Published in final edited form as:
Health Place. 2012 September ; 18(5): 991–999. doi:10.1016/j.healthplace.2012.06.008.

Assessing Urban Walking Trail Use and Changes in the Trail Environment Using Systematic Observational Protocols

Duncan C. Meyers[a], **Dawn K. Wilson**[a], **Kassandra A. Kugler**[a], **Natalie Colabianchi**[a], **Thomas L. McKenzie**[b], **Barbara E. Ainsworth**[c], **Julian Reed**[d], and **Sara C. Schmidt**[a]

[a]University of South Carolina

[b]San Diego State University

[c]Arizona State University

[d]Furman University

Abstract

This study evaluated the extent to which two systematic observation protocols which were modified for underserved communities (low income, minorities) could be utilized to reliably assess a) use of walking trails and b) physical environmental features of these trails. This study was a supplement to the Positive Action for Today's Health (PATH) walking trial. The modified tools were shown to be reliable methods for a) measuring trail use and b) assessing physical features of the trail in underserved environments. Reliability data for measuring trail use were found to be high (ICC = .98, $p < .01$). Reliabilities for measuring features of the trail ranged from fair to highly reliable (κ = .77 - 1.00; ICC = .34 - 1.00). The observation tools that were customized for this study were shown to be reliable instruments for measuring trail use and assessing physical features of walking trails in underserved communities.

Keywords

physical activity; observational methods; low income; African Americans; walking trails

Introduction

Systematic observation procedures are particularly important for assessing physical activity (PA) (McKenzie, 2010) because they can be designed for a particular purpose (e.g., observing walking behavior) and include protocols for quantifying specific behaviors or other categories of interest. Such protocols can be used to assess the extent to which community members engage in PA and the adequacy of physical supports for PA in communities (e.g., walking trails). Assessing such behaviors and supports is an important step in understanding how community-based interventions impact walking and other PA.

Corresponding author. Duncan C. Meyers, 1512 Pendleton St. Columbia, SC 29208, Tel.: 803.414.1090; fax: 803.777.9558. meyersd@mailbox.sc.edu.

Relative to assessing walking and other PA, self-report methodology has substantial limitations related to estimating time and recall bias, and accelerometer data do not provide information on activity contexts (McKenzie, 2010; Sallis and Owen, 1999). Given these limitations, systematic observation is important because it allows for the measurement of activity levels as well as when, where, and with whom it occurs. Measurement tools have been developed to count people in specific settings, including those using walking tracks (e.g., McKenzie and Cohen, 2007). Audit tools are often used to identify physical environmental qualities (e.g., quality of sidewalks) that are related to PA (Brownson et al., 2009). While there is no widely accepted standard protocol for measuring parks and recreation environments, the high-quality tools that are currently available enable researchers to expand the base of evidence on how parks and recreation areas contribute to PA (Floyd et al., 2009). A remaining challenge, however, is to continually improve these measures to ensure their relevance for diverse populations and contexts (Brownson et al., 2009). While much progress has been made regarding measures that help researchers and practitioners assess the built environment, these can be considered "first-generation" measures that need further development (Brownson et al., 2009).

Measures for assessing trail use and physical trail qualities have yet to be refined and modified for use in underserved communities. Two systematic observational protocols were developed and utilized as part of the present study as a supplement to the Positive Action for Today's Health (PATH) walking trial. PATH employed a police-patrolled walking intervention to increase safety (training community leaders as walking captains, hiring off-duty police officers to patrol the walking trail, and containing stray dogs) and access for PA (marking a walking route) in communities with a history of high crime rates and low median household income (Wilson et al., 2010). The purpose of the current study was to describe the protocols used in the PATH project and present the reliability data collected in two communities.

First, information is presented regarding a modified observational protocol to assess the use of an urban walking trail. We conceptualize a walking trail as a marked or established route with a starting and end point that can be used for exercise, recreation, and/or commuting. These walking trails could be located in a variety of settings (e.g., parks, neighborhoods, indoor facilities like gymnasiums) in any type of community (e.g., rural or urban). The walking trails investigated in this study were identified on existing sidewalks in underserved communities. Features of this protocol were designed to capture 1) participation in organized walking groups – a component of the PATH intervention – as well as 2) trail use at the community level. Next, a separate systematic observational protocol is discussed that assesses physical features of the trail in relation to design and functionality, safety, and trail maintenance. For both protocols, existing systematic observation procedures were modified for use in underserved neighborhoods (high crime, low income). Finally, reliability data (inter-observer agreement reliability) from baseline data are presented to document the reliability of using these modified protocols.

METHODS

Study Design

The PATH project is a randomized trial designed to examine a 24-month environmental intervention designed to improve safety and access for PA and trail use in three underserved communities that has been previously described (Wilson et al., 2010). Three communities were matched based on census tract level information (e.g., race, median income, poverty status, crime index, health index score, physical inactivity), with two communities receiving an intervention and the third serving as a comparison. The two communities that received the walking trail intervention (i.e., the PATH intervention communities) were the settings in which data for this study were collected.

In 2009 the populations of the PATH intervention communities were 38,412 and 32,180 (United States Census Bureau, 2012). Census data from 2009 indicate that 99% of persons in the census tracts that were selected for this study were African American (Wilson et al., 2010). The percent of foreign born persons in these communities for 2006 – 2010 was 2.6% and 4.1% (United States Census Bureau, 2012). These communities experienced high levels of chronic disease, poverty, and crime relative to state and national averages and were therefore considered underserved. For example, in 2009, 38-32% of households in the two PATH intervention communities (defined by the US census tracts they fall within) experienced poverty compared to national rates of 14.3% (DeNavas-Walt et al., 2010). Census data from 2009 also indicate that median household in the communities participating in this study ranged from $16,804 – $22,088 compared to a median income for all US households of $49,777 (DeNavas-Walt et al., 2010). For the counties that PATH intervention communities are located in, county-level crime rates (index total per capita crime) in 2006 were 666.9 and 570.2 compared to the state county-level average of 500.5 (Office of Research and Statistics, 2006).

