCCOR gratefully recognizes The JPB Foundation for funding the development of the Measures Registry User Guides. NCCOR also extends its sincere thanks to the NCCOR Steering Committee, NCCOR Members, and the following individuals for their contributions to the Guides.

Authors

Individual Diet
- Sharon Kirkpatrick, PhD, MHSc, RD
- Amanda Raffoul, MSc

Food Environment
- Leslie Lytle, PhD
- Allison Myers, PhD, MPH

Individual Physical Activity
- Gregory Welk, PhD
- James Morrow, PhD, FACSM, FNAX
- Pedro Saint-Maurice, PhD

Physical Activity Environment
- Jordan Carlson, PhD
- Kelsey Dean, MS, RD, LD, CCRP
- James Sallis, PhD

Reviewers

Food and Nutrition Expert Panel
- Alice Ammerman, DrPH, RD
- Carol Boushey, PhD, MPH, RD
- Karen Webb, PhD, MPH
- Gail Woodward-Lopez, MPH, RD

Physical Activity Expert Panel
- Genevieve Dunton, PhD, MPH
- Patty Freedson, PhD
- Brian Saelens, PhD

Measures Registry User Guide Project Team

NCCOR Members
- Rachel Ballard, MD, MPH
- David Berrigan, PhD, MPH (co-lead)
- Stephanie George, PhD, MPH
- Jill Reedy, PhD, MPH, RD (co-lead)

NCCOR Coordinating Center
- LaVerne Canady, MPA
- Adee Kennedy, MS, MPH
- Todd Phillips, MS
- Anne Rodgers
- Amanda Samuels, MS, MPH
- Namita Vaidya, MPH

We would also like to thank Seung Hee Lee-Kwan, PhD, MS; Latetia Freeman Moore, PhD, MSPH; Sohyun Park, PhD; Jesus Soares, ScD, MSc; and Tina Kauh, PhD, for reviewing the final draft of the Guides.
Introduction
Introduction

Measurement is a fundamental component of all forms of research and it is certainly true for research on childhood obesity. A top priority for the National Collaborative on Childhood Obesity Research (NCCOR) is to encourage consistent use of high-quality, comparable measures and research methods across childhood obesity prevention and research.

NCCOR’s Measures Registry—a free, online repository of articles about measures—helps achieve this aim. It is widely recognized as a key resource that gives researchers and practitioners access to detailed information on measures in one easy-to-search location. The Registry’s measures focus on four domains that can influence childhood obesity on a population level:

- Individual Diet
- Food Environment
- Individual Physical Activity
- Physical Activity Environment

Even with this resource, however, it can be challenging for users to choose the most appropriate measures for their work. To address this need, NCCOR began the Measures Registry User Guides project in 2015. Organized by the same four domains as the Measures Registry, the User Guides are designed to provide an overview of measurement, describe general principles of measurement selection, present case studies that walk users through the process of using the Measures Registry to select appropriate measures, and direct researchers and practitioners to additional resources and sources of useful information (Figure 1). The User Guides will help move the field forward by fostering more consistent use of measures, which will allow for standardization, meta-analyses, and synthesis.

Overview of the Physical Activity Environment Measures Registry User Guide

The overall goal of this User Guide is to help users of the NCCOR Measures Registry make informed decisions when selecting, processing, and interpreting measurement tools for physical activity environments. This Guide is not intended to be a comprehensive summary of all measures, an evaluation of the measures, or a compilation of research using the measures. Instead, it fills a different niche by emphasizing the measurement issues that should be considered when selecting and using physical activity environment measures for research and practice purposes. This User Guide offers examples and tips for navigating the Measures Registry.

Organization of this User Guide

This Guide provides an orientation to physical activity environment assessment methods, by setting, and discusses
considerations for selecting and using measures. It is designed to be useful for both researchers and practitioners, and includes case studies that show how both audiences can apply these considerations in practice.

In addition to this Introduction, this User Guide includes the following sections:

- **Section 2. Assessing and Defining Physical Activity Environments** provides a rationale for assessing physical activity environments and defines the key physical activity environment settings. Understanding the nature of these environments and the existing evidence will help users identify the most appropriate measures for the settings they wish to study.

- **Section 3. Key Concepts in Physical Activity Environment Assessment** describes the various methods of physical activity environment measurement across settings and other key concepts to consider when selecting environmental measures.

- **Section 4. Evaluating Existing Measures** provides an overview of the key measurement properties to consider when selecting environmental measures, including terminology, distinction between reliability and validity, single- vs. multi-item measures, response scales, and sensitivity to change.

- **Section 5. Examples of Measures with Reliability and/or Validity Evidence** provides examples of commonly used measurement tools with relatively extensive evidence of reliability and/or validity for assessing environments in various settings.

- **Section 6. Selecting Measures** outlines the process of selecting appropriate measurement tools for the given study population and research or evaluation aims. Additional considerations, such as resources required for data collection and analysis, are discussed, as are suggestions for using the Measures Registry.

- **Section 7. Collecting and Reporting Data** outlines methods and resources for successful and reliable data collection, including identifying local expertise, training staff, and deriving variables from the raw data.

- **Section 8. Case Studies of Selecting Measures** use hypothetical study designs to illustrate decision making about measures based on the information in this Guide. Both practice- and research-based examples are given, as well as selection considerations depending on the project purpose, study population, and intended audience of the study.

- **Section 9. Next Steps in Physical Activity Environment Assessment** highlights gaps in physical activity environment research and makes recommendations to facilitate continued advances in this field.

- **Section 10. Conclusion**

- **Section 11. Additional Resources on Physical Activity Environment Measurement** highlights additional resources that can supplement the information in this Guide.

---

**NCCOR: WORKING TOGETHER TO REVERSE CHILDHOOD OBESITY**

NCCOR is a partnership of the four leading funders of childhood obesity research: The Centers for Disease Control and Prevention (CDC), the National Institutes of Health (NIH), the Robert Wood Johnson Foundation (RWJF), and the U.S. Department of Agriculture (USDA). These four leaders joined forces in 2008 to continually assess the needs in childhood obesity research, develop joint projects to address gaps and make strategic advancements, and work together to generate fresh and synergetic ideas to reduce childhood obesity. For more information about NCCOR, visit [www.nccor.org](http://www.nccor.org).
Assessing and Defining Physical Activity Environments
Many factors influence whether or not a person engages in physical activity or meets the physical activity guideline of 150 minutes per week for adults and 60 minutes per day for youth. A vast body of research has shown that the built environment is a key determinant of physical activity, and the rationale and evidence behind this work are presented in this section.

Why Study the Physical Activity Environment?

The Institute of Medicine, Centers for Disease Control and Prevention, National Physical Activity Plan, American Heart Association, and World Health Organization all recommend interventions to change physical activity environments as a way to increase physical activity and reduce obesity in populations. The U.S. Surgeon General also emphasized the critical role of environments in the title and theme of the 2016 report Step it Up! The Surgeon General’s Call to Action on Walking and Walkable Communities. It is intuitive that some places are better than others for physical activity, but intensive study, mainly in the 21st century, has defined many environmental features that hold promise for increasing physical activity.

Many groups in communities across the United States and worldwide are working to create more activity-friendly environments. The goal of most of these efforts is to make it easier, safer, and more comfortable for people to choose to be physically active in their daily lives. It has become clear that different types of environments can be designed to facilitate different purposes, or domains, of physical activity. Well-designed parks, playgrounds, school grounds, trails, recreation centers, and health clubs can support leisure-time physical activity. Neighborhoods designed so people can walk to nearby destinations, along with safe facilities for pedestrians and bicyclists, can increase walking and bicycling for transportation. Buildings sited with access to public transit and multi-use trails, on-site walking paths, and designed with attractive stairs can encourage physical activity to, around, and in buildings. These environments are designed and managed by sectors of society outside of public health and medicine, including city planning, transportation, parks and recreation, architecture, landscape architecture, education, and private enterprise. Thus, researchers, practicing professionals, and policy makers in all of these sectors may be interested in measuring physical activity-related features of environments and be potential users of the physical activity environment measures in NCCOR’s Measures Registry.

Theoretical models that guide research and practice and that are widely used in the childhood obesity field include the built environment as an important factor in influencing behavior. In particular, ecological models of behavior (Figure 2) are based on the idea that behavior is influenced by a range of factors at the individual (psychology, biology), social and cultural (norms, values, interactions), environmental (social and built), and policy (laws, regulations, practices) levels. The most effective interventions are expected to be those that operate at multiple levels. For example, Safe Routes to School policies in transportation agencies can provide funding to improve sidewalks and street crossings near schools (environment level), support walking school buses and better enforcement of speed limits (social level), and deliver education programs to encourage walking and bicycling to school (individual level). A variety of measures are needed to evaluate the environmental, social, and individual changes produced by such multi-level interventions.
Physical Activity Can Be Supported in Diverse Environments

Multiple settings play an important role in overall physical activity, and many public health intervention recommendations are location-specific (e.g., school-based physical activity, home-based screen time, neighborhood walking). In youth, key physical activity settings include neighborhoods, parks and recreation areas, schools, and homes. Within each of these settings, the built environment can support or not support physical activity. Some environmental features have differential impacts on youth vs. adult physical activity, as highlighted in Section 3.

