Journal of Physical Activity and Health 2006, 3, Suppl 1, S158-S175

Development and Reliability and Validity Testing of an Audit Tool for Trail/Path Characteristics: The Path Environment Audit Tool (PEAT)

Philip J. Troped, Ellen K. Cromley, Maren S. Fragala, Steven J. Melly, Hope H. Hasbrouck, Steven L. Gortmaker, and Ross C. Brownson

Background: To determine how trail characteristics may influence use, reliable and valid audit tools are needed. *Methods:* The Path Environment Audit Tool (PEAT) was developed with design, amenity, and aesthetics/maintenance items. Two observers independently audited 185 trail segments at 6 Massachusetts facilities. GPS-derived items were used as a "gold standard." Kappa (k) statistics, observed agreement and intraclass correlation coefficients (ICCs) were calculated to assess inter-observer reliability and validity. *Results:* Fifteen of 16 primary amenity items had k-values ≥ 0.49 ("moderate") and all had observed agreement ≥ 81%. Seven binary design items had k-values ranging from 0.19 to 0.71 and three of 5 ordinal items had ICCs ≥ 0.52. Only two aesthetics/maintenance items (n = 7) had moderate ICCs. Observed agreement between PEAT and GPS items was ≥ 0.77; k-values were ≥ 0.57 for 7 out of 10 comparisons. *Conclusions:* PEAT has acceptable reliability for most of its primary items and appears ready for use by researchers and practitioners.

Key Words: audit, environment, trails, physical activity

The influence of the built environment, which includes places such as homes, schools, worksites, parks and recreational spaces, business and industrial areas, roads and highways,[1] is receiving increased attention in physical activity determinants research. Since the early to mid-1990s,[2, 3] there has been increasing interest in studies examining relationships between characteristics of the neighborhood physical environment and participation in physical activity. As recent reviews have demonstrated, there is evidence from public health[4] and the transportation and

Troped, Fragala, Melly, and Gortmaker are with the Prevention Research Center and Dept of Society, Human Development and Health, Harvard School of Public Health, Boston, MA 02115. Cromley is with the Dept of Geography, University of Connecticut, Storrs, CT 06269. Hasbrouck is with the School of Architecture, University of Texas at Austin, Austin, TX 78712. Brownson is with the Prevention Research Center and Dept of Community Health, School of Public Health, Saint Louis University School of Public Health, St. Louis, MO 63104.

planning disciplines[5, 6] that attributes of the built environment, such as access to facilities and land-use mix, are correlated with physical activity.

Thus far, a limitation of this research has been the quality of the environmental measures. Most studies have used individual perceived measures of the neighborhood environment; few used objective measures (e.g., using geographic information systems (GIS))[7] and often psychometric properties have not been assessed.[8] Recently, researchers have been working to develop reliable and valid tools to measure characteristics of the neighborhood built environment.[9-13] Studies in Australia [13] and the US[10, 11] have tested the reliability of audit tools that are used to assess characteristics of the pedestrian and cycling environment in neighborhoods.

However, to our knowledge, this important measurement work has largely excluded a substantial component of the built environment that may influence physical activity—namely, parks, community trails, and other outdoor recreational areas. Community trails/paths comprise a sub-set of outdoor facilities studied by public health researchers in recent years.[14-16] This growing attention is warranted given that trails/paths can provide free, equitable opportunities for physical activity among all members of a community and the popularity of trails for walking, bicycling, and other activities is increasing (accessed January 24, 2005: http://www.railtrails.org). To aid in the development of reliable measures of trail characteristics, this study had three objectives: 1) develop and test the inter-observer reliability of the Path Environment Audit Tool (PEAT); 2) briefly illustrate within-site variation in trail attributes; and 3) assess the validity of a sub-set of PEAT items using global positioning system (GPS)-derived measures.

Methods

Path Environment Audit Tool (PEAT): Development and Components

Our aim was to use an evidence-based approach to selecting specific measures for inclusion in PEAT that involved three specific strategies. First, we conducted a brief review of the parks and recreation, landscape architecture, transportation, and planning literatures to identify environmental characteristics of trails that have been documented to influence use. This review included peer-reviewed literature, as well as trail design manuals. We obtained lists of references and suggestions for search engines from our consultants and investigators, and searched websites such as Google and the Web of Science using key words that included "parks and recreation," "user preferences," "trail," "recreation," "physical activity," "walking," and "bicycle." Due to resource constraints, we did not include "fugitive literature" in our review. Approximately 35 articles, design guides, and books were retrieved and 28 were deemed to have relevant content and were subsequently reviewed. Based on this literature, we created a summary table, in which we listed any factor that was either empirically studied or received prominent attention in the document. Overall, we observed that the evidence of relationships between specific trail characteristics and trail use was limited to non-existent. However, we identified 90 different factors that potentially influence use, which we organized into five domains: 1) design features (further subdivided into *overall structure*, *site* *furnishings or amenities*, *condition/maintenance*, and *access*); 2) overall park or trail characteristics such as size; 3) aesthetics; 4) "human" environmental factors; and 5) "situation" characteristics or neighborhood contextual variables.

As a second strategy, we developed and implemented a brief intercept survey with a convenience sample of 73 adult trail users at two study sites during the spring of 2003. The purpose was to identify specific physical characteristics of sites that influenced use, based on users' perceptions. Two items asked respondents to report three things they liked the "most" and "least" about the trail or path. In terms of things most liked, users most frequently provided responses that related to design characteristics (e.g., trail distance), aesthetics (e.g., scenery), social factors (e.g., not too crowded), convenience (from home, work), and general atmosphere (e.g., quiet).

The third strategy for selecting PEAT items was to solicit input from our transdisciplinary team (e.g., parks and recreation, landscape architecture). A tabular summary of the literature review and trail surveys was sent to this group. Via e-mail and a follow-up conference call, they were asked to indicate items they believed were important to include in the audit tool, including additional items not captured in the summary.