Local community centers in these PATH intervention communities serve as the facilities for all project-related activities. The walking trails were identified by community stakeholders (e.g., community center staff, community residents) who were knowledgeable of the neighborhood climate, residents, and street system. The trails were on pre-existing sidewalks throughout each neighborhood, and these trails made a circular loop which allowed for walks to start and end at the same location (i.e., the community center).

Each walking trail was divided into segments to break the trail up into smaller – observable – units. In one of the two communities, there were 17 total segments equaling 1.33 total miles. In the other, there were 24 segments, totaling 2.35 miles. The majority of the segments were identified based on intersections (e.g., segments started whenever there was a road which intersected the trail). Although intersecting roads served as the most common means to divide the trail into segments, in some instances landmarks (e.g., municipal signs, fence posts) were used to disaggregate long segments which lacked intersecting roads. Some segments of the trail were adjacent to major roads, while others were adjacent to residential side streets. These identified trails were typified by mixed land-use and had segments that were adjacent to green space (e.g., fields and parks), commercial and civic properties (e.g., convenience stores, churches, restaurants, schools, strip malls), residential properties (e.g., apartment complexes, houses), and industrial facilities. Maps were developed for data collection which marked the trail segments.

Two coders completed the data collection protocols simultaneously, thereby providing a measure of inter-observer reliability. Having two coders also enhanced the safety of the data collection process. This study was deemed exempt by the University Institutional Review Board since de-identified data were used.

PATH Systematic Observation Protocols

Instruments were used to assess a) trail use and b) features of the trail environment. The System for Observing Play and Recreation in Communities (SOPARC; McKenzie and Cohen, 2007) was modified to assess trail use. For assessing features of the trail environment, a modified version of the Path Environmental Audit Tool (PEAT; Troped et al., 2006) was used. The data collection for these instruments took place on the walking trails in the PATH intervention communities.

SOPARC Modified Observational Tool of Community Walking (SOPARC PATH)

The modified version of the SOPARC (McKenzie and Cohen, 2007), was used to assess the number and characteristics of walking trail users and their PA levels. The original SOPARC protocol has been used to assess use of a bicycle path (Cohen et al., 2008), PA in parks located in underserved communities (McKenzie et al., 2006), and has been modified to enhance its relevance for assessing use of walking trails with multiple access points (Reed et al., 2011). This is the first study to modify the SOPARC tool to enhance its relevance for assessing use of a walking trail in underserved communities.

Through the use of momentary time sampling techniques, the SOPARC tool allows for systematic scans of each trail user within trail segments. During these systematic scans, the activity level of each trail user was coded as *sedentary* (e.g., lying down, sitting, standing, or being pushed in a stroller, wagon, wheelchair, etc.), *walking* (e.g., shifting weight back and forth, walking at a steady pace), or *vigorous* (e.g., any activity considered to have a greater energy expenditure than walking, such as running, in-line skating, biking, etc.). Since this tool uses momentary time sampling, the activity level that was first observed was recorded, even if the trail user changed activity levels on a given segment. Other user characteristics assessed in this study included age category (child – 0 to 12 years of age, teen 13 – to 20 years of age, adult – 21 to 59 years of age, and senior – 60 years of age and older), gender (male, female), and ethnicity (black, other ethnicity). In this way, the SOPARC tool allows for an estimate of the number of individuals engaged in PA, intensity with which they engaged in this activity throughout the day, plus demographic variables.

SOPARC observations were made during specified times in the morning, noon, afternoon, and evening. Having specified observation times permitted comparisons to be made between all PATH intervention communities in this study since all observations were made at the same time in both communities and on the same day. Data were collected during seven random non-contiguous days of clement weather for each day of the week (e.g., Sunday through Saturday; observation days did not directly follow one another) in a given month.

Reliability for SOPARC has been previously reported elsewhere (McKenzie et al., 2006; Tester and Baker, 2009). Specifically, both high percent agreements (McKenzie et al., 2006) and intra-class correlations (Tester and Baker, 2009) have been demonstrated. The tool's activity codes have been used and validated through similar observation systems used with youth in various PA settings, such as community parks (McKenzie et al., 2000; van der Mars et al., 2004). Furthermore, there is an abundance of evidence available indicating that momentary time sampling techniques yield valid behavioral samples (McKenzie, 2010). Support for the tool's validity and reliability has been demonstrated during its use in underserved urban community parks (e.g., McKenzie et al., 2006); however, whether this tool can be modified for use on *urban walking trails* and still be used reliably has yet to be established.

The modified SOPARC protocol – which is referred to as SOPARC PATH – utilizes both stationary observations (which were intended to capture participation in the organized walking groups that were a PATH intervention component in both communities) and mobile observations along the walking trail (which were intended to capture community level trail use).

Stationary Observations

The original SOPARC tool specifies procedures for coding trail use on a walking/jogging track (McKenzie and Cohen, 2007). These observations are to be made from a standard pre-determined location (i.e., the coding station) from which all observational scans are conducted. For this study, the purpose of the stationary observations was to capture participation in organized walking groups (a PATH intervention component in both communities). To accomplish this, the coding station was located on the PATH walking trails at a location within close proximity to the community center to ensure that participants in scheduled PATH walks would pass the coding station. Stationary observations took place in the early morning and evening to coincide with the times during which the organized PATH walks occurred (6:00 am – 8:30 am; 6:00 pm – 7:30 pm).

During stationary observations, all trail users on the walking trail who passed by the coding station were coded for activity level and characteristics. Similar to the original SOPARC tool, individuals were coded each time they passed the coding station. When trail users were observed twice during a stationary observation (e.g., if they made multiple laps around the trail and passed the coding station more than once) they were only counted as one observation. Coders indicated individuals who were observed twice during a given stationary observation on the data collection form to aid in data analyses.

Mobile Observations

Trail use was also assessed for each segment of the trails via mobile observations (e.g., Reed et al., 2011). These observations were added given that stationary observations were unable to capture use of the trail as a whole since community members could have accessed the trail at multiple points (e.g., intersections and sidewalks), and not just the specified starting points at the community centers. These mobile observations allowed trail use to be assessed at many points at which community members could access the trail since they required coders to walk the trail and code community-level trail use for each and every trail segment.