In general, neighborhood community design factors that support physical activity include greater residential density, a mix of residential and commercial land uses, and access to schools, parks, and recreation facilities. Neighborhood transportation system factors that support physical activity include street connectivity, cul-de-sacs (particularly as play areas for youth), bicycle infrastructure networks, and access to transit. Neighborhood streetscape features that support physical activity include safe street crossings, sidewalks, traffic speed bumps, and features that protect pedestrians from vehicles. Figure 3 contrasts commercial shopping areas that are designed around automobiles versus those designed with pedestrians in mind. Schools that support physical activity typically have safe outdoor play spaces as well as comprehensive school physical activity programs that include ample physical education, recess, classroom activity, and before- and after-school activity. Homes that support physical activity are those in which youth have ample access to play spaces and equipment and limited access to sedentary environments. Measures exist for assessing environments in each of these settings, though some environments do not have good coverage from existing measures, as outlined in Section 5.
Evidence on Environments and Physical Activity

The literature on social and built environments as they relate to physical activity has grown rapidly, made possible by a sustained output of new measures, as documented in the NCCOR Measures Registry. Most of the evidence about environments and physical activity is cross-sectional, and though the studies and their findings have many inconsistencies, some findings show good agreement. Among adults, walkability variables are supported as correlates of walking for transportation, and proximity to recreation facilities and neighborhood aesthetics are correlates of leisure-time physical activity. Total physical activity is related to recreation facilities, transportation facilities, and aesthetics. Similar associations have been found with children and adolescents, though less consistently. It is not possible to draw conclusions for older adults due to a paucity of studies. More prospective studies in adults and older adults are available now, also with inconsistent findings. Critiques of physical activity and built environment research note that self-selection—the idea that people who like to walk choose residential areas that are walkable, for example—limits evidence of causality. However, several studies have assessed self-selection and neighborhood preferences and concluded that associations between built environment and physical activity are sometimes attenuated but remain significant after accounting for self-selection. The large and long-term RESIDE study of people who relocated residences in Australia showed that changes in neighborhood environments were related to changes in physical activity (e.g., Hooper, 2014; Knuiman, 2014). Studies evaluating “natural experiments,” such as park improvements, bicycle facilities, and Safe Routes to School interventions, have shown promising results that improve evidence of causality.

Recently, interest has been growing in how built environments are related to sedentary behavior, including both total sitting time and more specific measures, such as sitting at work and screen time at home. The most consistent environmental correlate of screen time among children is having a television in the bedroom. A few studies of neighborhood walkability and sitting have been conducted, but results have been inconsistent.

Multiple Users Have Multiple Uses for Physical Activity Environment Measures

Researchers can use environmental measures to advance knowledge about environmental correlates (cross-sectional studies) and determinants (prospective studies) to explore equity in access to physical activity opportunities across socioeconomic classes and geographic areas, to document and evaluate environmental changes in intervention studies, and to evaluate the effect of policies designed to change built environments, such as Complete Streets, Vision Zero, and Safe Routes to School. Practitioners in many fields can use environmental measures to assess baseline conditions as part of planning processes and to evaluate their interventions. Policy makers can use environmental data to determine the need for community improvements, to target interventions and investments where they are needed most, and to assess the outcomes. Community groups can use environmental measures to provide data to support their advocacy for community enhancements. Information regarding how to select appropriate physical activity environment measures for each of these uses is presented in Section 6.
Key Concepts in Physical Activity Environment Assessment
The built environment can be measured at different levels, using different methods, and in various settings relevant to physical activity. Descriptions of these levels, methods, settings, and other key concepts are presented in this section. The main focus is on measures of environmental characteristics themselves. Related measures of how often a specific setting (e.g., park) is used, and measures of physical activity in a specific setting, such as a park, are discussed later in this section. Some measurement tools capture both environment characteristics and physical activity in the setting.

Levels of Built Environment Attributes

Physical activity environments are complex and can be measured at multiple levels that vary somewhat by the specific settings.

**Macro-scale Attributes**

"Macro-scale" attributes describe the layout of communities and include the concept of walkability. Walkability refers to the combination of mixed land use, moderate-to-high residential density, and high street connectivity that allows people to walk from home to common destinations, such as shops and schools. Proximity of parks, trails, bicycle networks (e.g., lanes, trails), and public transit can be considered macro-level attributes as well. Geographic information systems (GIS) are often used to capture these attributes when data are available; self-report measures also can assess these variables.

**Micro-scale Attributes**

"Micro-scale" attributes indicate the design of a setting and can affect the experience of being active or inactive in a given place. Streetscape attributes include presence, quality, and amenities of sidewalks; characteristics of street crossings, such as crosswalks and pedestrian signals; transportation variables, such as transit stops and signage, travel lanes, and speed limits; and aesthetic features of buildings and landscaping. Other micro-scale variables include activity zones and equipment for physical activity in parks and school yards; design of public spaces; characteristics of trail design; and attributes of buildings, including stairs. Micro-scale features are usually measured by direct observation (i.e., audits) and self-report, because they are not often available in GIS.

**Social Attributes**

A third level of measurement is mostly focused on social dimensions of the environment. Social environments are not well-defined, but commonly studied variables include people in the environment and evidence of people’s behavior. Examples are indicators of social disorder such as graffiti, litter, and abandoned buildings; people walking or children playing; maintenance of park or playground equipment; crime; traffic density and speed; and programming and supervision in recreation areas. Social indicators are usually measured by direct observation or self-report.

Methods of Environmental Assessment

Environmental assessment tools fall into three categories based on their methodology and data collection procedures (see Table 1). Within the Measures Registry, specific measures falling within each of these methods of assessment can be isolated using the check boxes under...
the “Measure Type” subheading. The Measurement Types are:

- **GIS**: GIS-based measures, which involve archival data sets that are layered and analyzed within GIS software.

- **Environmental Observation**: Observational measures, which are obtained using systematic observational methods (audits).

- **Questionnaire**: Self- or proxy-reports of environmental features, which can be administered by paper-and-pencil questionnaire, computer or online survey, or interview.

Additionally, the Measures Registry includes Electronic Monitor, Record or Log, and Other as modalities of environmental measures. The Electronic Monitor section includes studies that used Global Positioning Systems (GPS) to assess participants’ locations. The Record or Log and Other sections include some observational audit measures as well as direct observation measures of physical activity within specific settings (e.g., System for Observing Play and Recreation in Communities [SOPARC]). These measurement types are not covered in detail here because they are described in the Measures Registry User Guide: Individual Physical Activity.

### GIS-based Measures

GIS is a software system (rather than a measure itself) used for integrating and analyzing spatial and geographic data. GIS-based environmental measures are typically derived from existing databases, and these data sources are continuously evolving. Data collected through observations and questionnaires also can be integrated into GIS if they are spatially specific. Data in GIS need to have spatial reference, such as parcel or road database shapefiles. Existing databases are available from multiple sources, such as regional transportation planning agencies, local municipalities, and ongoing surveillance surveys (e.g., American Community Survey, U.S. Census). From this existing information, built environment measures that could be related to physical activity are computed. For example, road network spatial databases (i.e., shapefiles) can be obtained from many regional transportation planning agencies or local municipalities and used to calculate the number of street intersections within a defined neighborhood or area. The GIS measures in the Measures Registry are simply variable computations performed in GIS that have been used in physical activity research. Many of the publications listed in the “Results” view of the Measures Registry include multiple GIS-derived variables (see Section 6 for suggested process for using the Measures Registry).

Advantages of GIS-based measures are that they rely on (relatively) objective data and are easy to derive for large samples by someone with GIS software expertise. Disadvantages of GIS-based measures are that they are limited to community design (land use) and transportation system information. Data often do not exist for home, school, streetscape, and micro-level (i.e., interior) parks and recreation features. However, data collected in these settings can be integrated into GIS. Using GIS-based measures can be difficult. GIS-based projects often require collaboration with geographers or transportation or public affairs experts. Although GIS-based data on traffic, crime, sidewalks, and crossings exist in some areas, data can vary in quality, completeness, recency, and availability across jurisdictions. The way that data, such as crime data, are recorded across jurisdictions also can vary, making it difficult to compare and pool data across geographic areas. The quality of available GIS data should not be assumed.
Observational Measures

Observational measures, or audit tools, involve systematic observation of an environment, typically by trained research staff or community stakeholders. Good audits tools are accompanied by a manual with well-defined construct definitions and observer training procedures that result in acceptable inter-rater reliability. Audits exist for assessing micro-level streetscape, home, school, park, and recreation center features. In neighborhoods, audits can provide useful information that is not captured in GIS databases, such as availability of crosswalks and duration of pedestrian crossing signals. Audits can be favored over self-reported measures because the former provide more objective, detailed, and place-specific data than self-reported participant or informant perceptions.

Observational measures require substantial time and resources because trained observers must visit each environment being assessed. Although evidence is accumulating to support the validity of conducting observations remotely, such as through Google Earth and Google Streetview, these methods still require significant time and resources and extensive training of coders. Another limitation of audits is that they have poorer validity for capturing transient and subjective constructs, such as aesthetics or social disorder, because they are typically only conducted at one point in time. Audit measures must be processed and scored, which is a burden on users, and some measures do not have well-developed scoring procedures.

Audit tools serve several important purposes outside of research. They can be used to engage and educate community members about the impacts of environments on health. These tools are also useful for conducting needs assessments to inform local decision making and advocacy. Audit tools can point to specific environmental changes to target, including more modifiable micro-level features (e.g., sidewalk quality) than can be identified through GIS-based measures, which capture constructs that can take decades to change (e.g., street layout). For example, audits are often used by community advocacy groups to identify needed sidewalk and street-crossing improvements.

Questionnaires and Environment Perceptions

Questionnaires are commonly used in physical activity environment assessment and include participant-report and proxy-report tools. Proxy-reports involve the questionnaire being completed by someone other than the participant themselves, such as a parent, key informant, or school staff member. Most questionnaires in the Measures Registry ask participants to report on objective facts, such as presence of sidewalks, proximity to parks, and physical activity equipment in parks, schools, or homes. Objective indicators of aesthetics can even be reported, such as amount of greenery, views, and how well buildings are maintained. Some questionnaires ask participants to evaluate the quality of environments. Examples include perceived safety from traffic, attractiveness of parks, and comfort level while walking in the neighborhood. Self-report questionnaires are the only way to assess perceptive evaluations of the environment, which often do not align with objective measures of the environment.