PEAT items were then created within three general content areas that we termed *design features*, *amenities*, and *maintenance/aesthetics*. Most items focused on the immediate environment of the trail/path, although several items assessed the proximal neighborhood environment (e.g., commercial destinations). Items were scaled either as dichotomous (e.g., yes/no for presence of amenities) or on an ordinal scale (e.g., 1 = flat or gentle slope, 2 = moderate slope, 3 = steep slope). PEAT contained 11 trail design features items, some with sub-items. The primary design items were assessments of surface condition, slope, cross-slope, vertical clearance, sight distance, and vegetative cover or enclosure, presence of a shoulder adjacent to the trail, presence of a road adjacent to the trail, presence of official access points, presence of gates or bollards, and presence of viewpoints. Sub-items further assessed attributes of certain design features, for example, width of shoulders, wheelchair accessibility of access points and bollards, and specific types of viewpoints along the trail (e.g., water body).

PEAT also included 16 amenity items, some with sub-items. Trail amenities included presence of lighting, telephones, emergency call boxes, rest rooms, benches, picnic tables, drinking fountains, garbage cans, signage, parking, bike racks, exercise or play areas, services (e.g., food), public transit stations, cultural or civic destinations, and commercial destinations. Typically, sub-items involved a rating of the functioning, condition, and cleanliness of the amenity.

The third major category, maintenance/aesthetics, included seven items that assessed trail conditions, such as the amount of glass, litter, graffiti, vandalism, odor, noise, and dog/animal droppings. These conditions were rated on a 4-point scale: "none," "a little," "some," "a lot." The items were comparable to those originally described by Sampson and Raudenbush[17] and recently used in a neighborhood audit tool.[10] PEAT also included one item that assessed the presence of dogs on the trail.

In addition to the items designed to assess trail segments, PEAT included a 6-item module for roads that intersected trails, a feature that typically occurred at rail-trails or linear parks. Items included presence of traffic signals or stop signs for vehicles, curb cuts, crosswalks, raised crosswalks, and pedestrian crossing signals, and an overall rating of the intersecting road's safety (5-point scale from "very poor" to "excellent").

Although PEAT could be used as a paper and pencil tool, we decided to develop a computer-based instrument for easier field use and more efficient data processing. We developed a data entry form within Microsoft Access and loaded it on a tablet PC, which allowed observers to directly enter PEAT observations into an electronic database.

During December 2003 and January 2004, two members of the research team (PT, MF) pre-tested PEAT. A convenience sample of 43 trail and intersecting road segments were audited at two study sites (Danehy Park and Southwest Corridor). The data were imported into SAS and inter-observer reliability was assessed using Kappa coefficients (k) and intraclass correlation coefficients (ICCs). The team reviewed results item-by-item and attempted to identify factors that may have led to low reliability. Subsequently, items with low inter-observer reliability were either deleted from PEAT or the wording was modified. In some cases, where it appeared that one or both observers had been unclear about how to rate a particular item, operational definitions were discussed and further clarified in writing.

Study Setting

This study was conducted in Massachusetts at six trails/paths that were located in urban (n = 3), suburban (n = 2), and rural (n =1) communities. The primary selection criterion were that the trails: 1) supported walking and secondarily, other linear forms of physical activity, (e.g., bicycling, in-line skating); 2) represented both linear facilities (e.g., rail trail) and trail circuits (e.g., within a park); and 3) were located in urban, suburban, and rural settings. The goal was to select facilities that would have wide variation in design characteristics and amenities. Our research team identified 16 potential sites they had familiarity with and selected six that appeared to have diverse characteristics (see Table 1).

Trail Data Collection with Global Positioning System (GPS) Unit

Prior to auditing trails with PEAT, we searched MassGIS, a statewide clearinghouse for spatial datasets, for GIS trails data for the six study sites and found that data were only available for two rail-trails. We then contacted agencies responsible for the other facilities and found that spatial trail data were not available in a consistent format and accuracy (e.g., in some cases only paper maps existed). Therefore, we decided to collect data at the six sites with a high-accuracy GPS unit, since it would be the most efficient way to obtain consistently accurate trail data. The purposes were to: 1) create a spatial framework for all subsequent data collection with PEAT and integration of other GIS data layers; 2) collect detailed data on amenities and other point attributes along trails that could serve as "gold standard" validation measures for PEAT items; 3) create a spatially accurate geographic unit of observation for PEAT observations (i.e., trail segments); and 4) develop maps of trail segments that would guide PEAT observations.

During the summer of 2003, teams of three to four researchers walked all primary trails at study sites with a Trimble GPS Pathfinder Pro XR receiver and a

Table 1 Select Physical and Socio-demographic* Characteristics of Six Study Sites

Facility	Setting & type of facility	General characteristics	Trail segments observed (n)	Mean (SD) segment length	Family** income (mean)	Pop.*** density (mean)	Percent black (mean)
Cutler Reservation Needham, MA	Suburban conservation land	1.6 mile circular dirt path Highly wooded, water views Next to industrial/ office park	14	229 (115)	103,215	1341	1.8
Danehy Park Cambridge, MA	Urban-suburban neighborhood park	Built over landfill Paved walking & bicycle paths Internal circuits (about 3 miles) Other facilities: baseball & soccer fields	16	226 (146)	73,058	6062	14.5
Franklin Park Boston, MA	Large urban park	500 acre park, racially diverse area Trail circuits, paved and unpaved (primary loop = 2.1 miles) Other facilities: zoo, golf course	29	196 (98)	32,547	7974	60.7
Minuteman Bikeway Arlington to Bedford, MA	Suburban rail-trail	11 miles, asphalt America's "500th Rail Trail" Well-established, high use	56	288 (111)	90,471	2594	1.7
Nashua River Rail Trail (north central MA)	Rural rail-trail	12 miles, asphalt Dirt horse path parallel to trail	48	377 (114)	69,563	432	1.7
Southwest Corridor Boston, MA	Urban linear park with adjacent facilities	4 miles, asphalt Close access to public transit Adjacent facilities: play courts, playgrounds	22	259 (150)	47,304	10,529	26.9

Note. *Based on block groups within 400 meters using 2000 US Census data. **yearly income. ***population per square kilometer.