For SOPARC PATH mobile observations, coders started at the community center (i.e., the starting point of the trail) and remained there for 1.5 minutes. During this time, coders recorded *any person* they observed who set foot on the trail or who were observed to be sedentary on the trail (e.g., being pushed in a stroller, sitting on a bench with their feet touching the trail). After the 1.5 minutes at the community center had elapsed, the coders began walking to the next segment (all coders walked in the same direction along the trail as did the organized PATH walking groups). While walking to the next segment the coders continued to code any persons in front of them observed setting foot on the trail or observed to be sedentary on the trail. When coders arrived at the next trail segment they again remained at the beginning of that segment for 1.5 minutes coding trail use, after which they walked the segment coding trail use until they reached that segment's end. When trail users were observed on multiple segments (e.g., if a given trail user was seen on more than one segment), coders indicated so on the data collection form. These trail users were only counted once in the data analyses even though they were observed on multiple segments of the trail. This systematic process continued until the coders had walked the entire trail which ended at the community center.

Mobile observations took place at four specified observation intervals: 1) 7:30 am, 2) 11:30am, 3) 2:00 pm, and 4) 4:30 pm. These times were selected as they are the same intervals utilized in the original SOPARC protocol. These predetermined intervals are based in momentary time sampling techniques and allow for periodic snapshots of community-level walking trail usage over different time periods throughout the day.

SOPARC PATH Coding Conventions

Coding conventions are a prominent feature of systematic observation and should be based on a particular setting and context (McKenzie, 2002). Such conventions were especially important since the SOPARC tool requires coders to code each variable for each observed person even if the coder was not sure what to record. For example, activity level represents the intensity of PA, and codes for this variable were based on energy expenditure. More specifically, trail users were coded as walking if their energy expenditure appeared to be within the range of shifting weight from one foot to another up to walking at a normal pace. Any activity that appeared to expend more energy than walking at a normal pace (e.g., power walking, walking using weights, riding a bike or skateboarding) was coded as vigorous. Another convention was implemented for ethnicity. Because both communities were comprised of a predominately African American population, the default code for ethnicity was African American. For example, if a coder observed a trail user who appeared to be biracial, the coder would have coded “African American” since they cannot be sure of the individual's race unless they asked that person.

SOPARC PATH Logistical Considerations

Data collected via the SOPARC PATH protocol was recorded with pencil-and-paper versions of the SOPARC coding form (McKenzie and Cohen, 2007). One SOPARC coding form was completed for each trail segment, and these forms were customized from the forms that were published with the original tool. The form is a series of rows and columns, with a row for each person observed using the trail and a column for each of the variables of interest (e.g., gender, age category, ethnicity, activity level). Once data were recorded on these pencil-and-paper forms, they were then transferred onto PDAs during down time in between observations.

For each interval – both stationary and mobile – coders filled out a separate cover sheet. This cover sheet indicated the date of observation, name of the observer, which community they were observing, the weather and temperature, whether the trail was well lit, whether the coder felt uncomfortable at any point during the observation, whether an organized walking group was observed, whether police support was seen accompanying the organized walking group, and the number of stray dogs observed. More pertinent to the stationary observations were the items concerning whether an organized walking group was observed, and whether there appeared to be a walking leader accompanying the organized walking group (often identified by the presence of a required red first aid kit and walkie-talkie, both for safety purposes). Trail team members used stopwatches to accurately time observations. Another logistical consideration was that data were collected in two shifts in order to minimize coder fatigue.

Assessment of Environmental Features with Modified PEAT Tool

The PEAT tool was used to measure factors of the physical environment thought to influence trail use. It was used at multiple time points in the study to conduct environmental trail audits to assess changes in features of the trail environment over time. The items on the PEAT tool relate to design features and functionality (e.g., whether the path is adjacent to a road, condition of the path surface, slope of the path), amenities (e.g., presence of lighting, garbage cans, commercial or cultural institutions), trail maintenance (e.g., presence of litter and graffiti), and characteristics of any intersections (e.g., safety features like pedestrian crossing signals and curb cuts). The details of the development and of the PEAT have been previously published, and the tool has been found to have acceptable levels of reliability for most of its primary items (Troped et al., 2006). A subset of its items demonstrated adequate validity via global positioning system (GPS)-derived measures (Troped et al., 2006).

PEAT items are scaled either as dichotomous (yes/no) or on an ordinal scale (e.g., 1=very poor condition, 2=poor condition, 3=moderate condition, 4=good condition, 5= excellent condition). Data were collected over four days, with two days spent in each of the two PATH intervention communities. On a given data collection day, four trail team members audited the trail in pairs. Each pair audited half of the trail. Approaching data collection in this way proved to be an efficient method for collecting data for every trail segment in only one day. Trail auditing pairs audited the trail by entering their codes directly into an electronic data collection form via a PDA. The PEAT protocol encourages coders to re-walk the trail segments as needed to ensure accuracy. A separate electronic data collection form was completed for each trail segment.

Modifications to the PEAT protocol (PATH PEAT)

All items contained in the PEAT manual were retained for the PATH PEAT; however, three dichotomous items (i.e., yes/no) were added to the PATH PEAT based on guidance provided by one of the consultants on the project and feedback from community members at baseline. The additional sub-items assess specific attributes of the surface of the path which may impede or limit travel, including 1) presence of overgrowth from vegetation (e.g., shrubs or bushes), 2) presence of a man-made structure (e.g., fire hydrants, power transformers) or 3) presence of debris (e.g., pinecones).

Additionally, coding conventions were developed for several of the PEAT items (see Table 1). The purpose of the coding conventions was to adapt the instrument for use on urban walking trails in underserved communities. The coding conventions helped to increase consistency of the ratings by providing guidelines for the specified items. Coding conventions were implemented for items that were less straightforward, such as the distinction between “shoulders” and “buffers”. In other instances, coding conventions helped to make the protocol better suited for urban trail use, such as adding in additional types of temporary barriers that could be encountered.

Recruitment & Interview Process

A variety of methods were employed to recruit trail team members for both SOPARC and PEAT data collection. Specific university departments (psychology, exercise science, sociology, social work, public health, and geography) were contacted via respective administrative assistants to request that they forward a recruitment email and flyer to students and faculty.