Questionnaires range in length, with some using single items and others using multi-item scales to capture various constructs. They frequently use Likert response scales, such as 4- or 5-point scales, with response options ranging from “strongly disagree” to “strongly agree.” Tools that capture the presence or absence of environmental attributes often use tallies or checklists of available resources. Tools that capture value-based perceptions of environmental attributes, such as safety and aesthetics, typically use Likert response scales (e.g., agree vs. disagree). The Neighborhood Environment Walkability Scale (NEWS) is an example of a questionnaire whose scales include reporting of both objective features and evaluative perceptions, whereas the sedentary environment module of the Physical and Nutritional Home Environment Inventory is an example of a tally/checklist that can be completed with a good level of objectivity.

Questionnaires are commonly used because meaningful environmental information can be captured with few resources and little burden to the researcher or practitioner. Questionnaires are useful for capturing information that is difficult to assess when GIS and audit data do not exist or are difficult to capture. Evidence from perceived environment measures has shown that perceptions often have small associations with objective measures, yet the former are important for understanding physical activity over and above objective environment characteristics. Thus, many studies find it beneficial to include both perceived and objective environment measures. Limitations to questionnaires include participant burden, the potential for...
low response rates, and floor or ceiling effects. Floor or ceiling effects can occur when the majority of participants respond at the bottom or top of the response scale and thus result in limited variability and power for detecting associations.

**Proximity, Accessibility, and Quality of Environmental Attributes**

Distinguishing between measures of proximity, accessibility, and quality is important when assessing environments. Proximity refers to the presence, absence, or distance to an environmental feature. For example, having a greater number of parks in the neighborhood and a shorter distance to the nearest park have been associated positively with physical activity in youth.\(^{19,56}\) Accessibility refers to the ease of getting to the feature. For example, a park could be nearby, but a highway or dangerous intersection between the neighborhood and the park reduces accessibility. Accessibility can be affected by the street network, with a connected grid allowing direct access and a disconnected suburban-style network creating a long, indirect pathway. Quality involves rating existing features on attributes such as aesthetics, appeal, condition, ease of use, or safety. The Environmental Assessment of Public Recreation Spaces (EAPRS) tool, for example, can be used to capture ease of access, cleanliness, colorfulness, condition, and comfort of park facilities.\(^{57}\) Similarly, the school audit tool by Jones et al. covers quality attributes such as aesthetics, design, maintenance, and suitability for physical activity of the external grounds of the school.\(^{58}\) GIS-based measures are mainly used to assess proximity through the mixture of land uses in community design and accessibility through the connectivity of the transportation system. Many audit tools and questionnaires capture proximity, accessibility, and quality.

**Measurements Using Specific Environments and Physical Activity in Specific Environments**

Researchers and practitioners sometimes desire to assess people’s use of specific environments (e.g., whether or how often a park is used), or physical activity levels within specific settings (e.g., how much activity occurs in parks). This information can be used to better understand where physical activity occurs, how exposure to different environments may influence physical activity, or how characteristics of a setting are related to physical activity in that setting. Although measures of use of specific environments and physical activity in specific environments are not directly part of environmental assessment, they are sometimes assessed in conjunction with environmental variables. Specific measures that fall under this category are presented at the end of Section 5, and examples include questionnaires, observational measures, and GPS.

**Considerations for Youth**

The Measures Registry includes measures relevant for both youth and adults. The “Age” subheading can be used to narrow the number of measures in the “Results” view based on age group. The following factors should be considered when selecting a measure that is specifically relevant for youth.

**Environmental Correlates Unique to Youth**

Youth-specific measures and evidence should be considered when youth are the focus of study. Some evidence suggests that the community design and transportation environments sometimes have differential associations with physical activity in youth versus adults. For example, greater intersection density and street connectivity have been fairly consistently positively related to walking for transport purposes in adults and youth.\(^{19,59,60}\) However, presence of cul-de-sacs, a street design feature indicative of lower street connectivity, has been positively related to recreational physical activity in youth, likely because these features serve as important low-traffic play areas for youth.\(^{19,27,61-64}\) Recreation environments also may be more important for understanding physical activity in youth versus adults.\(^{19,27,61,65}\) School environments are clearly more important for youth than for adult physical activity.\(^{65}\) Less is known about differences in the importance of environmental attributes in other settings between youth and adults, but in addition to consideration of study purpose, the study population should be kept in mind when selecting appropriate measures.

**Self-reporting in Youth**

Although self-reports have seen vast use in adults, they are used less commonly in children. This is because young children have difficulty with comprehension and/or may not have sufficient knowledge of their environments. No specific age cutpoint has been established for when self-reports become appropriate, but self-reports of young adolescents can produce comparable results to parent reports.\(^{66}\) When assessing environments in younger children, questionnaires can be completed by parents or other proxy reporters.
4

Evaluating Existing Measures
A tool’s measurement properties should always be considered when evaluating the acceptability of a measure. The most important properties to consider, namely the tool’s reliability and validity, as well as other relevant measurement properties, are described in this section. Each measure and publication in the Measures Registry includes a tab that provides detailed information on the measure’s reliability and validity, when available.

**Reliability and Validity**

Reliability refers to the consistency with which something is measured, whereas validity refers to the accuracy with which the measure captures the construct to be measured. Validity can be further broken down into criterion validity, which is when the tool is compared to a gold standard measure of the same construct (i.e., truth), and construct validity, which is when the tool is positively correlated with a measure of a theoretically-related construct. For example, we expect neighborhood walkability to be positively correlated with physical activity, especially active transportation. The types of reliability and validity that are relevant to built environment assessment differ across the methods of assessment. Acceptable measures would ideally have evidence of both reliability and validity, although in environmental research it can be difficult (e.g., with GIS) or unnecessary (e.g., with direct observation and environmental perceptions) to assess criterion validity because of the lack of gold standard comparison measures. Thus, many environmental measures rely on evidence of reliability (i.e., do raters agree with each other and is the same respondent consistent over time), face validity (i.e., is the measure perceived by experts as consistent with the concept it is intended to measure), and construct validity (e.g., is the measure related as expected to physical activity) to evaluate their quality or utility.

**GIS-based Measures**

A primary component comprising reliable GIS-based measures is that variable computations should be replicable by multiple analysts (see Section 7 for more detail on variable computations). Few existing measures provide this evidence, but this type of inter-rater reliability can be maximized by using well-defined variables and a detailed scoring protocol. Another key component of reliability of GIS-based measures is temporal match, which refers to the match between the time period when the GIS data were collected and the time period when other variables (e.g., participant physical activity) were collected. Ideally, the GIS and participant variables would be collected within one year of one another. Although GIS variables can be stable over multiple years, some areas can change rapidly, such as those undergoing redevelopment, so local knowledge of the area is useful. A poor temporal match would make GIS variables that change rapidly less useful for explaining physical activity. Criterion validity in GIS-based measures is primarily affected by the completeness and accuracy of the GIS databases used. Unfortunately, it is often impossible to know whether errors exist in public geodatabases, and evidence is lacking on how data incompleteness and inaccuracies affect physical activity research. It is important to investigate the quality of the data source when possible; for example, by directly observing a small sample of GIS variables to determine accuracy. Construct validity is important and relevant to GIS-based measures, with the key consideration being that variables used should have evidence or rationale for associations with physical activity. Evidence of construct validity can be found in reviews of built environment and physical activity research.
**Observational Measures**

Inter-rater reliability is the most commonly assessed measurement property of observational measures and is a critical component in determining a tool’s quality. Inter-rater reliability involves multiple raters completing the audit tool for the same locations independently and comparing their responses for discrepancies. This is typically done for a sample of at least 30 to 40 instances of the environment being captured (e.g., street segments). Key metrics to consider include percent absolute agreement, Kappa for use with yes/no checklist audits and categorical data, and intraclass correlation coefficients (ICCs) for use with continuous responses (e.g., 1-5 Likert scales). Commonly used thresholds to represent good percent agreement are ≥80 percent, and for Kappas and ICCs ≥0.80; Kappa and ICC values between 0.60 and 0.80 are often considered acceptable. It is important to note that a measure cannot have acceptable validity if it does not have acceptable inter-rater reliability.

Criterion validity testing of observational measures is not typically necessary because direct observation is considered a gold standard objective method. Construct validity is commonly assessed for observational audits, and tools or constructs that have shown associations with participants’ physical activity are interpreted to have good construct validity.

**Questionnaires**

The primary measurement property needed to support questionnaires is test-retest reliability, which involves administering the tool to the same participants at two time points, such as two weeks apart. Similar to inter-rater reliability, percent agreement, Kappa statistics, and ICCs are used to interpret reliability.

Criterion validity can be assessed for presence or absence and tally-based report measures by comparing a participant’s responses to the same tool completed by a researcher using an observational audit, but this is not commonly done. Criterion validity is not typically assessed for perceptive evaluation report measures because evaluations are subjective and do not have a gold standard comparison. Similar to observational measures, construct validity should be established by testing associations with physical activity. Users also should note whether the measurement properties were established in a similar population as the user intends to study. When using a questionnaire in a new population or comparing across populations, the measure would ideally have evidence of invariance across subgroups (e.g., the measure performs similarly in men and women).