TSCI Asset surveyor (hand-held unit). The unit was programmed to collect attribute data on trail segments (e.g., type of surface material, trail width) and spatial data on 31 point features (e.g., lighting) using the Quickmark feature of the GPS. One person operated the GPS unit and the other two to three team members assisted by measuring the trail width at frequent intervals and calling out amenities as the team walked along the trails. By having two to three observers walk in unison along the trail, we estimated that few amenities were missed during GPS data collection.

GPS data were exported in a shapefile format, edited, and incorporated into a GIS application with other data layers (e.g., road networks) from MassGIS. This GIS database of trails and amenities models all six trails in a consistent way at the same scale and time period.

Geographic Unit of Observation: PEAT Trail Segments

In this study each trail was conceptualized as a series of smaller trail segments. During GPS data collection a new trail segment with its own unique segment identifier was created whenever one of the following conditions were met: the trail surface changed, the width of trail changed by greater than 10%, a feature such as an intersecting road was encountered, the level of trail circulation appeared to change (e.g., from primary trail to access segment), and whenever trail segments intersected.

For PEAT observations, two types of trail segments were defined: intersecting road segments and PEAT trail segments. Intersecting roads identified during GPS data collection were retained without further modification. We modified other GPS trail segments, prior to PEAT audits, to create new PEAT segments that were generally more uniform in length. We aggregated very short GPS segments and disaggregated longer GPS segments (e.g., at rail-trails, such as Nashua) to a target length of 400 meters. This distance, equivalent to one or two blocks in an urban setting, would be covered in several minutes by walking at a brisk pace. It defines what we hypothesized was a functionally relevant trail unit for assessing amenities, aesthetics, and maintenance conditions observable to trail users. Due to natural breaks in the trails (e.g., intersecting roads), the average distance of the PEAT segments across the six sites was 283 ± 133 meters. Figure 1 shows an example of PEAT trail and intersecting road segments at the Minuteman Bikeway.

Creating GPS Trail Measures ("Gold Standard")

Using ArcGIS 9.0, ten of the most commonly occurring amenities and design features captured by GPS Quickmarking were assigned to PEAT trail segments. Since amenities that were Quickmarked were offset from the trail, we used a buffer approach to associate amenities with the closest trail segments. In a few cases, dedicated bicycle trails are parallel to pedestrian trails (e.g., Southwest Corridor) and trail amenities are in proximity to and service two parallel trails. Therefore, a 3-meter buffer was created around single track trail segments and a 10-meter buffer was created around parallel trail segments. Amenities that fell within a buffer were assigned to the associated PEAT trail segment (see Figure 1). If an amenity fell within more than one buffer, it was assigned to all the associated PEAT segments. Microsoft Access was used to group the records by PEAT segment and calculate

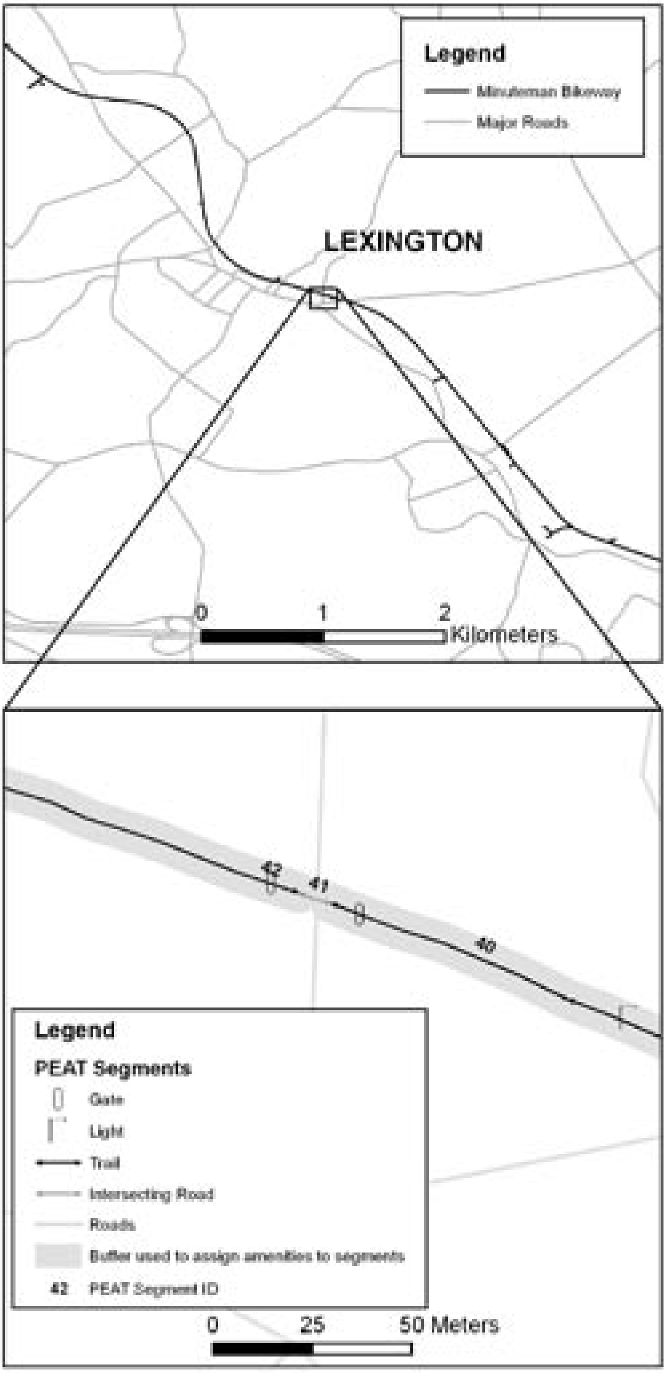

Figure 1—Representation of PEAT segments and buffers used to assign GPS Quickmarks to segments –Minuteman Bikeway, Lexington, MA

the count of each type of amenity for each segment. A derived binary variable was then created indicating the presence or absence of a GPS amenity along a particular segment. These GPS-derived measures served as a criterion measure of 10 trail design characteristics or amenities and were used to assess validity of the corresponding PEAT items.