SOPARC & PEAT Training

For each time point (baseline, 12-month follow up, 18-month follow up), SOPARC training was conducted before PEAT training. This was decided because it was desired that trail team candidates practice the SOPARC protocol in intervention communities without first being exposed to the PEAT tool. It was important to practice the SOPARC in the participating communities so the trail team candidates would be acquainted with the context of these communities and the different visual cues on the trail that can help ensure accurate coding of trail use (e.g., cues which help identify the end-point of particularly long trail segments).

Training for SOPARC consisted of two phases: didactic training and hands-on practice. The first day of SOPARC training involved an overview of the project, the certification process, the data collection calendar, the SOPARC PATH protocol, and other study measures. Applicants viewed slides which showed individuals and groups engaging in various levels of PA or sedentary behaviors to facilitate a discussion around strategies for accurately rating age category, ethnicity, and gender with the SOPARC PATH tool. Hands-on practice at a pre-determined location was also part of this training so candidates could practice their data collection skills and receive corrective feedback.

The second day of SOPARC training included a brief review of the previous day's training, as well as practice using the SOPARC tool in both PATH intervention communities.

Candidates first met community center staff and police officers. Then, they became acquainted with the trail and the neighborhood context through walking the entire trail with a training facilitator and practicing coding trail use with the SOPARC PATH tool as a group. One additional day of practice in each intervention community was conducted. Such an approach is in line with recent recommendations for a combination of classroom training and training sessions in the field where teams can practice the protocol and discuss results with a training facilitator (Brownson et al., 2009). Applicants practiced both stationary and mobile observations and received corrective feedback from the training facilitator.

Training for the PATH PEAT consisted of both didactic and behaviorally-based activities. The first day of training included didactic presentations focusing on the purpose of the tool, basics about its use, and an item-by-item discussion during which the coding conventions were discussed. At the end of the first day the candidates and training facilitators travelled to a local recreational walking trail to practice auditing the trail using the PATH PEAT. Segments used for training were pre-selected as they contained various trail conditions, amenities, signage, and other aspects of the protocol.

During the second day of PEAT training the candidates and training facilitators travelled to a neighborhood similar to PATH intervention communities. The neighborhood was specifically selected as a preferred training site since it had a similar mix of trail features (e.g., slope), amenities (e.g., public transit stops), and destinations (e.g., churches and schools). In addition, this neighborhood had various trail conditions that were to be coded on the PATH PEAT tool and therefore exposed the candidates to a wide variety of items. Training on this day consisted of auditing the trail with the PATH PEAT, and after each trail segment the group discussed their codes and was provided feedback by the facilitator.

Trail Team Certification

All applicants completed two certification tasks for the SOPARC PATH: the coding of PowerPoint slides and the coding of DVD segments. Slides contained static pictures of individuals engaging in PA. Candidates coded the following: gender, age, and ethnicity categories. Candidates' scores were compared to a criterion measure which was established by consultants on the project. Candidates needed to earn scores that were within at least 80% agreement or higher with the criterion. An "agreement" was defined as a match between the candidate's codes and the criterion codes for age category, ethnicity, and gender.

For the DVD component, candidates rated clips presenting individuals engaging in PA from "Systematic Observation: SOPLAY/SOPARC" (McKenzie, 2005). They were asked only to code the activity level (sedentary, walking, or vigorous) of the individuals in the clips. The activity level codes provided by the candidates were compared to a criterion measure which was as established by project consultants. Since the activity level data are at the ordinal level of measurement, Spearman's Rho was used as a reliability statistic (Cohen et al., 2003); candidates needed to reach $\rho > .80$ to meet the reliability standard.

Certification for the PEAT required the candidates to audit four segments in a community near the university campus. While the segments selected for certification were not in the PATH intervention communities, the segments had features that were similar to what would be encountered in the intervention communities. Applicants met reliability standards if their codes were in an acceptable range of agreement with a criterion measure established by project consultants (i.e., Kappa and intraclass correlation values above .80).

After candidates demonstrated adequate reliability for either the SOPARC, PEAT, or both of these measures, they were invited to participate in Phase II of certification. During this second phase, applicants met with the study coordinator for one hour to review the following: issues related to safety, what to do in the case of inclement weather, and other problem solving vignettes that facilitated discussion and learning around behaving respectfully in the intervention communities and how to interact with police support. This meeting helped ensure that candidates understood expectations and recognized the importance of the established safety protocols.

Data analyses

Baseline data were used for analyses, and these data were collected during October, November, and December 2008. Inter-observer reliability was assessed at the individual item level. For SOPARC PATH data, the total number of walkers was summed, along with the total number of walkers who were coded for gender, age category, ethnicity, and their level of activity. These sums were calculated for each trail segment. Having these sums for each segment allowed for a comparison of segment totals between coding partners. For PATH PEAT data, agreement between coders was assessed for each item on each individual trail segment, and each of the codes for every PATH PEAT item was compared between raters.

The intraclass correlation (ICC) was chosen as a metric of inter-observer reliability for ordinal (e.g., Likert-type PATH PEAT items) or continuous (e.g., frequency of SOPARC PATH variables) responses. ICCs were chosen for such items because this statistic is sensitive to both association and absolute differences between coders (Müller and Büttner, 1994; Shrout and Fleiss, 1979). Two-way random effects single rater ICCs were calculated for each of the continuous SOPARC PATH variables and ordinal PATH PEAT items given that coders were considered a random sample of all potential and future coders (Müller and Büttner, 1994; Shrout and Fleiss, 1979). The single measure ICC was selected because the unit of analysis is the individual coder who collected the data, rather than an average across coders. Inter-observer reliability for items with dichotomous values (yes/no) was assessed using the Kappa statistic, calculated via a contingency table. For all reliability analyses, values were interpreted as "poor" (< 0.00), "slight" (0.00 to 0.20), "fair" (0.21 to 0.40), "moderate" (0.41 to 0.60), "substantial" (0.61 to 0.80) and "almost perfect" (0.81 to 1.00) based on ranges described by Landis and Koch (1977). Although these divisions are arbitrary, they are conventional benchmarks used to qualitatively describe the strength of agreement between coders.

One issue with ICC and Kappa values may be encountered when there is little or no variability in codes by one or both raters. When there is little or no variability ICC and Kappa values can be misclassified as low. In extreme cases of low variability (e.g., when a given variable is a constant that has the same code for every instance), Kappas and ICCs cannot be computed by statistical software programs. In light of the possibility that values can be misclassified as low – or not computed – percent agreement was calculated as an additional measure of inter-observer reliability. Percent agreement ((number of agreements) / (number of agreements + number of disagreements)) is relatively straightforward to compute, but it can be considered a stringent statistic since a disagreement that is small in magnitude (e.g., by a value of one) impacts the resultant percentage in the same way as a disagreement that is large in magnitude (e.g., by a value of, say, 50). Binary items with low Kappa values, but a percent agreement ≥ 75%, and ordinal or continuous response items with low ICC values, but a percent agreement ≥ 60%, were deemed as having adequate reliability based on ranges described by Saelens et al. (2006). Descriptive statistics (i.e., frequency distributions, means) were used to summarize the trail for PATH PEAT items. While data for two coders were available for these items, one of these coders was randomly selected via a coin flip and descriptive statistics are based off of only the codes the selected coder provided. All analyses were conducted via SPSS version 18.0.

RESULTS

SOPARC PATH

Inter-observer reliability data for both communities indicate a high level of agreement (.98, p < .01). Table 2 displays ICCs for each variable of interest (e.g., total walkers, total number of males, total number of females, etc.) in both communities[i]. A total of 83% of these ICCs were in the almost perfect range, 8% were in the substantial range, 8% were in the moderate range, and 0% was in the fair range. While the vast majority of items were in the almost perfect to substantial range (1.00 to .61), the item with the lowest reliability that fell in the moderate range was the activity level-related variable “sedentary” (ICC = .42, p < .01).

SOPARC PEAT

Table 3 contains descriptive statistics and Kappa coefficients for the binary PEAT items for trails in both PATH intervention communities combined. Each item is listed along with the proportion of segments on which it was observed and coded. For example, 72% of the trail segments on the PATH trails started at an intersecting road, which is indicative of the urban environment in which the trails were identified. While the presence of some safety features are relatively prevalent on these trails (e.g., nearly all of the roads which intersected with the PATH trail had stop signs and curb cuts to enhance pedestrian safety and enable mobility for walkers using assistive devices, respectively), trail amenities such as benches, lighting, and drinking fountains were rarely present. Given the urban nature of the trails, it is not surprising to see that almost half of the trail segments (45%) included commercial destinations that trail users could access.The Kappa coefficients in Table 3 show reliability for the PEAT items ranged from .77 to 1.00. These values fall in the substantial to almost perfect range. The provided percent agreement values should be used as an estimate of reliability for the 14 out of 39 Kappa values that could not be computed due to low variability. For these low-variability PEAT items, percent agreement ranged from 92.8% to 100% agreement.

Descriptive statistics and reliability data for ordinal PEAT items can be found in Table 4. ICCs ranged from .34 to 1.00, and these values fall within the fair to almost perfect range. For the three items for which ICCs were not able to be computed (likely due to either low variability or low base rates), the percent agreement values provide a more appropriate assessment of inter-observer agreement. Percent agreement for these three items ranges from 98.6% to 100% agreement.

Table 4 indicates that the condition of the path surface and the safety of intersections on these urban trails are both between Fair and Poor – based on the metrics presented in the PEAT Manual. These data also indicate that the trails are relatively flat and that there is virtually no slope or cross-slope. In regard to trail maintenance, the trails had “a little” glass, litter, and noise, and data indicate that there was virtually no graffiti, vandalism, or animal droppings observed on days during which data were collected.

DISCUSSION

Much progress has been made in the ability to assess the extent to which community members use walking trails for PA, as well as the physical environmental qualities of these trails. In the spirit of continuing this progress, existing measures need to be improved through a process in which their utility, feasibility, and appropriateness for use in diverse

[i]Inter-observer reliability data specific to each community are not reported since they are redundant with the aggregated reliability data reported herein.

contexts are investigated. In this article a process has been described that can be used as a model for using observational protocols in underserved communities.

The modified tools have been shown to be reliable methods for a) measuring trail use and b) assessing physical features of the trail in underserved environments. Reliability data related to SOPARC PATH indicated a very high level of agreement among coders (ICC = .98, $p < .01$). Inter-observer reliability for individual SOPARC items was also very high as 10 of the 12 (83.3%) of the ICCs were in the almost perfect range. The ICC for the age category 'Senior' was in the moderate range (.67), and this may be due to coders having difficulty estimating whether a given trail user appeared to be 60 years of age or older vs. whether the trail user was in the upper-end for the 'Adult' age category (21 years old up to 59). The only other ICC that wasn't in the almost perfect range was for sedentary activity level. Although this ICC was lower when compared to all other SOPARC variables, it was considered to be of adequate reliability given that the percent agreement among coders for this variable was 86%. In fact, the low ICC value was likely a misclassification due to low variability (i.e., there were very few instances of sedentary behavior on the trail, and the low base rate of this code resulted in low variability).

Reliabilities for measuring features of the trail were also adequate, as these data were found to range from fair to almost perfect (κ = .77 – 1.00; ICC = .34 – 1.00). All but three of the binary items on the PEAT were reliably coded in the almost perfect range (≥ .81). The three binary items that were the exception were close to almost perfect but fell in the substantial range (.79, .79, and .77) and were all related to amenities that were coded on less than 9% of the trail segments (e.g., accessible telephones, play areas, and other services for trail users). Reliabilities for the ordinal PEAT items were not as high as they were for the binary PEAT items. Only two of the ordinal items had an ICC in the almost perfect range. One of these items was for slope of the trail, and since the trails were essentially flat and did not fluctuate, the high reliability with which this item was coded can be expected. The other item with an almost perfect ICC was related to the width of a buffer (an expanse that provides protection to trail users from an adjacent road). This item was also relatively straightforward and coders used a tape measure to determine width. Most of the ordinal items are subjective, which makes them more difficult to code. For example, indicating whether or not a garbage can is present (binary item requiring coder to indicate yes or no) is much more straightforward than rating the condition of the path surface (ordinal item on a 5-point scale). Such ordinal items require the coder to review the PEAT manual carefully when collecting data and take into account the entirety of the segment, and at times it can be difficult to select between two possible codes that are both marginally appropriate. Most of the ordinal PEAT items were in the substantial to fair range, and the resulting values are similar to those found in the study that validated and established the reliability of the PEAT. For example, two items that were considered very important to communities who participated in this study – condition of the path surface and safety of intersections – were found to have values very similar to those found in Troped et al. (2006). Extensive training and certification processes should be implemented as modeled in this study as a quality assurance process. The purpose of planning and implementing high quality training is to ensure that a "common lens" is established for all coders that is objective and nonjudgmental, and steps should be taken to support and maintain coders' skills after training to prevent against "coder drift" (McKenzie, 2010).

Modifying the SOPARC and PEAT tools was approached systematically. These modifications were conducted in the spirit of continually improving these measures to ensure their relevance for diverse populations and contexts. Consultants with extensive experience using the protocols (B. E. A., N. C., T. L. M., & J. R.) guided the development of other protocol modifications, including adjustments to how data were collected and establishing coding conventions. For the SOPARC PATH protocol, mobile observations were added to capture use of the trail as a whole since trail users could have accessed the trail at multiple points (e.g., intersections and sidewalks). The mobile protocol complemented the stationary protocols which were better-suited to capture participation in organized walking groups. For the PATH PEAT tool, the modification process involved community members who identified important items regarding trail safety in their community (e.g., overgrowth from vegetation, presence of stray dogs on the walking trail, lack of drug dealers). In addition coding conventions were added to existing PEAT items to enhance relevance to urban walking trails in underserved communities (e.g., adding in additional types of temporary barriers that could be encountered like parked cars and curbside garbage cans) and to enhance the clarity of items that coders found difficult during training and certification (e.g., the distinction between “shoulders” and “buffers”).

While both the SOPARC PATH and PATH PEAT protocols have been modified for use in underserved communities, their use in other underserved communities may reveal additional adaptations that may be needed to ensure relevance and accurate coding. It is suggested that any further modifications to these protocols or similar tools be done so with both community members within the communities of focus along with developers and/or content experts who have a deep understanding of the measures that are to be modified.

While these community- and expert-guided modifications are thought to strengthen the original protocols by enhancing their relevancy to the communities of focus, the modifications may not be appropriate for other contexts and populations that differ from those observed in this study. It is also important to note that these observational protocols require a major time commitment to conduct appropriately. Training for these protocols is often conducted over multiple days and involves didactic activities (e.g., operational definitions), behavioral skill- based activities (e.g., practicing systematic observation in diverse settings and receiving feedback on performance), and an assessment of trainees’ level of skill (e.g., McKenzie et al., 2006).

Despite all this certification and careful planning and coordination among the study coordinators, community members, police officers providing oversight, and the Trail Team who collected the data, inclement weather often disrupted data collection days such that they had to be repeated. These days were repeated given that inclement weather has been established as an environmental barrier to PA (e.g., Tucker and Gilliland, 2007) and we assumed that estimates of PA on these days would have been biased due to such weather. Being at the mercy of the weather is an unavoidable drawback of randomly selecting days to observe data on walking trails. That said, we acknowledge that this is a limitation of doing direct observations, and that collecting data only during clement weather may present the “best-case scenario” for trail use. The protocols described herein are based on published observational tools that have been used in multiple studies of park and trail use.

An additional limitation is that the interpretive ranges used for percent agreement statistics in this study (percent agreement ≥ 75% for binary items and ordinal or continuous response items with percent agreement ≥ 60% were deemed as having adequate reliability) are arbitrary and based off of a single study (Saelens et al., 2006). These arbitrary interpretive ranges are not as well established as those suggested by Landis and Koch (1977) for ICC and Kappa statistics. We feel the guidelines suggested by the Saelens et al. study are appropriate given its similarities with the present study.

This study provides a model for applying modified SOPARC and PEAT protocols to be used in an underserved community. The instruments were useful and able to be applied with a high level of reliability and their content are appropriate for assessing trail use and features of the trail environment in these communities. Future directions for this work include assessing reliability data for the 12- and 18-month follow up observations to determine the extent to which trail use changes over time, the extent to which features of the trail change over time, and the degree to which these constructs are related to the PATH project's intervention. While further modification of these protocols may be warranted, these initial steps have helped establish currently available tools for use in underserved communities.

References

Brownson RC, Hoehner CM, Day K, Forsyth A, Sallis JF. Measuring the built environment for physical activity: State of the science. American Journal of Preventive Medicine. 2009; 36:S99–S123. [PubMed: 19285216]

Cohen D, Sehgal A, Williamson S, Golinelli D, Lurie N, McKenzie TL, Capone-Newton P. Impact of a new bicycle path on physical activity. Preventive Medicine. 2008; 46:80–81. [PubMed: 17707495]

Cohen, J.; Cohen, P.; West, SG.; Aiken, LS. Applied multiple regression/correlation analysis for the social sciences. 3rd ed.. Lawrence Erlbaum Associates, Publishers; Mahwah, NJ: 2003.

DeNavas-Walt, C.; Proctor, BD.; Smith, JC. Income, Poverty, and Health Insurance Coverage in the United States: 2009, Current Population Reports. U.S. Government Printing Office; Washington, DC: 2010. p. P60-238.

Floyd MF, Taylor WC, Whitt-Glover M. Measurement of park and recreation environments that support physical activity in low-income communities of color: Highlights of challenges and recommendations. American Journal of Preventive Medicine. 2009; 36:S156–S160. [PubMed: 19285207]

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

McKenzie, TL. Use of direct observation to assess physical activity. In: Welk, GJ., editor. Physical Activity Assessments for Health-related Research. Human Kinetics Publisher, Inc.; Champaign, IL: 2002. p. 179-195.

McKenzie, TL. Systematic Observation: SOPLAY/SOPARC Introduction, Practice, and Assessment. San Diego State University; San Diego, CA: 2005.

McKenzie TL. 2009 C. H. McCloy Lecture. Seeing is believing: observing physical activity and its contexts. Res Q Exerc Sport. 2010; 81:113–122. [PubMed: 20527295]

McKenzie, TL.; Cohen, DA. System for observing play and recreation in communities (SOPARC) description and procedures manual. 2007.

McKenzie TL, Cohen DA, Sehgal A, Williamson S, Golinelli D. System for Observing Play and Recreation in Communities (SOPARC): Reliability and feasibility measures. Journal of Physical Activity & Health. 2006; 3:S208–S222. [PubMed: 20976027]

McKenzie TL, Marshall SJ, Sallis JF, Conway TL. Leisure-time physical activity in school environments: an observation study using SOPLAY. Preventive Medicine. 2000; 30:70–77. [PubMed: 10642462]

Müller R, Büttner P. A critical discussion of intraclass correlation coefficients. Statistics in Medicine. 1994; 13:2465–2476. [PubMed: 7701147]

Office of Research and Statistics. South Carolina Statistical Abstract. 2006.

Reed JA, Hooker SP, Muthukrishnan S, Hutto B. User demographics and physical activity behaviors on a newly constructed urban rail/trail conversion. Journal of Physical Activity & Health. 2011; 8:534–542. [PubMed: 21597126]

Saelens BE, Frank LD, Auffrey C, Whitaker RC, Burdette HL, Colabianchi N. Measuring Physical Environments of Parks and Playgrounds: EAPRS Instrument Development and Inter-Rater Reliability. Journal of Physical Activity and Health. 2006; 3:S190–S207.

Sallis, JF.; Owen, N. Physical activity and behavioral medicine. Sage Publications; Thousand Oaks, CA: 1999.

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979; 86:420–428. [PubMed: 18839484]

Tester J, Baker R. Making the playfields even: Evaluating the impact of an environmental intervention on park use and physical activity. Preventive Medicine. 2009; 48:316–320. [PubMed: 19463491]

Troped PJ, Cromley EK, Fragala MS, Melly SJ, Hasbrouck HH, Gortmaker SL, Brownson RC. Development and reliability and validity testing of an audit tool for trail/path characteristics: The Path Environment Audit Tool (PEAT). Journal of Physical Activity & Health. 2006; 3:S158–S175.

Tucker P, Gilliland J. The effect of season and weather on physical activity: A systematic review. Public Health. 2007; 121:909–922. [PubMed: 17920646]

United States Census Bureau. State & County QuickFacts. 2012.

van der Mars H, Rowe P, Schuldheisz J, Fox S. Measuring students' physical activity levels: Validating SOFIT for use with high school students. Journal of Teaching Physical Education. 2004; 23:235–251.

Wilson DK, Trumpeter NN, St. George SM, Coulon SM, Griffin S, Van Horn ML, Lawman HG, Wandersman A, Egan B, Forthofer M, Goodlett BD, Kitzman-Ulrich H, Gadson B. An overview of the “Positive Action for Today's Health” (PATH) trial for increasing walking in low income, ethnic minority communities. Contemporary Clinical Trials. 2010; 31:624–633. [PubMed: 20801233]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Highlights

- Modified two systematic observation protocols for use in underserved communities.
- Modified protocols assessed use of a walking trail and their physical features.
- Reliability for measuring trail use was found to be high (ICC = .98, $p < .01$).
- Reliability for measuring trail features was κ = .77 – 1.00; ICC = .34 – 1.00.
- Modified protocols were shown to be reliable for use in underserved communities.

Table 1

Coding Conventions for the PATH PEAT

Item	Coding Convention
Access Points	Only include the community center as an official access point related to the organized walking program (component of the PATH intervention).
Animal Droppings	Bird droppings included.
Buffer Along Adjacent Road	Buffer is defined as “something used to protect a trail user from an adjacent road”. This could be a strip of grass or other expanse – or a physical barrier like a concrete wall – that separates the trail from a road. The identifying feature for a buffer is that it serves to protect trail users whether or not it was designed for safety purposes. “Buffers” enhance safety while “shoulders” (described below) provide a means to pass others (which enhances trail functionality).
Commercial Destinations	Only code if on the segment of interest; Only code once (do not code if located on an adjacent segment); If the destination is on the segment of interest but is directly across the street only code if pedestrians could easily access the destination (e.g., they didn't have to cross multiple lanes of busy traffic) *and* it is reasonable that pedestrians would utilize the trail to access such a site (e.g., retail stores, restaurants).
Condition of path surface	Overgrowth from vegetation is not considered part of path surface (there is a separate item on the PATH PEAT for overgrowth).
Cultural or civic institutions	Only code if on the segment of interest; Only code once (do not code if located on an adjacent segment); If the institution is on the segment of interest but is directly across the street only code if pedestrians could easily access the institution (e.g., they didn't have to cross multiple lanes of busy traffic) *and* it is reasonable that a pedestrians would utilize the trail to access such a site (e.g., churches, schools).
Exercise or play equipment	Only code if on the segment of interest; Only code once (do not code if located on an adjacent segment); If the equipment is on the segment of interest but is directly across the street only code if pedestrians could easily access the equipment (e.g., they didn't have to cross multiple lanes of busy traffic) *and* it is reasonable that pedestrians would utilize the trail to access the equipment (e.g., parks that are accessible to the public and not behind locked fences).
Lighting	Only consider lighting on the segment of interest (do not include lighting across the street). Coder does not need to determine if light actually functions properly as this is assessed via a separate lighting assessment conducted in the evening.
Presence of Shoulder	A shoulder is defined as “an intentionally cleared area used to pass others”. Shoulders could be a strip of grass, gravel, or other expanse that allows trail users to safely pass others on the trail.
Temporary Barrier	Can include cars and garbage cans that impede travel on path.

Table 2

Inter-observer reliability for all baseline SOPARC observations in both intervention communities separated by item

		95% Confidence Limit			
Item	**ICC**	**Lower Bound**	**Upper Bound**	**df**	**% Agmnt**
Total walkers	.99*	.99	1.00	49[a]	74.0%
Female	.83*	.72	.90	49[a]	72.0%
Male	.89*	.81	.93	49[a]	55.2%
Child	.95*	.92	.97	49[a]	92.0%
Teen	.88*	.80	.93	49[a]	62.0%
Adult	.92*	.86	.95	49[a]	22.0%
Senior	.67*	.49	.80	49[a]	60.0%
Black	1.00*	.99	1.00	49[a]	74.0%
Other ethnicity	.92*	.86	.95	49[a]	94.0%
Sedentary	.42*	.17	.62	49[a]	86.0%
Walking	.99*	.98	1.00	49[a]	68.0%
Vigorous	.98*	.96	.99	49[a]	84.0%

ICC = Intraclass correlation; df = degrees of freedom (n -1); % Agmnt = Percent agreement (number of agreements – (number of agreements + number of disagreements))

* $p < .01$

[a] Two cases were excluded from the analysis via listwise deletion due to only one coder being available for two out of the four observation intervals during one observation date in Community 1; Four cases were excluded from the analysis via listwise deletion due to only one coder being available being available for all four observation intervals during one observation date in Community 2.