**Single Items, Scales, and Indices**

Measure developers have used a variety of methods to reduce measures with many items to a small number of scales and indices. However, there is a trade-off between greater feasibility but lower reliability and validity of shorter scales or single-item indicators. Both indices and scales can be used to reduce a large number of variables to a small number of useful metrics. Although the terms “scale” and “index” are sometimes used interchangeably, they have differences. Scales comprise inter-related items capturing a narrow (usually unobservable) construct, typically a perception or attitude. One advantage of scales is that they often improve reliability over single items. In contrast, items comprising an index do not need to be inter-related and often capture a broad concept. Environmental measures typically assess a wide range of features, but consensus is growing that no single feature is the most important for physical activity. One advantage of indices is that they can be used to investigate additive effects of single items or features, particularly when composed of a sum of dichotomous items. A growing number of studies show that multi-item indices are most strongly related to physical activity outcomes.

**GIS-based Measures**

Walkability indices, which involve a composite score derived from multiple environmental attributes, are commonly used in GIS-based measures. Such indices can be computed by summing dichotomous (e.g., yes = 1, no = 0) variables across a number of environmental attributes. For example, a neighborhood that has connected streets (1), access to shops and restaurants (1), and low residential density (0) would have a score of 2 on an index ranging from 0-3. When creating indices from continuous variables representing different units of measurement, a common procedure is to transform each indicator or attribute score into a standardized z-score based on the sample mean and then take a sum or average across the z-scores to derive the index (e.g., Frank’s walkability index). These indices are advantageous because they can reduce a large number of variables to a manageable number of metrics and represent the combination of attributes sometimes viewed as more important than any single attribute. Another reason to use scales and indices is that environmental variables often are correlated with each other. Because components of walkability can be inter-correlated, it is sometimes not possible to include all components in the same analytic model. Running models with one walkability component at a time will underestimate associations with
outcomes, but the multi-component index should yield a more accurate estimate of the association of the pattern of the built environment with the outcome. Limitations of walkability indices are that one variable may be unknowingly driving an association, and the metrics are difficult to interpret outside of the area or region used to generate the standardized scores. Thus, single variable attributes with easily interpretable metrics (e.g., number of intersections per square km, used to capture street connectivity) are sometimes more desirable for planning and decision making than are indices.

Observational Measures
Indices lend themselves well for use in observational assessment, given that many features are rated dichotomously (e.g., presence or absence), but indices are unfortunately underused in this area. The Microscale Audit of Pedestrian Streetscapes (MAPS) is an example of an observational tool that uses indices. Indices can be used in observational measures in the same manner as in GIS-based measures and should be considered by users.

Questionnaires
Many questionnaires include multiple scales to group items according to specific attributes. For example, a perceived safety from traffic scale may include items covering perceptions of sidewalks, street crossings, traffic volume, and traffic speed, each rated on a Likert response scale. This type of data reduction simplifies analyses and can improve reliability properties of the measure. A disadvantage of scales is that multiple items are needed to assess each construct, so measures assessing multiple constructs can become lengthy. Some tools, such as the Physical Activity Neighborhood Environment Scale (PANES), use one or two items to assess a construct and cover multiple constructs. Such tools can have acceptable measurement properties and should be considered when survey length is a concern, but in general, reliability properties are most favorable when multiple items are used to assess each construct.

Other Measurement Issues

Response Scales
Response scales are used in both questionnaires and observational measures. Response scales in environmental assessment tools typically include anywhere from 2 (e.g., yes/no) to 10 or even more response options. Likert-type response scales can include three or more response options, with all or some options including anchor points, such as “none,” “some,” and “many” or “strongly agree” and “strongly disagree.” Anchors need to be balanced so that the intervals between every two sequential numbers are roughly equivalent. Response scales that include an odd number of response options often include a “neutral” or middle category, which can be problematic if responses will later be dichotomized (e.g., agree vs. disagree). An advantage of continuous response scales is that they typically result in greater variability and thus provide more power for detecting associations than dichotomous response scales. However, dichotomous response scales such as yes/no and agree/disagree are more easily interpretable and sometimes sufficient.

Stability vs. Sensitivity to Change
Measures that have good inter-rater or test-retest reliability focus on environmental attributes that are stable and show little change over short time periods. Although these traits are beneficial for establishing reliable measures, such tools may have limited use for assessing changes over time, such as evaluating interventions. GIS-based measures, in particular, have low sensitivity to change or have not yet been evaluated for sensitivity to change because macro-level community design features can require several years or even decades to change. For example, changes to increase population density or mixed land use within a neighborhood dominated by single-family homes would require significant policy changes and extensive redevelopment.

Although most observational measures and reports are generally not designed specifically to capture constructs that change over time, some may be useful for assessing changes over somewhat longer periods of time (e.g., ≥1 year) or when interventions target the constructs being assessed. For example, a cross-walk audit tool could be useful for capturing changes if an intervention specifically targets cross-walk improvements. If the user’s goal is to assess changes over time, special consideration is needed when identifying appropriate measures. Development of a new measure or new items for an existing measure may be needed to evaluate changes in the specific variables targeted by interventions. In some situations, it may be beneficial to use mixed-methods (i.e., both qualitative and quantitative data) to capture environmental changes. For example, a key informant with knowledge of the area and/or environmental changes could be interviewed to provide additional information on the environmental changes and the process by which they were achieved.
Examples of Measures with Reliability and/or Validity Evidence
GIS-based Measures

Community Design and Transportation System
GIS-based community design variables with evidence of construct validity include residential density, land use mix, and access to parks and recreation facilities. Transportation variables include street connectivity and access to transit. Walkability indices with evidence of construct validity include those by Cervero and Kockelman,74 Ewing et al.,75 and Frank et al.76 The latter index involves a summation of z-score variables, whereas the other two indices are computed using principal components analysis. Brownson et al. provide a detailed list of GIS-based variables and associated data sources.68 Walk Score is an example of an index with evidence of validity, but it is unique in that walkability scores are provided on a publicly available website free of charge. The user simply types in an address, and a Walk Score ranging from zero to 100 is provided. The Walk Score takes into account distance to amenities such as schools, parks, businesses, and other common destinations, as well as measures of street connectivity and residential density.77 The Walk Score creators have made an effort to align their algorithm with academic research, and several studies have found that Walk Score correlates well with measures of street connectivity, residential density, density of retail destinations, and access to other destinations.78,80 However, limited evidence exists regarding associations between Walk Score and physical activity, and validity of other metrics provided by Walk Score, such as Bike Score, Transit Score, and Crime Grade.

Observational Measures

Streetscapes
Streetscape audit tools with evidence of reliability and validity include the Microscale Audit of Pedestrian Streetscapes (MAPS),81 Systematic Pedestrian and Cycling Environment Scan (SPACES),82 Irvine-Minnesota Inventory (IMI),83 and Pedestrian Environment Data Scan (PEDS).84 MAPS includes a comprehensive scoring system, a limitation of other tools,81 as well as a short version designed for practitioners.85 Some streetscape audit tools were designed for specific settings, such as around schools (e.g., Texas Childhood Obesity Prevention Policy Evaluation School Environmental Audit Tool),86 and worksites (e.g., Workplace Walkability Audit),87 and in street alleys (e.g., Systematic Pedestrian and Cycling Environmental Scan for Alleys [SPACES for alleys]).88

Trails, Parks, and Recreation
The Path Environment Audit Tool (PEAT) has evidence of reliability and validity for assessing trails.89 Park assessment tools with evidence of reliability and validity include the Environmental Assessment of Public Recreation Spaces (EAPRS)87 and Community Park Audit Tool (CPAT).90 Both are comprehensive, but CPAT may be more feasible for use by practitioners. The Physical Activity Resource Assessment (PARA) can be used to assess trails, parks, and recreation areas, and has the advantage of being brief and easy to complete.91 A more comprehensive tool for assessing recreation environments is the Recreation Facility Audit Tool (REFAT), which was adapted from the PARA and has evidence of construct validity.92
Table 1: Examples of Commonly Used Built Environment Measures with Evidence of Reliability and/or Validity

<table>
<thead>
<tr>
<th>Setting</th>
<th>Method of Assessment</th>
<th>Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community design</td>
<td>Indices developed by Cervero and Rockman, Ewing et al., Frank et al., and Walk Score</td>
<td>Neighborhood Environment, Walkability Scale (NEWS)</td>
</tr>
<tr>
<td>Transportation system</td>
<td></td>
<td>Neighborhood Physical Activity Questionnaire (NPAC)</td>
</tr>
<tr>
<td>Streetscapes</td>
<td>(Microscale features are not often available in public databases, few GIS measures have been developed. However, data from observational measures can be entered into GIS. The same issues apply to the settings below)</td>
<td>Physical Activity Neighborhood Environment, Scale (PANES)</td>
</tr>
<tr>
<td>Parks</td>
<td>Community Park Audit Tool (CPAT)</td>
<td>Research on Urban Trail Environments (ROUTES)*</td>
</tr>
<tr>
<td>Recreation</td>
<td>Environmental Assessment of Public, Recreation Spaces (EAPRS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Path Environment Audit Tool (PEAT)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physical Activity Resource Assessment (PARA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recreation Facility Audit Tool (RFAT)</td>
<td></td>
</tr>
<tr>
<td>Schools and child care</td>
<td>Environment and Policy Assessment and Observation (EPAO)</td>
<td>School Physical Activity Nutrition and Physical Activity Self-Assessment for Child Care (NAP SACC)</td>
</tr>
<tr>
<td></td>
<td>Audit Tool for Primary School Environments</td>
<td>School Physical Activity Policy Assessment (S-PAPA)*</td>
</tr>
<tr>
<td>Homes</td>
<td>Home-Inventory Describing Eating and Activity Development (Home IDEA) for Preschoolers</td>
<td>Mapping Home and Neighborhood Environments</td>
</tr>
<tr>
<td></td>
<td>Healthy Homes Survey</td>
<td></td>
</tr>
<tr>
<td>Workplaces</td>
<td>Audit of Physical Activity Resources for Seniors (APARS)</td>
<td>Office Environment and Sitting Scale</td>
</tr>
<tr>
<td>Other buildings</td>
<td>Checklist of Health Promotion Environments at Worksites (CHEW)</td>
<td>Worksite and Energy Balance Survey (WEBS)</td>
</tr>
<tr>
<td></td>
<td>Environmental Assessment Tool (EAT)</td>
<td>Worksite Supportive Environments for Active Living Survey (SEALS)*</td>
</tr>
<tr>
<td>Rural</td>
<td>Rural Active Living Assessment (RALA)</td>
<td>Rural Active Living Perceived Environmental Support Scale (RALPES)</td>
</tr>
</tbody>
</table>

**Schools and Child Care**
A comprehensive school environment audit tool with evidence of reliability and validity was developed by Jones et al.58 Audit tools also exist for the childcare environment, such as the Environment and Policy Assessment and Observation (EPAO).93

**Homes**
Home audit tools with evidence of reliability include the Home Food and Activity Assessment94 and Healthy Homes Survey.95 The latter is unique because it can be conducted through a telephone survey, and involves a parent doing an audit of his or her home.