PEAT Training

During the spring of 2004, two graduate students attended in-office and field training for PEAT that lasted approximately 2 d in total. Trainees were oriented to the study protocol and to each PEAT item using a detailed manual that included operational definitions and illustrative figures and photographs. Following in-office training, the students met with the research team at two non-study sites in Boston where they practiced using PEAT for several hours. Observers were encouraged to refer to the PEAT manual as they completed items. Also, as observations on a segment were completed, the team reviewed responses with the goal of reaching consensus on the ratings. In some cases, as a result of these discussions, operational definitions for PEAT items were further modified.

PEAT Data Collection

The two students conducted PEAT observations at the six study sites during June 2004. At each of the three smaller sites (Cutler Reservation, Danehy Park, Franklin Park), the two observers completed observations on the same day. One observer typically conducted observations in the morning and the other would do so in the afternoon. At the three remaining sites, observations were completed by the observers on different days within a two-to-seven day period. Prior to observations at a given site, the project director gave explicit instructions regarding the trail segments that would be audited and the direction that observers would walk on the trail. Both observers followed the same procedures at a given site (e.g., auditing segments in the same sequence).

Observers brought the following equipment and materials with them: a tablet PC with additional rechargeable batteries, detailed maps that indicated PEAT trail segment and intersecting road ID numbers, study fact sheets for trail users, and a cell phone. In addition, for several trails that did not have clear landmarks at the beginning or end of all PEAT trail segments, we downloaded end-points for segments onto a hand-held Garmin eTrex Vista GPS unit and used these to identify the end and start of new segments. During all observations, a third member of the research team (MF) accompanied the observer. Her primary role was to help read the trail maps and identify PEAT segment IDs, troubleshoot problems with the tablet PC, and operate the hand-held GPS unit.

Statistical Analyses

Descriptive statistics (i.e., frequency distributions, means) were used to summarize the proportion of PEAT intersecting road and trail segments with a particular attribute (e.g., lighting) and the mean rating on a given characteristic across segments (e.g., slope). We report these data for the first observer only.

We assessed inter-observer reliability for all PEAT items, both primary items and corresponding sub-items. However, we are not reporting results for sub-items with ≤ 10 observations (i.e., where both observers completed the item for the same trail segment), since these estimates were considered to be unstable. These included three sub-items each for telephones, public bathrooms and water fountains (working condition, cleanliness, wheelchair accessibility); one sub-item related to accessibility of emergency call boxes; four sub-items specifying the type of trail services available (e.g., food, information); and six items identifying the type of cultural or civic institution adjacent to the trail segment.

For binary PEAT items (yes/no), we estimated Kappa coefficients (k) and calculated observed agreement (%) from two-by-two tables. For ordinal and Likert-scaled items, intraclass correlation coefficients (ICCs) were estimated with a SAS macro (available from the SAS website: http://ftp.sas.com/techsup/download/stat/intracc.html). We used the Shrout-Fleiss reliability random set value that this macro generates. The proportion of observed agreement between the two observers was also calculated.

To assess variability of certain trail characteristics within each site, we calculated Spearman rank correlations between adjacent trail segments for four design items (e.g., slope) and four aesthetics/maintenance items (e.g., glass), using data from the first observer. We hypothesize that certain attributes may vary substantially over the course of a trail, while some may be quite uniform. Therefore, analogous to the determination that multiple days of accelerometer data are needed to derive a reliable estimate of physical activity,[18] it is likely that observations of multiple trail segments are needed to derive a reliable estimate of certain trail characteristics.

Finally, to assess validity of a sub-set of 10 PEAT items for which we had "gold standard" GPS-derived measures, we calculated Kappa coefficients (k) and percent observed agreement. These statistics were calculated separately for both observers. For all analyses, k 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.[19]

Results

Inter-observer Reliability for PEAT Items

Items for Intersecting Roads. Inter-observer reliability was estimated from data that the two observers collected on 45 intersecting road and 185 trail segments. As shown in Table 2, for four out of the five binary items used to rate attributes of intersecting road segments (presence of curb cuts, crosswalks, raised crosswalks, and pedestrian crossing signals), k was high (≥ 0.85). However, an item for presence of signals or signs for vehicular traffic had only "fair" agreement (k = 0.38). The ICC for an overall rating of the intersecting road segment safety, scaled from 1 ("very poor") to 5 ("excellent"), was moderate (Table 3).

Trail Design Items. K-values for the eight binary trail design items (i.e., primary items) varied greatly (Table 2). For presence of access points, gates/bollards, and

Table 2 Inter-observer Reliability for Binary Measures in PEAT, Based on Observations at Six Trails/Paths in Massachusetts