Table 3

Inter-observer reliability for binary measures in PEAT based on observations in both PATH intervention communities

Item[a]	Segments[b]	Proportion of segments with attribute – 1st observer	df[c]	Kappa	% Agmnt
Intersecting road	40	72%	68	1.00 *	100%
Presence of stop sign or traffic signal	29	97%	50	1.00 *	100%
Presence of curb cut at the intersection	29	97%	50	1.00 *	100%
Presence of crosswalk	29	7%	50	1.00 *	100%
Crosswalk is raised	29	0%	50	No kappa[e]	98.4%
Presence of pedestrian crossing signal	29	7%	50	1.00 *	100%
Portion of the path under repair	40	0%	68	No kappa[e]	100%
Presence of temporary barrier on path	40	3%	68	1.00 *	100%
Vegetation overgrowth impedes/limits travel	40	53%	68	0.86 *	92.8%
Man-made structure impedes/limits travel	40	13%	68	0.86 *	97.1%
Debris impedes/limits travel	40	23%	51[d]	0.82 *	94.2%
Sufficient vertical clearance	40	62%	68	0.97 *	98.6
Presence of a shoulder	40	0%	68	-	100%
Road parallel to trail	40	100%	68	No kappa[e]	97.1%
Presence of a buffer from the adjacent road	24	58%	68	1.00 *	100%
Presence of official access point	40	5%	68	1.00 *	100%
Presence of gates or bollards	40	0%	68	No kappa[e]	100%
Deliberate viewpoints	40	0%	68	No kappa[e]	100%
Lighting on trail	40	50%	68	0.88 *	94.2%
Accessible telephones	40	8%	68	0.77 *	97.1%
Presence of emergency call boxes	40	0%	68	No kappa[e]	100%
Accessible restrooms	40	5%	68	0.85 *	98.6%

Item[a]	Segments[b]	Proportion of segments with attribute – 1st observer	df^c	Kappa	% Agmnt
Accessible benches	40	3%	68	1.00*	100%
Accessible picnic tables	40	0%	68	No kappa[e]	100%
Accessible drinking fountains	40	5%	68	1.00*	100%
Accessible garbage cans	40	8%	68	1.00*	100%
Signage	40	0%	68	No kappa[e]	100%
Adjacent parking lot	40	8%	68	1.00*	100%
Accessible bicycle racks	40	3%	68	1.00*	100%
Accessible exercise/play areas	40	8%	68	0.79*	97.1%
Services for trail users	40	3%	68	0.79*	98.6%
Mass transit stops along trail	40	0%	68	No kappa[e]	100%
Presence of cultural or civic institution(s)	40	25%	68	1.00*	100%
Cultural or civic institution is a community facility	10	30%	68	No kappa[e]	100%
Cultural or civic institution is a cultural institution	10	10%	68	No kappa[e]	100%
Cultural or civic institution is a municipal building	10	10%	68	No kappa[e]	100%
Cultural or civic institution is a secular/religions building	10	50%	68	No kappa[e]	100%
Presence of commercial destination(s)	40	45%	68	0.91*	97.1%
Stray dogs	40	8%	68	0.92*	98.6%

df = (number of cases – 1); % Agmnt = Percent agreement (number of agreements – (number of agreements + number of disagreements)

[a]Items that were not coded on at least 10 segments of the trail are not included as these estimates were considered to be unstable.

[b]Community 1 has 17 segments, Community 2 has 23, which adds to a total of 40 segments among both trails.

[c]Degrees of freedom for analyses differs from total number of segments. During one observation day in Community 2 only one observer was available for 11 segments.

[d]Degrees of freedom for analyses differ from total number of segments due to missing data for this item on first data collection date in Community 1.

[e]Kappa could not be calculated due to low variability

*p < .01

Table 4

Inter-observer reliability for ordered categorical measures in PEAT based on observations in both PATH intervention communities.

Item	Segments[a]	Mean rating – 1st observer	*df*[b]	Lower bound	Upper bound	ICC	% Agmnt
Condition of the path surface	40	2.88	68	0.37	0.70	0.56*	65.2%
Safety of intersection	40	2.76	50	0.38	0.74	0.59*	49%
Slope of trail	40	1.08	68	1.00	1.00	1.00	100%
Cross slope of trail	40	1.00	68	-	-	No ICC[c]	100%
Sight distance	40	1.13	68	0.53	0.79	0.68*	92.8%
Width of buffer	24	1.33	68	0.88	0.95	0.93*	100%
Lateral visibility	40	1.97	68	0.16	0.57	0.38*	95.7%
Amount of glass	40	2.00	68	0.60	0.83	0.73*	63.8%
Amount of litter	40	2.33	68	0.34	0.68	0.53*	65.2%
Amount of graffiti	40	1.02	68	0.51	0.78	0.66*	98.6%
Amount of vandalism	40	1.00	68	-	-	No ICC[c]	98.6%
Amount of odor	40	1.43	68	0.11	0.53	0.34*	55.1%
Amount of noise	40	2.65	68	0.40	0.72	0.58*	59.4%
Amount of animal droppings	40	1.00	68	-	-	No ICC[c]	100%

df = (number of cases – 1); ICC = Intra-class correlation; % Agmnt = Percent agreement (number of agreements – (number of agreements + number of disagreements)

[a]Community 1 has 17 segments, Community 2 has 23, which adds to a total of 40 segments among both trails.

[b]Degrees of freedom for analyses differs from total number of segments. During one observation day in Community 2 only one observer was available for 11 segments.

[c]ICC was unable to be calculated due to low variability.

*$p < .01$