**Workplaces and Other Buildings**
Physical activity environment assessment tools for worksites have been reviewed by Hipp et al.96 The Checklist of Health Promotion Environments at Worksites (CHEW)97 and Environmental Assessment Tool (EAT)98 are comprehensive tools for assessing worksites that have evidence for reliability and validity. Evidence of reliability and validity also exists for tools that assess specific building types, such as congregate living facilities for older adults (Audit of Physical Activity Resources for Seniors [APARS]).99

**Rural**
The Rural Active Living Assessment (RALA) includes a street segment audit tool that has evidence of reliability and validity.100
Questionnaires

Community Design, Transportation System, and Streetscapes

Several report tools are commonly used and have evidence of reliability and validity for assessing neighborhood walkability factors. These include the Neighborhood Environment Walkability Scale (NEWS), an abbreviated NEWS, St. Louis Environmental Instrument, and Neighborhood Physical Activity Questionnaire (NPAQ). The Physical Activity Neighborhood Environment Scale (PANES) has acceptable reliability and validity properties but differs from the aforementioned tools because it is brief (17 items).

Trails, Parks, and Recreation

Trails, parks, and recreation environments are typically assessed using audit tools. Slater et al. developed an adolescent questionnaire for assessing park environments. For trails, the Research on Urban Trail Environments (ROUTES) trail use report tool includes items assessing trail environment characteristics.

Schools

The ActiveWhere? study developed and evaluated reliability and validity for a set of questions on the availability of several physical activity-related school facilities. The School Physical Activity Policy Assessment (S-PAPA) is a detailed report tool with reliability and validity that is primarily used for assessing school physical activity practices and policies, but also includes a small number of questions on the physical environment. For assessing child care environments, the Nutrition and Physical Activity Self-Assessment for Child Care (NAP SACC) has evidence of reliability and validity.

Homes

Report tools developed by Hume et al. and Rosenberg et al. have evidence of reliability and validity for assessing home environments related to youth physical activity and sedentary behaviors.

Workplaces and Other Buildings

The Worksite and Energy Balance Survey (WEBS) includes 72 items on physical activity as well as nutrition environments in the worksite and has evidence for reliability and validity. A briefer tool with evidence of reliability and validity is the Worksite Supportive Environments for Active Living Survey (SEALS). Also with evidence of reliability and validity, the Office Environment and Sitting Scale specifically assesses the sitting environment in the worksite.

Rural

The Rural Active Living Assessment (RALA) includes interview questions that are completed by community members or key informants. The questions cover availability of recreation facilities, programs, and policies, and have evidence of reliability and validity. Also with evidence of reliability and validity, the Rural Active Living Perceived Environmental Support Scale (RALPESS) can be used to assess participants’ perceptions of their environment. The tool covers: (1) church facilities, (2) town center connectivity, (3) indoor areas, (4) around the home/neighborhood, (5) town center physical activity resources, (6) school grounds, and (7) outdoor areas.

Assessing Use of Specific Environments and Physical Activity in Specific Environments

Questionnaires and Observational Measures

Tools measuring use of specific environments are not presented in the Measures Registry. Examples of questionnaires to measure these constructs can be found on Dr. James Sallis’ website. Counting users, pedestrians, and/or bicyclists within specific environments can provide an objective way of assessing use of environments, and counting can be done manually or with automated devices. Many cities and counties use automated counters to track pedestrian and bicycle traffic, and some parks track users in a similar manner. Issues to consider with manual counting include when and how frequently to collect assessments. Both manual and automated counting strategies may need to take into account non-independence (i.e., people being counted at multiple sites).

Commonly used tools for assessing physical activity in specific environments include the observational audit tools System for Observing Play and Recreation in Communities (SOPARC) and System for Observing Play and Leisure Activity in Youth (SOPLAY). These tools include checklists of environmental features to describe the environments being captured, and can be found in the Measures Registry.

GPS

Several studies have used person-worn GPS tracking devices in physical activity research to capture the...
environments that an individual encounters across the
day rather than simply the environment around a specific
origin (e.g., home neighborhood). These methods allow
investigation of the environments in which physical activity
occurs as well as how exposure to various spatial contexts
may influence physical activity. Several tools exist for making
the use of GPS data in built environment research feasible
(e.g., PALMS).121

**Gaps and Limitations of Existing Measures**

As is apparent in Table 1, few to no strong measures exist
for some methods and settings (e.g., home audits, park
self-report measures). In some settings multiple measures
exist, making it difficult to choose (e.g., streetscape audit
and report tools). Several existing measures with evidence
for reliability and validity are lengthy and require complex
training and/or scoring procedures because they were
designed for research. Such tools may not fit the needs of
practitioners because of their cost, complexity, and/or user
burden. Much progress still needs to be made before both
research- and practice-quality tools exist for all methods
and settings.
Selecting Measures
Defining Project Purpose

Physical activity environment projects can have a wide variety of purposes and objectives, as highlighted in the Behavioral Epidemiology Framework presented in Figure 4. The figure distinguishes five project purposes: Basic, Epidemiology (or Health Outcomes), Surveillance, Theory and Correlates, and Intervention. The majority of projects in the practice domain fall under Surveillance and Intervention Research. Environmental research fits within each of the five levels, but has less of a role in Basic Research. As outlined in Section 2, physical activity environment research is guided by ecological models that suggest that physical activity is influenced by factors from multiple levels. Although many study purposes are multi-level, this section specifically focuses on selecting measures at the environmental level.

Once the study purpose is refined using the Behavioral Epidemiology Framework, clearly defined project objectives can be identified to facilitate selecting appropriate environmental assessment tools. Objectives should be SMART: Specific, Measureable, Achievable, Realistic, and Time-bound. The primary objectives of conducting environmental assessments include:

- To conduct periodic measures for public health planning purposes and assessing trends over time (i.e., surveillance).
- To observe associations between environmental attributes and physical activity, sedentary behavior, or health (i.e., observational/correlates study)
- To conduct intervention research or evaluation involving one or more of the following:
  - To educate and/or engage community members in advocacy around environments and health (i.e., community engagement)
  - To identify areas and/or attributes in need of improvement (i.e., needs assessment); can also include assessment of disparities in physical activity environmental attributes
  - To evaluate whether and how environmental attributes change over time or as the result of an intervention (i.e., evaluation)

Whether grounded in research or practice, most projects fall into one or more of these three purposes. An example of a multiple-purpose project is when an investigator may conduct a physical activity intervention and wants to use built environment assessment to 1) engage participants in advocacy efforts, and 2) evaluate whether environmental improvements occur. Correlates research also can be combined with health outcomes research, for example when physical activity is investigated as a mediator of the link between the built environment and health outcomes.

Considerations in Selecting Measures

Once the project purpose is defined using SMART objectives, multiple factors should be considered in selecting the most appropriate tool(s) for environmental assessment. These factors include the setting of interest, comprehensiveness, reliability and validity, relevance to population being studied, resources and expertise required, and flexibility and adaptability to local needs.
Figure 4: Behavioral Epidemiology Framework

Welk, 2002. The figure was adapted to feature physical activity environment measures.
Setting of Interest

Selecting the setting of interest is often evident once the project objectives are finalized. A researcher should select the setting(s) that affect the population being studied or the setting in which the intervention is taking place. In youth, this often involves homes, schools, neighborhoods, parks, and recreation settings. Because it is often not feasible to assess all of these settings, settings should be prioritized based on their relevance (e.g., where youth spend a large amount of time), availability of partners (e.g., where partners work and/or have influence), and/or potential for intervention. However, because it is uncommon for a single measure to cover multiple settings, multiple measures may be needed to adequately account for the variance in youth’s overall physical activity.

A practitioner or researcher should select the settings that have potential for being affected by policy or environmental changes. For example, a needs assessment could be conducted on streetscape characteristics around schools seeking or receiving Safe Routes to School funding. When using environmental assessment for community engagement, the involvement of community members could inform which setting(s) to target.

Comprehensiveness

Comprehensiveness is an examination of whether a measure or set of measures sufficiently assesses the aspects of the physical activity environment that need to be assessed. For example, if the aim of a project is to examine changes in the neighborhood built environment following the installation of a new bike/pedestrian trail, measures of the quality of as well as the access to (e.g., pedestrian features linking the trail to existing neighborhood) the bike/pedestrian trail could be examined. Comprehensiveness of measures includes the number of settings as well as the number of attributes or constructs within each setting being selected. Comprehensiveness should be balanced with project goals, feasibility, resources required, and measurement properties (e.g., reliability and validity). However, comprehensiveness is not necessarily the equivalent of length. Long measures can include multiple inter-related items representing single constructs (e.g., multi-item scales) and brief measures can cover multiple constructs using one to two items per construct.