Measure	Segments (*n*)	Proportion of segments with attribute – 1st observer	Observed agreement (2 observers)	Kappa (adjectival rating)
Intersecting road*				
Signal or sign at intersection for vehicles	45	0.86	0.73	0.38 (fair)
Curb cut	45	0.98	1.00	1.00 (almost perfect)
Crosswalk	45	0.91	0.98	0.85 (almost perfect)
Raised crosswalk	45	0.02	1.00	1.00 (almost perfect)
Pedestrian signal	45	0.20	1.00	1.00 (almost perfect)
Trail: design **				
Is surface under repair?	185	0.01	0.99	–0.01 (poor)
Temporary barrier?	185	0.05	0.84	0.19 (poor)
Sufficient vertical clearance?	183	0.33	0.72	0.44 (moderate)
Presence of shoulder	184	0.71	0.74	0.23 (fair)
Road adjacent to trail	184	0.43	0.77	0.54 (moderate)
Buffer along road	61	0.93	0.92	–0.03 (poor)
Access points	184	0.52	0.82	0.63 (substantial)
Gate or bollard	183	0.42	0.86	0.71 (substantial)
Viewpoint	183	0.13	0.92	0.66 (substantial)
Trail: amenities				
Lights	185	0.19	0.94	0.83 (substantial)
Phone	185	0.03	0.98	0.49 (moderate)
Emergency call boxes	185	0.06	0.97	0.69 (substantial)
Restrooms	185	0.06	0.96	0.61 (substantial)
Benches	184	0.30	0.81	0.58 (moderate)
Picnic tables	184	0.12	0.96	0.81 (almost perfect)
Drinking fountain	185	0.06	0.98	0.79 (substantial)
Garbage can	182	0.40	0.92	0.83 (almost perfect)
Signs	185	0.76	0.88	0.70 (substantial)
Cautionary sign	125	0.44	0.82	0.65 (substantial)
Directive sign	125	0.70	0.72	0.24 (fair)
Interpretive sign	125	0.43	0.69	0.32 (fair)
Objective sign	125	0.12	0.81	0.04 (poor)
Regulatory sign	125	0.72	0.87	0.68 (substantial)
Parking	185	0.12	0.88	0.52 (moderate)
Bike racks	183	0.15	0.95	0.76 (substantial)
Exercise or play equipment	185	0.27	0.96	0.89 (almost perfect)
Courts	45	0.31	0.91	0.79 (substantial)
Playground	44	0.30	0.89	0.72 (substantial)
Field	44	0.50	0.93	0.86 (almost perfect)
Pool	44	0.00	1.00	No Kappa
Rink	44	0.05	0.95	No Kappa
Exercise equipment	44	0.05	0.93	–0.03 (poor)
Garden	44	0.14	1.00	1.00 (almost perfect)
Golf course	44	0.20	0.98	0.93 (almost perfect)
Track	44	0.09	1.00	1.00 (almost perfect)
"Other" play	44	0.14	0.70	0.18 (poor)
"Other" exercise	44	0.03	0.95	No Kappa
Services	185	0.04	0.99	0.92 (almost perfect)
Public transit station/stop	185	0.04	0.98	0.70 (substantial)
Cultural or civic destinations	184	0.17	0.86	0.38 (fair)
Other destinations	184	0.15	0.90	0.52 (moderate)
Trail: "other"				
Presence of dogs	185	0.09	0.90	0.43 (moderate)

Note. *Completed on roads and streets that intersected rail–trails or linear parks;** completed on primary and secondary trail segments.: Adjectival ratings for Kappa were based on Landis and Koch.[19]

Table 3 Inter-observer Reliability for Ordered Categorical Measures in PEAT, Based on Observations at Six Trails/ Paths in Massachusetts

Measure	Segments (*n*)	Mean rating – 1st observer	Observed agreement (2 observers)	ICC
Intersecting road segments*				
Safety of intersection	45	3.49 (1.06)	0.40	0.57
Trail: design**				
Condition of path surface	185	4.01 (1.01)	0.38	0.52
Slope	184	1.25 (0.55)	0.75	0.63
Cross slope	183	1.02 (0.18)	0.98	–0.01
Sufficient sight distance?	185	3.34 (0.76)	0.58	0.56
Vegetative cover	182	3.88 (2.66)	0.34	0.32
Width of buffer from road	61	2.79 (0.54)	0.87	0.52
Access point wheelchair accessible?	87	2.73 (0.59)	0.66	0.38
Gate or bollard clearance ≥32 in.	51	2.87 (0.47)	0.90	0.64
Trail: amenities (sub-items)				
Bench condition and cleanliness	47	3.43 (0.74)	0.49	0.07
Bench accessible to wheelchair?	47	1.75 (0.86)	0.57	0.23
Picnic table condition and cleanliness	17	3.09 (0.43)	0.76	0.15
Picnic table accessible to wheelchair?	17	1.45 (0.60)	0.88	0.79
Garbage cans overflowing	61	1.15 (0.39)	0.84	0.005
Number of parking spaces	16	3.26 (1.10)	0.69	0.84
Bicycle rack condition	20	3.85 (0.60)	0.70	0.25
Trail: maintenance/aesthetics				
Glass	185	1.14 (0.39)	0.87	–0.02
Litter	185	2.00 (0.67)	0.49	0.03
Graffiti	185	1.76 (0.91)	0.51	0.50
Vandalism	185	1.16 (0.44)	0.85	0.09
Odor	185	1.11 (0.38)	0.74	–0.04
Noise	185	2.12 (0.79)	0.59	0.40
Dog droppings	185	1.15 (0.47)	0.89	0.07

Safety of intersection; Condition of Path Surface; Bench condition and cleanliness; Picnic Table condition and cleanliness; Bicycle rack condition: 1 = Very Poor; 2 = Poor; 3 = Fair; 4 = Good; 5 = Excellent.

Slope; Cross Slope: 1 = Flat or gentle; 2 = Moderate; 3 = Steep.

Vegetative cover / Built enclosure: 1 = Continuous lateral visibility; 2 = Moderate lateral visibility; 3 = No lateral visibility.

Width of buffer: 1 = < 1 meter; 2 = 1 – 3 meters; 3 = > 3 meters.

Sufficient Sight Distance; Access point wheelchair accessible?; Gate or bollard clearance ≥ 32 in.; Bench accessible to wheel chairs; Picnic table accessible to wheel chair; Garbage cans overflowing: 1 = No (None); 2 = Some; 3 = Yes (All).

Number of parking spaces: 1 = 10 or less; 2 = 11 – 25; 3 = 26 – 50; 4 = more than 50.

Glass; Litter; Graffiti, Vandalism; Odor; Noise; dog droppings: 1 = None; 2 = A little; 3 = Some; 4 = A lot.

viewpoints, k was "substantial." Two items indicating whether there was sufficient vertical clearance along the trail segment and a road adjacent to the trail had "moderate" reliability. The remaining three binary design items had "fair" or "poor" k-values, yet observed agreement was ≥ 0.74.