Reliability and Validity

Having evidence of acceptable measurement properties is important for both research- and practice-based projects. Tools with poor reliability and/or validity limit the ability to detect differences among groups, changes over time, or associations between environmental attributes and health outcomes. Reliability and validity are less important when environmental assessment is used for community engagement, but if the results of the assessments will be used for a needs assessment or to target environmental changes, it is important that the tool have acceptable measurement properties. Visit the “Validity” and “Reliability” tabs for each measure in the Measures Registry, where results are presented when available.

Relevance to the Population Being Studied

It is important to note that reliability and validity can differ across geographic areas and participant populations, the latter being particularly relevant for self-report measures. A tool validated in urban areas or with a predominantly white non-Hispanic sample may not be reliable or valid for use in other geographic areas, cultures, or samples. Similarly, a measure’s utility is specific to the purpose of the project. A tool may have acceptable validity for observational research but have poor sensitivity for detecting environmental changes over time, often because the response categories are very broad or there are no items to assess the planned intervention. Thus, users should consider whether the tool was designed and evaluated in a similar geographic area and/or population and for a similar purpose as the user intends. A majority of the measures in the Measures Registry have not been tested in multiple geographic areas and/or populations, particularly diverse populations in terms of race/ethnicity and socioeconomic status.

Rural and international environments are two key areas where environmental assessment tools diverge, particularly those related to community design. A strength of the NCCOR Measures Registry is that rural measures are separated from urban/metropolitan measures so they can be easily identified. When applying a questionnaire or audit tool to another country, adaptations are often needed, and GIS-derived data and relevant variables can vary drastically across countries due to regional and historical differences in community design.
Resources and Expertise Required

The practicality of using any environmental assessment tool comes down to its complexity and cost for data collection and management. The “How To Use” tab in the Measures Registry provides useful information about resources needed to administer the measure, when available. GIS expertise is specialized and costly, and not all GIS technicians will have expertise in the variables related to physical activity. Accessing GIS data from multiple sources and deriving variables are complex and time-consuming activities. However, GIS-based methods are advantageous in large samples of individuals from few geographic regions because the cost is primarily based on the number of jurisdictions and variables rather than number of people. Audits, on the other hand, carry a per-person or per-environment (e.g., school, neighborhood) cost, making them less feasible for use in large studies. Audits require expertise in training in observational methods, ongoing attention to quality control, and management of complex data. The primary limitation to questionnaires is participant burden. Knowledge and careful consideration of the study participants and the time required to complete all study questionnaires will inform the selection of appropriate questionnaires. Data management demands are least complex with questionnaires.

Complexity and cost also vary within each method of assessment. Data collection, training, and scoring guides greatly reduce the complexity and cost of using GIS, audit, and questionnaire tools. The quality and comprehensiveness of the guides should be assessed before a measure is selected. Measure developers may or may not be able to answer questions or provide assistance, so thorough documentation is highly desirable. Smartphone and tablet applications can significantly reduce data management burden with audits and questionnaires as long as they have been fully vetted.

Flexibility and Adaptability

It is important to choose a flexible measure in situations where the tool may need to be adapted or tailored to specific populations or settings. New items in audit tools and questionnaires should be pilot tested for inter-rater or test-retest reliability when possible. Tools that use smartphone or tablet applications or require complex scoring algorithms may be less flexible because of difficulties in making software or scoring modifications. GIS-based measures can be flexible; for example, a variable can be computed multiple ways, and the options will be determined by the datasets available.
Suggested Process for Using the Measures Registry

The following section outlines the main steps that users can follow to use the Measures Registry to select the most appropriate measures for a designated purpose.

**Searching and Filtering Results**

The Measures Registry can be accessed from the NCCOR home page, under the “Tools” tab. The physical activity environment measures in the Measures Registry are organized by “Measure Type” (i.e., measurement method), “Ages” covered, and “Context” (i.e., urban or rural). The check boxes within each category can be used to narrow the results field. Once the user selects the desired measure type and context, a list of measures fitting those criteria is displayed. When a measure has been evaluated (i.e., tested for reliability or validity), the Registry will most likely include the measurement development or evaluation publication. If multiple measurement evaluation publications exist for a given measure, the Registry will typically include each of the publications. If a measure development or evaluation publication does not exist, the Registry will include the first paper published using the measure, typically a physical activity correlates study. Many of the publications in the Registry are not measure development or evaluation studies.

The Measures Registry does not include a setting code specific to physical activity environments to organize studies. When a user is interested in a particular setting, it is recommended to use the search function. Entering keywords such as “home environment” or “school” into the search box results in a list of publications that included those keywords in the title or abstract. The search function can be used in combination with the category check boxes to further refine the search, such as limiting the search to a specific setting and measurement method.

**Navigating the Information Tabs Within Each Publication**

Clicking on a publication’s title will open a link with more detailed information about the measurement tool. It is recommended that each tab be viewed in detail while keeping in mind the selection considerations outlined above. The “At A Glance” tab includes helpful information when available, such as the length, constructs covered, and how to obtain the measure. The “Study Design” tab reports the characteristics of the sample used to develop and evaluate the measure, so users can consider whether the tool is appropriate for the population they intend to study. The “How To Use” tab includes information on how the tool is administered and whether data collection and/or analysis protocols exist. In circumstances where the Measures Registry does not include a link to the measure or protocols, the user should contact the authors of the study. Finally, the “Validity” and “Reliability” tabs include specific results from the publication on the tool’s measurement properties. If a tool has multiple publications in the Measures Registry, the user should view the tabs for each publication.
Collecting and Reporting Data
Several factors should be considered in the planning and implementation of data collection, as well as when analyzing and reporting data.

Data collection considerations are detailed in this section, organized by measurement method, and followed by general statistical considerations when working with environmental data. Many users will benefit from consulting an expert in the field of physical activity environment assessment or someone with experience using a specific tool.

GIS-based Measurement Considerations

Several considerations are important when determining whether to use GIS-based measurements in a study or project, and these are described below.

Needed Expertise

Although GIS software such as ArcGIS is user-friendly, the level of data processing and analyses involved in built environment assessments often requires GIS expertise. Fortunately, GIS experts can be found through local universities and government departments. GIS is commonly used in Geography, Public Affairs, City and Transportation Planning, and Health Departments. GIS is used with a wide variety of data, so project staff should make sure a GIS expert is recruited whose skills and interests are relevant to the study or project.

Access to Databases

One challenge of GIS-based environmental assessment is that GIS databases can be difficult to obtain. Local Metropolitan Planning Organization’s (MPO) websites are the best places to start when searching for GIS databases because many MPOs maintain a fairly large collection of GIS data for their region, such as road networks, parcels, and land use shapefiles. Other sources of GIS data include local municipalities, which often will share their GIS databases, sometimes for a fee. GIS data can also be obtained or purchased from sources such as ESRI and the U.S. Census. It is also possible to add an original audit or survey data into GIS.

Variable Computations

After determining geographic coordinates (e.g., for building addresses or from GPS) and obtaining the necessary geodatabases, built environment variables need to be created. Creating interpretable and useful built environment variables requires careful thought and multi-step computations involving multiple indicators and intermediary variables. Net residential density, for example, involves dividing the number of residential housing units by the sum land area across all residential parcels. Intersection density involves the sum number of intersections divided by the total land area. Mixed land use variables can be more complicated to compute. The International Physical Activity and Environment Network created a detailed protocol covering computations of numerous GIS-related built environment variables. This protocol serves as an excellent resource for those working on GIS-based built environment assessments and is publicly available online. A similar resource has been created mainly for application in the United States.

It is important to note that GIS variables can be computed in various ways and differ across studies. This should be taken into account when comparing findings across studies and selecting computations. Appropriate computations are those with the strongest evidence in relation to physical activity, relevance to the population being studied, and ability to be compared across geographic areas.

Buffer Size, Type, and Origin

GIS-based (e.g., community design) built environment assessments for physical activity have traditionally captured the spatial context around people’s homes. This method involves creating spatial boundaries around people’s homes in GIS to represent the environmental context believed to influence physical activity. The primary approaches to creating the spatial boundaries are: (1) radial/Euclidean buffers, which involve drawing a straight line from the home for a specified distance (e.g., 1 km) and using the line as the radius to create a circle, and (2) street network-based buffers, where a line is drawn a given distance from the home through the street network. Street-network buffers are believed to better represent opportunities for walking and are generally supported in previous studies, though results have been similar across buffer types.
Buffer size also is important to consider, and the selection may vary based on the population of interest. For example, smaller buffer sizes, around 500 meters, are more appropriate for younger children with limited mobility freedom, whereas larger buffers, around 1 kilometer, better represent an adolescent’s walking space. Although the home neighborhood is almost always the origin of the buffer, evidence is accumulating suggesting that community design aspects outside of one’s home neighborhood, such as the school and “neighborhoods,” also are important. Some studies have used GPS to determine all of the locations/geographic coordinates a person encounters and calculated GIS-based physical activity environment variables from this information. This dynamic approach allows a comprehensive assessment of environmental exposures but may be limited by data availability.

Observation Measurement Considerations

As with GIS-based measures, users interested in working with observation measures should first consider the following issues.

Training

Conducting observational audits can require significant resources and training, particularly for detailed research-oriented tools. Good audit tools are accompanied by a user guide that includes information on training coders. The following training steps should be used when conducting observational audits for research or when high levels of accuracy are required.

1. Before beginning data collection, each coder should be trained by a “master trainer” who has been trained and certified by the tool developer(s) or another master trainer.
2. Initial coder training should involve going through the audit tool together in the field (e.g., while walking the route being coded) while the master trainer provides definitions, instructions, and explanations for each selected code.
3. Coders should be deployed to audit a small number of environments (e.g., five segments or two parks) that the trainer also has audited. Each coder’s data should be compared to the trainer’s gold standard data. Each coder should exhibit a high level of agreement (e.g., >80 percent or 85 percent) with the master trainer before being certified for data collection.
4. Project staff should continue to monitor inter-observer agreement on a small number of environments throughout data collection, then provide feedback and retraining as needed. Ongoing monitoring is needed to ensure high quality of data.

These procedures are ideal for research studies but may not be applicable to practice-based projects, such as those that engage community members in audits for educational purposes or to identify areas needing improvement. Whenever possible, community members should be trained to conduct the audits, and simple certification procedures should be implemented based ideally on agreement with the trainer or at least an assessment of inter-rater reliability.

Selection of Environment Samples for Auditing

Because it is rarely feasible to audit all streets in a community or all parks in a city, a systematic approach to sampling is needed. One approach is to select generalizable samples of the environment/setting of interest when using an audit tool. For example, when using a streetscape audit tool, the user must decide which and how many street segments and/or routes to assess. When using a park audit tool, the user must decide which and how many parks to audit. However, little research is available to guide this selection process.

Another approach is to select samples for a specific purpose. If the goal of the study is to examine the safety of streetscapes around schools, a random sample of schools can be drawn. If the goal is to assess disparities in park quality, then low- and high-income areas can be sampled, and parks within those areas can be assessed. If the goal is to link the built environment to participants’ physical activity, the sample environments can be selected based on proximity to each participant’s home. The environments can also be selected based on the nearest cluster of destinations (e.g., shops, restaurants) to represent the most likely walking path between the participant’s home and these destinations. When the objective is to classify a neighborhood for activity-friendliness, a random sample of environments (e.g., parks, streets) can be selected to represent the neighborhood. Some evidence suggests that assessing between 25 percent and 50 percent of the street segments in the neighborhood may be sufficient to provide a representative estimate of the area.

Use of Google Street View for Observational Audits

Google Street View is a feature of Google Maps that provides images from views along streets across the world. Several recent studies have investigated whether streetscape audits can be reliably and validity completed using Google Street View. Each of these studies concluded that Google Street View is acceptable for such tasks. One important consideration when using Google Street View is to be cognizant of the date the images were collected. If the image is old, the environment could have changed, thus limiting the validity of the audit. Images can be taken during different
advocate for specific environmental improvements.

useful in needs assessment projects that aim to identify and studies that engage community members. It can also be relevant to the participant. Photovoice has utility in intervention rich qualitative information on specific environment factors narratives, can supplement quantitative measures by providing The photographs, sometimes paired with written or verbal

The photographs, sometimes paired with written or verbal

advocate for specific environmental improvements.

seasons, so if an obstruction (e.g., tree foliage) appears in an image, the user could search for other (newer or older) images of the same area. Google Earth has been available since 2007, so several images have been captured for many areas across the world over this time span. A helpful guide for conducting streetscape audits using Google Street View was developed by Wilson and Kelly.\textsuperscript{160}

**Smartphone and Tablet Applications**

Traditionally, observational audits have been completed using pen and paper. More recently, several audit tools have been incorporated into smartphone and tablet computer applications, such as the SOPARC Online App: System for Observing Play and Recreation in Communities (SOPARC).\textsuperscript{141} Advantages of applications are that data are recorded directly into the smartphone or tablet, skip-patterns can be automated, data can be transferred and stored on a secure server, and in some instances data can be scored automatically. Applications can be costly to build and maintain, but more are continuing to emerge in environmental assessment.

**Photovoice**

Photovoice has seen recent use in physical activity environment assessment and allows participants to capture photographs of their environment (e.g., Buman, 2013).\textsuperscript{142} The photographs, sometimes paired with written or verbal narratives, can supplement quantitative measures by providing rich qualitative information on specific environment factors relevant to the participant. Photovoice has utility in intervention studies that engage community members. It can also be useful in needs assessment projects that aim to identify and advocate for specific environmental improvements.

**Statistical Considerations**

In additional to measures-related considerations, users will also need to consider several statistical issues.

**Nested Data**

Environmental assessment often involves nested (i.e., multi-level) data when multiple participants are included from each environment. Users of environmental assessment tools need to consider the unit of analysis in their data collection and analysis efforts. When personal characteristics are of interest, such as physical activity, the participant is the unit of analysis. For example, when data from a school environment tool (i.e., audit or self- or proxy-report) are compared to physical activity data from multiple children from that school, the design is considered nested. In this circumstance, the sample size and statistical power (i.e., ability to detect a true association) are driven primarily by the number of schools rather than the number of participants. An ideal design would include a large number of schools and a small number of randomly selected students (e.g., 20–30) per school. Assessing more students per school would only minimally improve power. When analyzing nested data, mixed-effects models must be used to account for the non-independence of participants within settings.\textsuperscript{143} Most statistical packages are capable of handling these types of models (e.g., Singer, 1998).\textsuperscript{144}

**Standardized and Unstandardized Regression Coefficients**

Unstandardized regression coefficients are interpreted as the change in the dependent variable for every 1-unit change in the independent variable. The unit of measurement is retained, so if the dependent variable is minutes of physical activity, for example, the unstandardized coefficient represents minutes of physical activity. For example, an unstandardized regression coefficient of $B = -8.6$ for the association between distance to the nearest park in kilometers (independent variable) and minutes per day of physical activity (dependent variable) means that for every additional one kilometer in distance to the nearest park, participants had 8.6 fewer minutes per day of physical activity. So a person living one kilometer from a park would have 8.6 fewer minutes per day of physical activity than a person living zero kilometers from a park, and a person living two kilometers from a park would have 17.2 fewer minutes per day of physical activity than a person living zero kilometers from a park. Thus, unstandardized regression coefficients are particularly useful when the unit of measurement is meaningful, such as a quantity of something (e.g., minutes of physical activity, number of parks, number of intersections per acre). Standardized regression coefficients involve standardizing the variances of the dependent and independent variables to 1, similar to creating z-scores. Standardized coefficients cannot be interpreted in the unit of measurement of either the dependent or independent variable, but are useful for comparing effect sizes across variables and statistical models. Both types of coefficients add value and are commonly used in environmental assessment, but unstandardized coefficients are most informative to decision making.

**Multicollinearity**

Independent variables with high correlations should be grouped into scales or indices to prevent multicollinearity, which violates the assumptions of multiple regression analysis. Investigators commonly use a threshold correlation of $r = 0.7$ or $r = 0.8$ to denote “high” correlations among independent variables. Highly correlated items can be combined into scales using factor analysis or by summing or averaging values across items. See Section 4 for more information on scales and indices.
Case Studies of Selecting Measures
Three case studies are provided below to walk users through the processing of selecting appropriate measures for their project. The case studies cover a wide range of research and practice project purposes and apply several of the selection considerations covered in Section 7.
A local bicycle and pedestrian advocacy organization is working with the city planning department to improve environments around schools to support active living. The organization plans to apply for grant funding to support specific environmental improvements and would like the improvement targets to be identified through a community needs assessment. Their goal is to identify specific locations and types of improvements for which to seek funding.

The organization is interested in environmental attributes that would be feasible to modify during the two-year grant period. They do not have much influence over home, school, and other building environments, but they have some influence over neighborhood and park environment modifications through their partnership with the city planning and parks departments.

The organization also wants to assess neighborhood and park environments, but not to use GIS-based measures because community design attributes available in GIS are at the macro-level and not easily changed (e.g., residential density and land use mix). The organization considers using self-reports but ultimately rejects that idea because (1) they do not have expertise in community surveys, and (2) audit tools are more specific with regard to identifying areas and attributes to target for improvement. Thus, both streetscapes and park audit tools appear to be well-suited for this project. Because the organization wishes to involve community members in the project, they need brief tools that will require little training, pinpoint modifiable features in need of improvement, and produce simple results metrics that can be easily communicated to the grant funder.

In looking for a streetscape tool, the organization selects the Measures Registry filter options “Physical Activity Environment” and “Environmental Observation,” and types “street” into the search field. Approximately 15 results are shown. Some results are immediately ruled out for various reasons (e.g., Google Street View tool, rural tool, senior tool). About 10 tools are compared using the “Compare” check box. The organization narrows the results to the tools that have the instrument available in the Registry, approximately five tools. Based on the content desired, the organization considers the Active Neighborhood Checklist, Microscale Audit of Pedestrian Streetscapes-Mini (MAPS-Mini), Pedestrian Environmental Quality Index (PEQI), and Systematic Pedestrian and Cycling Environment Scan (SPACES). PEQI is ruled out because little evidence for reliability and validity is available. SPACES is ruled out because it is too lengthy.

The organization uses the same search parameters to identify park audit tools, and “park” is typed into the search box. About 15 results are provided but some are immediately ruled out because they are not exclusively focused on parks and thus lack the detail the organization desires. Using the “Compare” check box, the organization narrows the candidates to the Bedimo-Rung Assessment Tool, Environmental Assessment of Public Recreation Spaces (EAPRS), Physical Activity Resource Assessment (PARA) and Community Park Audit Tool (CPAT). PARA is ruled out because it lacks the level of detail desired.

For the streetscape audit tool, the organization selects the Active Neighborhood Checklist because it is brief, includes a codebook and training materials, has been previously used by practitioners, and has evidence of reliability and validity. They strategically select street blocks to audit near elementary schools to represent points of pedestrian access to each school. The CPAT park audit tool is selected because of its previous use by practitioners and evidence of reliability and validity. However, CPAT includes 140 items, and the organization is concerned about resources and time burden. Thus, they select a subset of items based on community input and published evidence of reliability and validity of the subscales.
A project team is planning a large multi-region study to better understand how the transportation and recreation physical activity of children ages 6 to 10 years is related to the supportiveness of the various environments the children encounter on a regular basis.