ICCs for trail design items that characterized surface condition, slope, sight distance, and enclosure/vegetative cover were in the range of 0.32 to 0.63 (Table 3). The ICC for a cross slope item was extremely low; however, observed agreement was high. Two sub-items characterizing the accessibility of access points and gates/bollards and one assessing the width of a buffer between the trail and a road had fair to moderate ICCs.

Trail Amenity Items. As Table 2 shows, overall the k-values for trail amenities were higher than the values obtained for design items. For the 16 primary trail amenity items, k ranged from 0.38 (cultural/civic destinations) to 0.92 (presence of services). Using Landis and Koch's adjectival ratings, the coefficient for 11 items was "substantial" or "almost perfect." Four items (presence of phones, benches, parking, and "other" destinations) had "moderate" k-values, although observed agreement was ≥ 0.81. The k-values for cautionary and regulatory sign sub-items were "substantial," but were relatively low for other types of signs. Six of the 11 sub-items characterizing the type of exercise or play equipment next to the trail had k values ≥ 0.72. Table 3 includes results for seven PEAT amenity sub-items that had a sample size ≥ 10. Except for an assessment of picnic table accessibility and a rating of the number of parking spaces, these items had low ICC values.

Maintenance/Aesthetics Items. Five of the seven PEAT maintenance/aesthetics items had ICC values less than 0.10, although observed agreement was ≥ 74% for four of these items (Table 3). ICCs for a rating of noise and graffiti were in the fair to moderate range. A summary of reliability results for all primary items in PEAT is shown in Figure 2.

PEAT Item Correlations Between Trail Segments

Overall, the correlations between adjacent trail segments varied greatly by the type of attribute and type of site (Table 4). No coefficient was generated for a number of PEAT attributes, since there was no variation in the ratings made by the observer. This occurred for the slope item at two rail-trails and for glass and vandalism at Cutler Reservation, where there was no glass or vandalism observed on any segment.

Overall, the majority of correlation coefficients for the eight PEAT items were low. For surface condition, the highest correlation between trail segments was found at the suburban and rural paved rail-trails (both $r = 0.50$). The highest correlation coefficients for slope were found at Danehy Park ($r = 0.57$) and Southwest Corridor ($r = 0.46$), two distinct facilities. For the sufficient sight distance item, the only significant correlation was found at Nashua River Rail Trail, which has numerous long and straight trail segments. Nevertheless, this correlation was still relatively low ($r = 0.29$). There was a fairly strong statistically significant correlation between adjacent trail segments for vegetative cover/enclosure at Danehy Park ($r = 0.76$), which has numerous sections of trails bordered by open ball fields. The correlations at Franklin Park and Nashua were also statistically significant, but not as strong

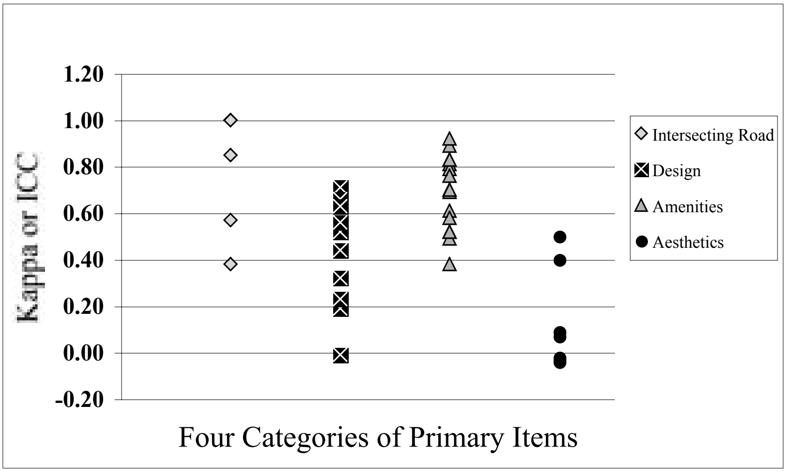

Figure 2—Inter-observer reliability for 4 categories of primary items in PEAT (n = 42)

Table 4 Site-Specific Spearman Rank Correlations Between Adjacent Trail Segments for Eight Select PEAT Design and Aesthetics/Maintenance Items for First Observer

	Cutler Reservation (n = 13)	Danehy Park (n = 13)	Franklin Park (n = 28)	Minuteman Bikeway (n = 55)	Nashua River Rail Trail (n = 47)	Southwest Corridor (n = 21)
Surface condition (1 = very poor to 5 = excellent)	0.29 (0.34)	0.41 (0.17)	0.20 (0.31)	0.50 (0.0001)	0.50 (0.0004)	0.22 (0.34)
Slope (1 = flat or gentle, 2 = moderate, 3 = steep)	–0.51 (0.08)	0.57 (0.04)	0.21 (0.27)	No variation (all = 1)	No variation (all = 1)	0.46 (0.04)
Sufficient sight distance (1 = all of segment, 2 = most, 3 = some, 4 = none of segment)	0.36 (0.23)	–0.47 (0.11)	0.24 (0.22)	0.05 (0.74)	0.29 (0.05)	–0.31 (0.18)
Vegetative cover (1 = continuous lateral visibility, 2 = moderate visibility, 3 = no lateral visibility)	–0.08 (0.79)	0.76 (0.003)	0.46 (0.01)	0.25 (0.08)	0.29 (0.05)	–0.04 (0.85)
Glass (1 = none, 2 = a little, 3 = some, 4 = a lot)	No variation (all = 1)	–0.08 (0.79)	0.43 (0.02)	0.12 (0.38)	–0.04 (0.77)	–0.11 (0.65)
Litter (1 = none, 2 = a little, 3 = some, 4 = a lot)	–0.10 (0.75)	–0.08 (0.79)	–0.17 (0.40)	0.16 (0.24)	0.20 (0.17)	0.04 (0.86)
Graffiti (1 = none, 2 = a little, 3 = some, 4 = a lot)	–0.37 (0.22)	0.12 (0.70)	0.21 (0.28)	0.18 (0.18)	0.14 (0.35)	0.14 (0.53)
Vandalism (1 = none, 2 = a little, 3 = some, 4 = a lot)	No variation (all = 1)	–0.08 (0.79)	0.19 (0.34)	0.11 (0.44)	–0.09 (0.53)	0.33 (0.14)

Note. P-values are in parenthesis.

as the one for Danehy ($r \leq 0.46$). For the four aesthetics/maintenance items, most correlation coefficients were low and not statistically significant. The one exception was at Franklin Park, where we found a significant correlation between adjacent trail segments for glass.

Validity of PEAT Items

For a sub-set of PEAT amenity and design items, for which we had GPS-derived measures ("gold standard"), we were able to determine validity of the PEAT assessments for each observer (Table 5). For the first observer, k-values were in the "moderate" to "substantial" range for 7 of 10 PEAT items: presence of access points, gate/bollard, lighting, benches, picnic tables, drinking fountains, and garbage cans. Observed agreement between the first observer's PEAT assessments and the GPS-derived items was 77% or higher for all comparisons. Generally, the k-values for the second observer were lower than for the first, although the overall pattern of agreement between PEAT and GPS measures was similar.

Table 5 Validity of Select PEAT Trail Design and Amenity Measures for Two Observers, Using GPS Quickmark Derived Items as "Gold Standard" Measure

Design and Amenity Measures	Segments (*n*)	Observer 1		Observer 2	
Presence of:		**Observed agreement with GPS**	**Kappa**	**Observed agreement with GPS**	**Kappa**
Access point	184–185	0.86	0.72	0.74	0.50
Gate or bollard	182	0.85	0.69	0.88	0.76
Lighting	185	0.88	0.68	0.86	0.65
Phone	185	0.97	–0.01	0.98	–0.01
Emergency call box	185	0.94	0.25	0.97	0.39
Bench or seating	184–185	0.84	0.59	0.76	0.44
Picnic table	184–185	0.95	0.74	0.95	0.69
Drinking fountain	185	0.96	0.57	0.96	0.44
Garbage can	185	0.91	0.80	0.91	0.80
Exercise or play equipment	185	0.77	0.22	0.78	0.23

Discussion

In what may be the first study to develop and test both the reliability and validity of an audit tool for community trails and paths, PEAT demonstrated good inter-observer reliability for the majority of its primary design and amenity items. For seven aesthetics/maintenance items that were more subjective in nature and did not represent "fixed" physical characteristics of trails, inter-observer reliability was fair. In addition, six items for assessing streets that intersect trails had fairly good inter-observer reliability overall. Our study illustrated that for certain trail attributes, observations of multiple trail segments may be needed to derive a reliable estimate for the entire trail. Lastly, we were able to demonstrate the validity of a sub-set of PEAT items, using "gold standard" measures based on a more resource-intensive GPS data collection process.

The current study measures many of the same domains and attributes as those in Brownson et al. [10] and Pikora et al.,[13] yet our study was conducted in a specific subset (trails) of the broader environment (neighborhoods) examined in those two studies. By briefly comparing our findings to those of Brownson and colleagues,[10] it may provide some perspective on the ability of observers to reliably measure certain attributes of the built environment. Although the items were scaled differently, the k-value for an item on "levelness and condition of the sidewalk" in the Brownson et al. study was somewhat higher than the value we obtained for an item on "condition of path surface" (0.66 versus 0.52). However, the observed agreement was almost exactly the same (0.40 and 0.38). Observed agreement for rating the presence of public telephones and trash bins was almost identical in the two studies (above 0.90); although in our study the k-value was lower for telephones (0.49 versus 0.79) and higher for trash bins (0.83 versus 0.47). We used two aesthetics/maintenance items that were worded and scaled similar to items used by Brownson and colleagues. The ICC for noise was higher in our study (0.40 versus 0.14), but the agreement between observers was comparable (0.59 versus 0.68). Interestingly, observed agreement for graffiti was much lower in our study (0.51 versus 0.90), but we found a somewhat higher ICC (0.50 versus 0.32).

PEAT could benefit from further refinements and testing. In some cases, improving response categories, operational definitions, and observer training may help to improve the tool's inter-observer reliability—even for items that had poor reliability in the current study. In other situations, it may be judicious to remove items altogether, since through further testing they may be determined to be too difficult for the "novice" observer to rate. Nevertheless, some might consider a trail attribute such as sight distance a challenging construct for observers to rate (e.g., adequate sight distance can be different for someone who is walking versus cycling), yet in our study this item had moderate inter-observer reliability. To further refine PEAT, correlates studies that assess relationships between trail characteristics and trail use or trail satisfaction may need to be conducted. This type of study may help to determine a core set of items that need to be audited with trails and community paths—which in turn will make auditing more efficient.

We also note that for less prevalent trail features (e.g., public telephones in our study), any missed observations can have a large impact on the inter-observer reliability (as indicated by the k-value or ICC). Finally, our overall approach to developing PEAT was to be more inclusive and by doing so we likely included some items that represented constructs that were too difficult to operationally define in a clear, simple manner. One example was the item for temporary barriers along the trail, which had a poor k-value. As illustrated in our study, a low reliability coefficient does not necessarily equate with low observed agreement. There are several possible reasons for low ICCs,[20] including actual low inter-observer reliability, little variation between trail segments on a given attribute, instability due to small cell sizes, and a change in the environment from initial to re-audit (e.g., litter cleaned up between audits). All of these possible explanations should be taken into account when designing future audit tools.

To illustrate how certain physical characteristics, such as surface condition, slope, litter, and glass may vary within a site, for each study site we estimated correlations between adjacent trail segments for select characteristics. One question of interest pertains to the number of segments (i.e., observations) needed to derive a reliable estimate of a given characteristic. Although we did not resolve this issue for auditing trails, our results indicate that multiple measurements are important. The precise number appears to be dependent on both the type of attribute being assessed and type of trail.

Strengths of this study include the fact that the PEAT tool was one component of a comprehensive GIS database developed for these trails. By conducting a fairly extensive GPS data collection process, we created a valid spatial framework for our study sites that in turn guided the PEAT data collection process. A related strength was that we conceptualized trails as a series of shorter trail segments that served as our unit of observation for PEAT. This allows us to examine variation in characteristics within trails. Finally, our validation of select PEAT measures with "gold standard" GPS data represents an alternative to validating a trail audit tool by assessing associations with trail use. A limitation of this latter approach is that lack of an association does not prove that the trail measures are not valid—it may simply indicate that there is no association between trail characteristics and trail use.

This study has several limitations. Since we only surveyed trail users and not non-users, it is possible that certain trail characteristics that act as barriers to use were not included in PEAT. For a number of items, such as sub-items related to the condition and cleanliness of restrooms, the sample size was too small to assess inter-observer reliability. Although the data generated from PEAT can be viewed as objective data, in fact a number of items were fairly subjective in nature (representing observers' perceptions). For example, although we attempted to operationally define slope, no special equipment was used to make the assessment—it was essentially the observer's perception of the slope along various trail segments. It may be unrealistic to expect that independent observers will have a high level of agreement for trail attributes such as slope. In the future, it may be preferable to have observers use measurement devices or to assess slope using other objective data sources (e.g., GIS data).

Conclusions

The majority of primary items in the PEAT tool have moderate to high inter-observer reliability, suggesting the tool is ready for use by other researchers and practitioners. Generally, reliability was highest for trail amenity items and lowest for more subjective items related to maintenance/aesthetics. In terms of future research applications, PEAT could be used in studies that examine associations between trail characteristics and use, which could be measured via random digit-dial phone surveys of community residents or objective monitoring along various sections of trails. Finally, several actions related to trail audit tools deserve consideration: 1) develop and test a shorter instrument that captures critical trail attributes; 2) integrate PEAT into a mobile GIS unit, which may produce a more efficient data collection process; and 3) refine sampling schemes and units of observation—to better ensure efficiency and that reliable estimates of trail characteristics are obtained.

Author's note: The Path Environment Audit Tool (PEAT) and manual are available on the Active Living Research website (http://www.activelivingresearch.org/). Contact Philip Troped (ptroped@hsph.harvard.edu) for information on the use of PEAT.

Acknowledgments

Preparation of this paper was supported by a grant from The Robert Wood Johnson Foundation, Active Living Research program. We extend a sincere thank you to our local consultants, Deneen Crosby, Aldo Gherin and Jim Purdy, for their input into the design of the trail audit tool. We are also grateful to Tyler Huffman, for his GPS expertise, and the time and energy he devoted to collecting GPS data at study sites in Massachusetts. Finally, we wish to thank our two observers, Bhavna Babani and Joseph Allen, for their diligence in conducting the PEAT observations.

References

1. Srinivasan S, O'Fallon LR, Dearry A. Creating healthy communities, healthy homes, healthy people: initiating a research agenda on the built environment and public health. *Am J Public Health.* Sep 2003;93(9):1446-1450.
2. Sallis JF, Hovell MF, Hofstetter CR, et al. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. *Public Health Rep.* Mar-Apr 1990;105(2):179-185.
3. King AC, Jeffery RW, Fridinger F, et al. Environmental and policy approaches to cardiovascular disease prevention through physical activity: issues and opportunities. *Health Educ Q.* Nov 1995;22(4):499-511.
4. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: a review. *Am J Prev Med.* Apr 2002;22(3):188-199.
5. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med.* Spring 2003;25(2):80-91.
6. Ewing R, Cervero R. Travel and the built environment: a synthesis. *Transportation Research Record.* 2001;1780 (4):153-158.
7. Owen N, Leslie E, Salmon J, Fotheringham MJ. Environmental determinants of physical activity and sedentary behavior. *Exerc Sport Sci Rev.* Oct 2000;28(4):153-158.
8. Sallis JF, Johnson MF, Calfas KJ, Caparosa S, Nichols JF. Assessing perceived physical environmental variables that may influence physical activity. *Res Q Exerc Sport.* Dec 1997;68(4):345-351.
9. Brownson RC, Chang JJ, Eyler AA, et al. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. *Am J Public Health.* Mar 2004;94(3):473-483.
10. Brownson RC, Hoehner CM, Brennan LK, Cook RB, Elliott B, McMullen KM. Reliability of two instruments for auditing the environment for physical activity. *JPAH.* 2004;1(3):189-207.
11. Emery J, Crump C, Bors P. Reliability and validity of two instruments designed to assess the walking and bicycling suitability of sidewalks and roads. *Am J Health Promot.* Sep-Oct 2003;18(1):38-46.
12. Saelens BE, Sallis JF, Black JB, Chen D. Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health.* Sep 2003;93(9):1552-1558.

13. Pikora TJ, Bull FC, Jamrozik K, Knuiman M, Giles-Corti B, Donovan RJ. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med.* Oct 2002;23(3):187-194.
14. Brownson RC, Housemann RA, Brown DR, et al. Promoting physical activity in rural communities: Walking trail access, use, and effects. *Am J. Prev. Med.* 2000;18(3):235-241.
15. Troped PJ, Saunders RP, Pate RR, Reininger B, Ureda JR, Thompson SJ. Associations between self-reported and objective physical environmental factors and use of a community rail-trail. *Prev Med.* Feb 2001;32(2):191-200.
16. Wang G, Macera CA, Scudder-Soucie B, Schmid T, Pratt M, Buchner D. Cost effectiveness of a bicycle/pedestrian trail development in health promotion. *Prev Med.* Feb 2004;38(2):237-242.
17. Sampson RJ, Raudenbush SW. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol.* 1999;105(3):603-651.
18. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? *Med Sci Sports Exerc.* Feb 2000;32(2):426-431.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* Mar 1977;33(1):159-174.
20. Streiner DL, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. New York: Oxford University Press; 1995.