The project team would like to assess several of the key environments where children spend their time. Based on previous studies, the team identifies the home, school, and neighborhood as key environments to include in their study. Because of the large sample size needed and desire to cover multiple environments, the team needs to identify measures that do not pose a high data collection burden or cost. Although streetscape audit tools would inform this project, the team ultimately decides not to use audit tools because of the large amount of time required to travel to participants’ neighborhoods and collect audit data (though observations using Google Street View would be an option in this circumstance).

For community design and transportation system measures, the project team considers GIS-based measures, given their relative ease in large samples, but they make sure to investigate the availability and comparability of GIS databases across the regions being considered for the study. For home, school, and streetscape environment measures, the team prefers proxy-report (e.g., parent report) when available because audits would be too costly, and the children are viewed as being too young to use self-report. Because this is a novel project, the team desires comprehensive measures with strong evidence of reliability and validity.

The team searches the Measures Registry by selecting the filter options “Physical Activity Environment” and “GIS.” Age category is not selected because of the relevance of GIS variables across age groups. Rather than using the “Compare” check box, the team views each of the approximately 25 results individually. Measures without evidence of construct validity, as indicated in the “Validity” tab, are ruled out. The remaining results are reviewed and examined for consensus (i.e., GIS variables that appear across multiple results).

More than 75 results are displayed when the team selects the “Physical Activity Environment” and “Questionnaire” check boxes, and type in “neighborhood.” Results that have the instrument attached and that appear to meet their criteria based on title are compared using the “Compare” check box. Those without evidence of reliability and validity are ruled out. As the results narrow, the team views the tools for their comprehensiveness and frequency of use. Some of the candidate measures are entered into a scientific literature search engine to estimate frequency of use.

The team searches for home environment report tools using the search term “Home” and selecting the “Questionnaire” check box. In reviewing the approximately 50 results, only those with a link to the instrument are considered, and the project team has difficulty identifying physical environment tools that are specific to the home setting (most are related to the home neighborhood or parenting behaviors). Two potential candidates are identified and reviewed for more detail: the Healthy Home Survey and Perceived Environment Survey on Safety, Aesthetics and Physical Activity. The same process is used to search for school physical environment report tools, with similar difficulties. Many of the search results are for school neighborhoods, child care facilities, or school policies/practices, or lack a link to the instrument. The team expands their search to include observational audits. Very few candidates are identified from the Registry, suggesting a lack of assessment tools in this area.

The project team selects GIS-based residential density, land use mix, street connectivity, and park access for assessing community design and transportation system factors because of their consistent use in various publications in the Measures Registry and their associations with children’s physical activity in previous studies. They consider GIS-based composite measures/indices, such as the walkability index by Frank et al., and choose the Cervero and Kockelman index because of its level of detail and...
comprehensiveness, compared to the Frank index, which comprises only four variables.

The parent report Neighborhood Environment Walkability Survey (NEWS) is selected because it covers features not available through GIS, such as sidewalk attributes and aesthetics, and has evidence of reliability and validity in relation to youth physical activity. The parent report tool developed by Hume et al. is selected as the home environment report tool because of its strong measurement properties and previous use in research. The school environment audit tool developed by Jones et al. is selected as the school environment report tool because of its level of detail, but it is first tested in a subsample to investigate its reliability and validity when completed by trained school key informants rather than using in-person audits.
A redevelopment grant is awarded to a city with special emphasis on pedestrian-oriented renovation projects through the city’s main urban corridor. A project team would like to evaluate the extent to which the grant results in improvements in streetscape features known in previous studies to be associated with physical activity.

The team chooses a pre-post design so that environmental changes can be captured across the one-year project. Because the renovation projects are focused on streetscape features, the team narrows in on streetscape report and audit tools. Although community member perceptions are important, audit tools are desirable because they provide a higher level of specificity than report tools. The constructs assessed need to have shown consistent relationships with physical activity in previous studies, so construct validity is especially important. The tool needs to be sensitive to change, particularly the changes being targeted by the renovation projects. Although most environmental assessment tools have not been evaluated for whether they can capture changes over short periods of time (e.g., one-year), the level of specificity provided by some audit tools is likely to be sensitive to the changes being targeted.

The team selects the Measures Registry filter options “Physical Activity Environment” and “Environmental Observation,” and types “street” into the search field. However, several relevant tools are left out when “street” is used in the search box, so this term is omitted. Simple checklist tools are ruled out because high detail and specificity are desired. A handful of relevant tools are selected based on their title and availability of the instrument. The candidates are compared using the “Compare” check box. The team considers the Irvine-Minnesota Inventory (IMI), Microscale Audit of Pedestrian Streetscapes (MAPS), and Pedestrian Environment Data Scan (PEDS) because they are extensive instruments with similar content and have supporting evidence.

The IMI tool is selected because of its specificity, evidence of reliability and validity, and use in multiple previous studies. In reviewing previous publications using IMI, the team identifies several items and scales that (1) have consistent associations with walking for transportation and/or leisure across age groups, and (2) appear to have the ability of being affected over the one-year time frame of the grant.
Next Steps in Physical Activity Environment Assessment
Continued development in the field of measuring the physical activity environment would benefit from attention to several areas, including: (1) considering the level of variability obtained in food environmental measures, (2) moving beyond observational data and increasing the evaluation of longitudinal relationships and change over time, and (3) increasing attention paid to the expected associations between environmental measures and outcomes, and (4) continued efforts to promote use of common measures where possible.

Though several validated measures are available to choose from in most settings, some clear gaps remain. The bigger need identified by an expert working group is for validated measures that are short enough to be feasible for use by practitioners. Interventions to improve physical activity environments are recommended by international and national authoritative organizations, and measures should be used to determine intervention needs and evaluate outcomes. Few brief, validated measures are available, but examples include MAPS-Mini for streetscapes, PARA for parks, and PANES for perceived neighborhood walkability.

Some measures in the Measures Registry, particularly audit measures, have evidence of reliability, but evidence is still needed on whether the constructs are consistently associated with physical activity (i.e., construct validity). For measures with evidence of validity in one population subgroup, it would be useful to expand evidence by studying other age groups, income groups, and specific race and ethnicity subgroups to evaluate generalizability of effects. An important consideration is whether measures are sensitive to change, but very few have been evaluated for this ability. An important improvement for GIS measures is to document the procedures used to create variables, so the variables become more comparable. Some guides help with standardization of procedure, but they depend on comparable, accurate, recent data being available. Thus, improvement in the quality and availability of physical activity-relevant data in GIS is a high priority. This will require collaboration with diverse government agencies such as transportation, city planning, and taxation, though crowdsourcing of data is a possibility.

Continuing barriers to the use of physical activity environment measures are the burden and cost of data collection, complexity of data management, lack of clarity in scoring and interpreting results, and need for adapting measures. The Built Environment Assessment Training (BEAT) Think Tank has recommended several strategies for advancing the use environmental measures:

- Develop simplified but validated measures to encourage more use by researchers and practitioners.
- Develop technological tools to simplify data collection, data management, scoring, and analysis.
- Improve scoring systems, with freely available syntax or software.
- Create guidelines for adaptation of existing measures and online access to various versions.
- Provide online training for data collection, data management, and interpretation of results.
- Encourage professional organizations to support the use of validated environmental measures and help the field develop consensus about definitions of constructs, preferred methods, and scoring.
- Incorporate training in environmental assessment in university courses in multiple departments.
- Use social media to support use and interpretation of environmental measures.
Conclusion
Many measures of physical activity environments have been developed over a relatively short time period. Options exist to measure almost all common physical activity settings, and measures are available in several modes, though not for all settings. Thus, available measures are likely to meet most needs for research and practice. Development of new measures and refinement of existing measures will continue, and this continued evolution is important for the field as new concepts are examined. However, the number and diversity of physical activity environmental measures in the Measures Registry can make it difficult to select appropriate measures and may discourage their use altogether. The goal of this User Guide, therefore, is to provide an overview of the field of physical activity environment measurement, offer general guidance about selecting measures to suit each user's needs, and make the Measures Registry a more user-friendly and valuable resource. This guidance should be applicable even as more measures are added to the Registry. We hope this User Guide encourages greater use of physical activity environment measures in research and practice.
Additional Resources on Physical Activity Environment Measurement
Several published resources that are useful in physical activity environment research can be used to supplement the information provided in this Guide:

- **Step It Up! The Surgeon General’s Call to Action to Promote Walking and Walkable Communities**
- **Active Living Research (ALR) website**, which also contains a compilation of environment measures and supporting materials
  
  **ALR Briefs Specific to Measurement**
  - **The Role of Self-Selection in Explaining the Effect of Built Environment on Active Travel**
  - **Counting Bicyclists and Pedestrians to Inform Transportation Planning**

  **Review Papers**
  - **Measuring the Built Environment for Physical Activity: State of the Science**
  - **Assessing Perceptions of Environments for Active Living**

  **GIS Protocols**
  - **Built Environment and Physical Activity: GIS Templates and Variable Naming Conventions**
  - **Environment and Physical Activity GIS Protocols Manual**
References

2. Heath GW, Brownson RC, Kruger J et al. The effectiveness of urban design and land use and transport policies and practices to increase physical activity: Systematic review. J Phys Activ Health. 2006; 3(Suppl 1): 555-76.


12. Institute of Medicine (IOM). Educating the student body: taking physical activity and physical education to school. Washington (DC); 2013.


References • 47
55. Prins RG, Oenema A, van der Horst K et al. Objective and perceived availability of physical activity opportunities:


