Journal of Physical Activity and Health 2006, 3, Suppl 1, S176-S189

Development of a Direct Observation Instrument to Measure Environmental Characteristics of Parks for Physical Activity

Ariane L. Bedimo-Rung, Jeanette Gustat, Bradley J. Tompkins, Janet Rice, and Jessica Thomson

Background: The study's purpose is to describe the development and evaluate the reliability (inter-observer agreement) and validity (rater agreement with a gold standard) of a direct observation instrument to assess park characteristics that may be related to physical activity. *Methods:*. A direct observation instrument of 181 items was developed based on a conceptual model consisting of the following domains: features, condition, access, esthetics, and safety. Fifteen pairs of observers were trained and sent to two parks simultaneously to assess two Target Areas each. *Results:* Overall domain reliability was 86.9%, and overall geographic area reliability was 87.5%. Overall domain validity was 78.7% and overall geographic area validity was 81.5%. *Conclusions:* Inter-rater reliability and validity were generally good, although validity was slightly lower than reliability. Objective items showed the highest reliability and validity. Items that are time-sensitive may need to be measured on multiple occasions, while items asking for subjective responses may require more supervised practice.

Key Words: audit, Bedimo-Rung Assessment Tools, reliability, validity

Despite the many health benefits of regular physical activity, half of all adults[1] and a third of all students in grades 9 to 12[2] did not meet national guidelines for physical activity in 2003. Because of this trend, researchers and practitioners from multiple disciplines are looking to the built environment for ways to engineer physical activity back into people's daily lives.[3] Urban parks are one common feature of the built environment. They provide opportunities for the public to engage in physical activity at little or no cost, and are located in most American cities.[4] As such, they provide a promising opportunity to promote physical activity. Previous studies cite several factors that are positively associated with physical activity. These include access to and satisfaction with recreational facilities,[5-13] the presence of enjoyable scenery,[7, 8, 14, 15] and frequency of seeing others exercise.[7-9, 14] While a number of

Bedimo-Rung, Tompkins, and Thomson are with the Louisiana State University School of Public Health, New Orleans, LA 70112. Gustat and Rice are with the Tulane University School of Public Health and Tropical Medicine, New Orleans, LA 70112.

audit instruments have been developed to assess the walkability and bikability of environments,[16] little is known about the specific characteristics of parks that are related to physical activity. Understanding the park characteristics that are related to physical activity could inform park design and amenity standards that might support more physical activity in the population. The objective of this paper is to describe the development and evaluate the reliability and validity of a direct observation instrument to assess park characteristics.

Methods

Bedimo-Rung Assessment Tools (BRAT) Overview

The direct observation instrument assessed in this study was developed as part of a larger project, the goal of which was to develop and test a set of instruments (the Bedimo-Rung Assessment Tools, or BRAT) designed to measure the physical, social, and policy environments of parks. The BRAT is a multi-method approach incorporating the use of direct observation (BRAT-DO), informant interviews, aerial photography, geographic information systems (GIS), and archival data to collect information about parks that may be related to physical activity occurring in them.

Items in the BRAT were developed based on the conceptual framework of park environmental characteristics developed by Bedimo-Rung et al. (Figure 1).[17] Six domains of park characteristics have been identified: 1) Features, 2) Condition, 3) Access, 4) Esthetics, 5) Safety, and 6) Policies. *Features* include the number, size, and type of facilities and programs offered at parks, as well as the diversity of users and uses found within them. *Condition* covers the routine upkeep, maintenance, and repair of park facilities, as well as incivilities, or cues in the environment, that provide signals about how to behave. Evidence of incivilities is found in disorderly physical surroundings (e.g., litter, graffiti) and disruptive social behaviors (e.g., drinking, loitering). Four different types of *Access* are considered in the framework. Availability includes the amount of total park space available in a given city; equitable access considers how park space is distributed across communities; individual access refers to an individual's ability to get to a park; and within park access refers to the ability of people to move around easily inside the boundaries of a park and access specific facilities. *Esthetics* incorporates the perceived attractiveness and appeal of the various design elements of a park as well as how the physical features of parks are laid out. *Safety* refers both to the personal security of park users from crime and to the ability of park features to prevent injury. *Policies* refer to park design policies, management practices, and budget procedures.

In addition to the six domains of park characteristics, the framework proposes four geographic areas in which data should be collected.[17] *Activity Areas* are the areas within a park that are specifically designed or commonly used for physical activity. They can include sports fields and courts, swimming pools, paths or trails, playgrounds, open green spaces, or other areas where physical activity occurs. Park *Supporting Areas* include those facilities and equipment that make physical activity in parks attractive and safe to a variety of users. They may contain features that do not directly promote physical activity but are nonetheless an integral part of the park visitation experience, such as community buildings, shelters, restroom/

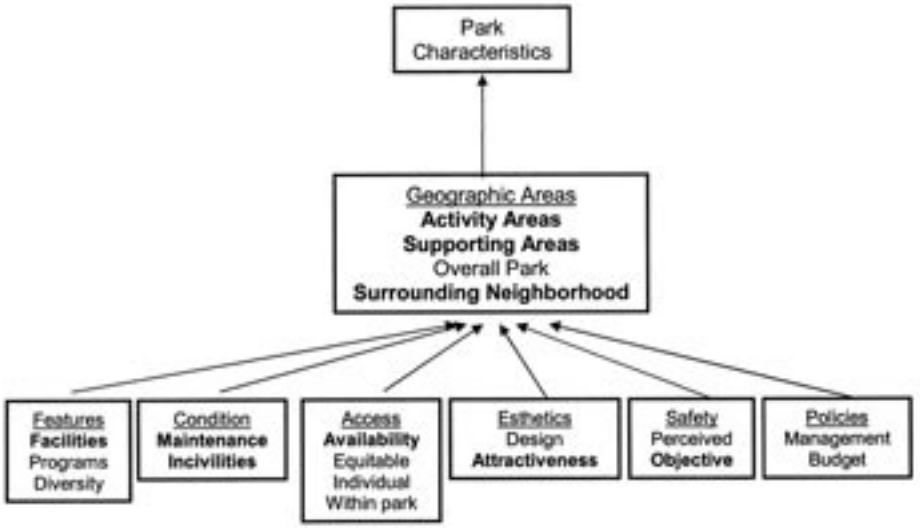

Figure 1—Conceptual model of park attributes that may be related to physical activity. Items in bold represent geographic areas and domains that are assessed in the direct observation instrument of the BRAT.

changing facilities, picnic areas, parking lots, etc. The *Surrounding Neighborhood* is another important area to consider when evaluating park characteristics, since people must cross through the surrounding neighborhood in order to enter the park. A variety of neighborhood characteristics across several domains are likely to have an effect on how people perceive and use a park, including traffic, blighted or abandoned housing, crime, and resident demographics. Finally, the *Overall Park* must be considered. Certain park characteristics, such as esthetic appeal, size, and diversity of programs, are not limited to specific areas of the park and should be applied to the whole park area.

While the framework is comprehensive in describing the types of attributes that may be relevant to physical activity, it is likely not possible to collect data on all aspects using only one method of data collection. This paper focuses on just one of the BRAT methods, the direct observation instrument (BRAT-DO), which captures data on many but not all aspects of the conceptual model. The domains and geographic areas that are covered in the BRAT-DO are shown in bold in Figure 1.

Instrument Development

The BRAT-DO instrument is a paper-and-pencil assessment used by observers in the field to visually identify and evaluate the physical characteristics of parks. It was developed through a series of meetings with an expert panel using the Delphi method.[18] The expert panel consisted of regional planners, park directors, physical activity specialists, study investigators, physical activity enthusiasts, and park users. An initial brainstorming session took place where all potential park characteristics were outlined and discussed. Study investigators then operationalized

the characteristics into measurable instrument items. An iterative process of refining this list of items then followed, which included periodic meetings of the study investigators and preliminary field testing sessions during which items were discussed, revised, or discarded according to their measurement feasibility and their concordance with the conceptual model.

The final BRAT-DO instrument includes items in five of the six conceptual domains: Access, Condition, Esthetics, Features, and Safety, as well as subcategories of these domains. Several administrative items that identify specific areas and observers as well as dates and times of observation, are also included. The park geographic area types considered include Activity Areas, Supporting Areas, and the Surrounding Neighborhood. The instrument contains one form for each Target Area, Street, and Activity Area (Court, Green Space, Path, Playground, and Sports Field). Target Areas are mutually exclusive subdivisions of the park and are predetermined by the investigators. The Target Area form includes questions on esthetics and condition, landscaping, trash, litter, sounds and smells, benches, bike racks, shelters, restrooms, concession stands, buildings, drinking fountains, picnic tables, water features, art and monuments, parking areas, and park staff. Each Activity Area is considered to be located within a Target Area, so that all the items on the Target Area form apply to the Activity Areas as well. The Surrounding Neighborhood is captured through a series of questions on the Street form, where streets border, cross, or are located entirely within a Target Area. Items relevant to the Overall Park were assessed as well, but are not included in this analysis due to the limited number of Target Areas that were sampled. Thus, seven forms covering five domains and three geographic area types were used, for an overall total of 181 items (Table 1).

Items on the BRAT-DO instrument consisted of a variety of question types. There were checklists, asking observers to note the presence or absence of specified features (e.g., What structures are present on the sports field? Circle all that apply: football goal posts, soccer goal posts, fence around home plate, dugouts, seating, batting cage/warm-up area, scoreboard, stadium, no structures present). Some items

Table 1 Number of Items in the Direct Observation Tool Grouped by Geographic Area and Domain

	Domains					
Geographic Areas	**Access**	**Condition**	**Esthetics**	**Features**	**Safety**	***Total Items***
Target Area Items	5	24	15	38	3	85
Street Items	11	0	0	0	2	13
Court Items	6	3	0	17	0	26
Green Space Items	1	1	1	1	0	4
Path Items	1	1	0	7	2	11
Playground Items	2	4	0	8	7	21
Sports Field Items	5	2	0	14	0	21
Total Items	31	35	16	85	14	181

asked the observers to rate a characteristic using a 5-point rating scales (e.g., How much graffiti is visible in the Target Area? Choose one: none, very little, some, a moderate amount, a lot). Some items asked observers to select one of a variety of categorical responses (e.g., What type of surfacing is under the play equipment? Choose one: hard, grass/turf/soil, loose fill, rubber tiles or unitary synthetic surface). Other items asked observers to write in information (e.g., If there is a speed limit sign, specify speed limit.).

A first draft (Round 1) was completed and pilot-tested by one data collector in a 400-acre regional test park in fall 2003 and then further refined. A second version of the instrument was tested for reliability in spring 2004 (Round 2) in the same test park, using a representative sample of eight of the 24 Target Areas, and 16 teams of observer pairs. Revisions were again made and a third and final instrument (Round 3) was developed and tested for reliability in fall 2004. The results from this third and final round of testing are presented here.

Training, Mapping, Data Collection Procedures, and Gold Standard

During Round 3 of testing, 15 teams of observer pairs were trained in the BRAT-DO methodology. Observers were students in a masters-level Survey Methodology course. A comprehensive BRAT-DO reference manual was developed by the investigators for use in training and as a reference guide. The manual outlines the procedures to be used for data collection, specific instructions on how to answer individual items, and illustrative photos of item examples. Observer pairs were asked to read the manual prior to the training. During the training session, a PowerPoint presentation summarizing the BRAT-DO reference manual was presented. This was followed by field practice and discussion in a park.

On the day of data collection, seven teams of observer pairs were sent to the original test park, and eight teams were sent to a second 200-acre regional park. Two large regional parks were chosen for testing because of the diversity of their features. Four representative Target Areas containing a variety of targeted park characteristics were selected for evaluation in each park. Area boundaries were identified by the investigators prior to the data collection. Aerial photographs of each park were used to identify logical Target and Activity Areas, as well as visible park features. Investigators then visited each park to refine boundary lines and identify features that were not visible on the aerial photographs. Final maps were supplied to each observer. Each team surveyed two Target Areas on the same morning so that each of the eight target areas was surveyed a total of three to four times by three to four different teams.

A team of two investigators closely involved in the development of the BRAT-DO assessed each area by consensus 2 wk prior to data collection. These assessments were used as the gold standard.

Analysis

Analysis was conducted in SPSS and Excel. Inter-observer agreement was used to assess reliability by calculating percent agreement among the observers. This method is the simplest and most frequently used index of agreement.[19] The drawback

to using this method, versus another index of agreement such as Cohen's kappa, is that it overestimates true agreement because it does not account for agreement due to chance. This study was not designed to reliably calculate Cohen's kappa values. For this study, Cohen's kappa would optimally be used when two (or more) observers rate an item on a dichotomous scale for a number of Target Areas. In this study, however, 15 pairs of observers were used to rate the various features of the two parks. While all eight Target Areas were measured between three and four times, no single pair of observers rated more than two Target Areas for the study. Hence, kappa values would be calculated using a sample size of two Target Areas, not a useful or reliable sample size on which to base conclusions. Similar constraints occur for the other Activity Areas.

Continuous items assessed on a five-point rating scale were dichotomized, and frequency distributions for individual items in each area were calculated. Reliability for each item was assessed by summing the number of observations with the majority response in an area divided by the total number of areas observed for the item. Reliability on individual items was then summarized and grouped into domains and geographic areas. Average, maximum, and minimum agreement for each domain and geographic area was calculated, along with the total and percentage of items in each domain or geographic area that showed agreement above 70%, which we consider to be sufficiently reliable.[19, 20]

Validity analysis was conducted by comparing the responses of the data collectors to the responses of the gold standard evaluation conducted by the two experts prior to the third round of testing. Validity for each item was assessed by summing the number of observations with the correct response (compared to the gold standard) in an area divided by the total number of areas observed for the item. Validity of individual items was summarized by grouping items into domains and geographic areas. Average, maximum, and minimum agreement for each domain and geographic area was calculated, along with the total and percentage of items in each domain or geographic area that showed agreement above 70%, which is considered to be sufficiently reliable.[19, 20]

Results

Reliability

Table 2 shows the inter-rater reliability of individual items summarized by domain. The average agreement within each domain ranged from a low of 83.7% in the domain of Esthetics to a high of 91.9% in the domain of Features. Agreement within each domain ranged from 63.6% to 100%. Average agreement over all the domains was 86.9% (range 67.3% to 100%). All of the domains exhibited high overall agreement (defined as having a high percentage of individual items with greater than or equal to 70% agreement), with a minimum of 87.5% of items within domains at greater than or equal to 70% agreement. Table 3 shows the same data, summarized instead by geographic area. Average agreement by area ranged from a low of 81.8% in Path to a high of 93.3% in Court. Agreement within each area ranged from 63.6% to 100%. Average agreement over all the geographic areas was 87.5% (range 69.6% to 99.5%). All geographic areas had at least three-quarters of their items exhibit high agreement.

Table 2 Inter-Rater Reliability by Domain

	# Items	Average Agreement	Highest Agreement	Lowest Agreement	# items ≥ 70% Agreement	% items ≥ 70% Agreement
Access	31	88.2%	100.0%	66.7%	30	96.8%
Condition	35	85.1%	100.0%	63.6%	32	91.4%
Esthetics	16	83.7%	100.0%	66.7%	14	87.5%
Features	85	91.9%	100.0%	66.7%	83	97.6%
Safety	14	85.6%	100.0%	72.7%	14	100.0%
Overall	181	86.9%	100.0%	67.3%	173	95.6%

Table 3 Inter-Rater Reliability by Geographic Area

	# Items	Average Agreement	Highest Agreement	Lowest Agreement	# items ≥ 70% Agreement	% items ≥ 70% Agreement
Target Area	85	88.9%	100.0%	66.7%	83	97.6%
Street	13	89.3%	100.0%	81.6%	13	100.0%
Court	26	93.3%	100.0%	66.7%	25	96.2%
Green Space	4	84.0%	100.0%	69.4%	3	75.0%
Path	11	81.8%	100.0%	66.7%	9	81.8%
Playground	21	88.3%	100.0%	63.6%	19	90.5%
Sports Field	21	87.2%	96.6%	72.4%	21	100.0%
Overall	181	87.5%	99.5%	69.6%	173	95.6%

Validity

Table 4 shows the validity of individual items summarized by domain. The average agreement with the gold standard assessment within each domain ranged from a low of 68.3% in the domain of Esthetics to a high of 88.3% in the domain of Features. Agreement within each domain ranged from 33.3% to 100%. Average validity over all the domains was 78.8% (range 42.8% to 99.3%). Three of the domains (Access, Features, and Safety) showed high overall agreement, where most (at least 85.7%) of their individual items exhibited individual item agreement over 70%. Table 5 shows the same data, summarized by geographic area. Average validity by area ranged from a low of 72.9% in Green Space to a high of 88.8% in Court. Agreement within each area ranged from 33.3% to 100%. Average validity over all the domains was 81.5% (range 48.7% to 99.5%). Every geographic area except for Playground showed high overall validity, with at least three-quarters of the individual items exhibiting individual item agreement over 70%.

Table 4 Validity Assessment by Domain

	# Items	Average Agreement	Highest Agreement	Lowest Agreement	# items ≥ 70% Agreement	% items ≥ 70% Agreement
Access	30	84.5%	100.0%	62.5%	28	93.3%
Condition	35	72.6%	100.0%	36.4%	21	60.0%
Esthetics	16	68.3%	96.7%	33.3%	10	62.5%
Features	85	88.3%	100.0%	45.5%	80	94.1%
Safety	14	79.7%	100.0%	36.4%	12	85.7%
Overall	180	78.7%	99.3%	42.8%	151	83.3%

Table 5 Validity Assessment by Geographic Area

	# Items	Average Agreement	Highest Agreement	Lowest Agreement	# items ≥ 70% Agreement	% items ≥ 70% Agreement
Target Area	85	80.5%	100.0%	36.7%	70	82.4%
Street	13	85.7%	100.0%	73.7%	13	100.0%
Court	26	88.8%	100.0%	50.0%	22	84.6%
Green Space	4	72.9%	100.0%	33.3%	3	75.0%
Path	10	78.4%	100.0%	45.5%	9	90.0%
Playground	21	80.9%	100.0%	36.4%	15	71.4%
Sports Field	21	83.1%	96.6%	65.5%	19	90.5%
Overall	180	81.5%	99.5%	48.7%	151	83.9%

Reliability and Validity of Selected Items

Table 6 illustrates selected items, the number of sites in which the items were assessed, the number of observations that occurred for each item, the number of possible response options for each item, individual item reliability, and individual item validity based on comparisons with the gold standard assessment for the different domains and various geographic regions. For example, in the Access domain, the item "Can the entire Target Area be locked?" was assessed on the Target Area form in eight different sites (or Target Areas) by a total of 30 observer pairs. This item had an item reliability agreement of 93.8% and an item validity agreement of 93.8%.

Table 6 Selected Items, Individual Item Reliability, and Individual Item Validity from the Direct Observation Instrument

Domain	Geographic Area	Item	*N* (sites)	*N* (obs)	# Possible Response Options	Individual Item Reliability	Individual Item Validity
Access	Target Area	Can the entire TA be locked?	8	30	2	93.8%	93.8%
Access	Street	What is the traffic volume of the street?	10	38	2	90.0%	80.0%
Access	Court	Are there sources of light that would allow the courts to be used at night?	3	12	2	100.0%	100.0%
Access	Sports Field	Are there signs specifying that reservations are required to use the sports field?	8	29	2	85.4%	72.9%
Condition	Target Area	Rate the condition of the landscaping in the Target Area.	8	30	5*	87.5%	87.5%
Condition	Target Area	How much litter is present in the Target Area?	8	30	5*	74.0%	55.2%
Condition	Court	How much of the court structures appear broken or missing on the courts?	3	12	5*	91.7%	58.3%
Condition	Green Space	Rate the condition of the surface of the green space.	10	36	5*	89.2%	84.2%
Condition	Path	Rate the overall condition of the surface of the path or path segment.	4	15	5*	87.5%	87.5%
Condition	Playground	How much deterioration or corrosion is evident on the play equipment?	3	11	5*	63.9%	47.2%
Condition	Sports Field	Rate the condition of the field.	8	29	5*	83.3%	72.9%
Esthetics	Target Area	Rate the appeal of the view from within the Target Area.	8	30	5*	81.3%	81.3%
Esthetics	Green Space	What portion of the Green Space could potentially be in the shade?	10	36	5	69.2%	33.3%
Features	Target Area	Are there any picnic Tables in the Target Area?	8	30	2	100.0%	100.0%
Features	Court	Is the court outdoor or indoor?	3	12	3	100.0%	100.0%

Domain	Geographic Area	Item	*N* (sites)	*N* (obs)	# Possible Response Options	Individual Item Reliability	Individual Item Validity
Features	Green Space	Describe the surface area of the Green Space/Open Area.	10	36	2	100.0%	100.0%
Features	Path	What is the surface of the path or path segment made of?	4	15	4	75.0%	75.0%
Features	Playground	What type of playground equipment is present: Climbing Apparatus	3	11	2	91.7%	91.7%
Features	Sports Field	What structures are present on the sports field: Scoreboard	8	29	2	96.9%	90.6%
Safety	Target Area	How many of the restrooms are gender-labeled?	8	30	3	83.3%	77.1%
Safety	Playground	If playground surfacing is a loose material, how deep is it?	3	11	3	83.3%	83.3%

Note. *Assessed on a 5-point scale, then dichotomized for this analysis.

Discussion

While several studies have attempted to measure the suitability of the general environment for walking and bicycling,[16, 20, 21] no similar audit instruments have been developed for assessing the specific characteristics of park environments for physical activity. This study provides information on the development, reliability, and validity of an original tool designed and guided by a theoretical model.[17]

In general, inter-rater reliability was good, whether assessed by domain or geographic area; that is, there is high reliability in every domain and geographic area. Validity in all geographic areas and domains was also high, albeit somewhat lower than reliability (particularly in the domains of Condition and Esthetics and in the geographic area of Playground). The finding that raters tended to be in higher agreement amongst themselves than they were when measured against the gold standard raters was not unexpected. This could be explained by the fact that the gold standard raters conducted their assessments 2 wk prior to those of the rest of the observers; hence, any changes that took place in the parks during those 2 wk, particularly for time-sensitive items, could affect the ability to measure validity. For example, the item asking "How much litter is present in the Target Area?" had 74% reliability but only 55% validity. Similarly, the item asking the observer to "Rate the general cleanliness of the restrooms" had 76% reliability and 67% validity. It is reasonable to expect, however, that the amount of litter or the cleanliness of a

restroom can change significantly in a 2-wk period, as they both depend on routine maintenance occurring on a frequent basis. Future research should explore the types of items that are likely to vary within a short time period and determine how many observations over time would be necessary to obtain more stable estimates of these characteristics.

Some items, particularly those asking observers to estimate the amount of shade in an area, performed very poorly and will be modified in the final instrument. Another technique to assess shade, such as the Vertical Sighting Tube method used in the field of ecology,[22] may be more objective and ultimately more reliable and valid. This technique involves observers systematically and with a random starting point locating transects in a Target Area and then locating a calculated number of points to sample on each transect. At each sampling point, the observer looks directly through a sighting tube; if the view through the tube contains foliage or leafless branches, this constitutes a "hit." If the view is devoid of leaves and branches, this constitutes a "miss." The number of hits and misses is then tallied and the percent of potential shade (ratio of hits to misses) is calculated.[22] In a comparison of various techniques to measure tree coverage, the Vertical Sighting Tube method was deemed both the most accurate and the most efficient.[23] This technique can be used in any season and avoids the "time-dependent" bias mentioned above as any natural growth, whether it is leaves or branches, counts as shade.

Items assessing the condition of features or the esthetics of an environment tended to have the lowest reliability and validity ratings. This is likely due to the majority of the items in these domains being inherently subjective. Pikora et al. also found lower reliability among subjective items in the SPACES audit instrument, suggesting that scores varied based on observers' previous experiences.[20] Although these items scored lower than the less subjective items, they are still likely to be important correlates of physical activity and therefore should not be eliminated entirely. Individual ratings on items in the Condition and Esthetics domains were similar to ratings in other domains. Thus, we feel it is possible to design a more comprehensive training program which would incorporate multiple structured practice observations in varied park environments so that observers share more common experiences. Future work on the instrument will focus training in this direction.

Conversely, the Features domain had both the highest inter-rater reliability and validity agreement. This domain contains numerous items asking observers to denote presence or absence of various features. These types of items intrinsically leave little room for interpretation by the observer and therefore helped produce higher agreement values.

A limitation of this analysis is the grouping and summarizing of individual items on the instrument. We attempted to group items by logical domains and geographic areas according to our previously defined conceptual model.[17] This resulted in some groups having significantly more or fewer items than others (e.g., Green Space contained 4 items compared to 85 items in Target Area), thus affecting summarized results. For this reason we grouped items both into domains and geographic areas and found that overall averages varied little despite the method of grouping.

Another limitation to this study is that the BRAT-DO instrument was developed in a single location and has not yet been tested in other geographic areas of the country. It will likely need to be adapted to other locations with different climates

and different local characteristics, necessitating a collaborative approach to develop a truly generalizable instrument. The instrument is flexible enough that additional modules covering other Activity Areas can also be added easily.

One major lesson learned from this study is the importance of good training. Ideally, training should involve both classroom time and multiple field practices. Time spent in a classroom setting going over both the study objectives and the intricacies of each individual item on the instrument is necessary for a clear understanding of concepts, definitions, and procedures. This should be followed by substantial practice time in a variety of park settings, so that data collectors are exposed to multiple types of park environments and various conditions while still being supervised. In fact, with every subsequent round of testing of the BRAT-DO, we have increased the amount of time devoted to training. Our current recommendation is to spend a morning in the classroom setting and an afternoon or more in the field doing supervised practice. Practice time should incorporate simultaneous environmental assessments by the data collectors and by the experts, followed by a comparison and discussion of each observer's responses. In this way, observation skills can be honed and data collectors can begin to approach subjective items from a consistent and standardized knowledge base.

Another important lesson is the necessity of good maps for use in the field and training on their use. It is possible that some observers got lost and mistakenly assessed an environment outside of the study area, contributing to some of the low scores. Maps of study areas should be created using GIS and aerial photography to ensure accuracy, and major landmarks should be clearly noted on each map. Ample time during the training should be spent on recognizing landmarks and identifying appropriate boundaries.

Future work on the BRAT-DO will include all the modifications suggested above. Because the number of items on the BRAT-DO is currently very high, further development of scales and summary scores based on the theoretical model and statistical analysis will be explored to simplify data analysis. Confirmatory factor analysis could be used to determine if variables are statistically associated with the theoretical domains. This technique could help reduce the overall number of variables by removing ones not associated with any of the theoretical domains. Additionally, correlation matrices of the variables within each Activity Area could be examined to measure the strength of the linear relationship between variables. Variables with strong linear relationships to other variables could be removed from the tool due to the redundancy of information obtained from them. Future work will also focus on combining the BRAT-DO with the other BRAT components (the Informant Interview, aerial photography, GPS, and archival data), such that each method works seamlessly together, in the field, during desktop processing, and in analysis. We will also investigate the possibility of using personal digital assistants to capture the direct observation data in the field. This will allow for simultaneous data collection and data entry and may easily be combined with the GIS component of the BRAT methods. Finally, we will be using the BRAT instruments in a study designed to measure physical activity in parks. This will allow us to test associations between the various park environmental characteristics and physical activity, and thus inform the discussion of how to engineer physical activity back into our daily lives.

Conclusions

While this direct observation instrument provides a method for collecting reliable and valid data on park environments, it is important to remember that not all relevant characteristics can be captured adequately and objectively with one methodology. Other methodologies such as GIS, informant interviews, and the use of archival and survey data should also be explored. This project will continue to modify and refine the BRAT-Direct Observation instrument, as well as further develop and incorporate the use of these other methodologies. The BRAT-Direct Observation tool provides a methodology that allows rigorous and detailed studies to be conducted that evaluate associations between specific park characteristics and physical activity among individuals.

Acknowledgments

The authors wish to thank S. Jones, F. Hamilton, R. Scribner, R. Meriwether, and students and teaching assistants in Dr. Gustat's Survey Methodology classes. The BRAT-DO and supporting materials are available at http://www.publichealth.lsuhsc.edu/bedimo-rung/research/index.htm.

References

1. Division of Adult and Community Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System Online Prevalence Data, 1995-2003. Available at: http://apps.nccd.cdc.gov/brfss/display.asp?cat=PA&yr=2003&qkey=4418&state=US. Accessed May 27, 2005.
2. National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. YRBSS Youth Online: Comprehensive Results, 2003. Available at: http://apps.nccd.cdc.gov/yrbss/QuestYearTable.asp?path=byHT&ByVar=CI&cat=6&quest=509&year=2003&loc=XX. Accessed May 27, 2005.
3. Killingsworth R, Earp J, Moore R. Supporting health through design: challenges and opportunities. *Am J Health Promotion.* 2003;18(1):1-2.
4. Harnik P. *Inside City Parks*. Washington: Urban Land Institute; 2000.
5. Sallis JF, Johnson MF, Calfas KJ, Caparosa S, Nichols JF. Assessing perceived physical environment variables that may influence physical activity. *Res Q Exerc Sport.* Dec 1997;68(4):345-351.
6. Sallis JF, Bauman A, Pratt M. Environmental and policy interventions to promote physical activity. *Am J Prev Med.* Nov 1998;15(4):379-397.
7. Brownson RC, Baker EA, Housemann RA, Brennan LK, Bacak SJ. Environmental and policy determinants of physical activity in the United States. *AJPH.* Dec 2001;91(12):1995-2003.
8. Wilcox S, Castro C, King AC, Housemann RA, Brownson RC. Determinants of leisure time physical activity in rural compared with urban older and ethnically diverse women in the United States. *J Epidemiol Community Health.* Sep 2000;54(9):667-672.
9. Sallis JF, Hovell MF, Hofstetter CR. Predictors of adoption and maintenance of vigorous physical activity in men and women. *Prev Med.* Mar 1992;21(2):237-251.
10. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc.* May 2000;32(5):963-975.

11. Booth ML, Owen N, Bauman A, Clavisi O, Leslie E. Social-cognitive and perceived environment influences associated with physical activity in older Australians. *Prev Med.* Jul 2000;31(1):15-22.
12. Sternfeld B, Ainsworth BE, Quesenberry CP. Physical activity patterns in a diverse population of women. *Prev Med.* Mar 1999;28(3):313-323.
13. MacDougall C, Cooke R, Owen N, Willson K, Bauman A. Relating physical activity to health status, social connections and community facilities. *Aust New Zealand J Public Health.* Oct 1997;21(6):631-637.
14. King AC, Castro C, Wilcox S, Eyler AA, Sallis JF, Brownson RC. Personal and environmental factors associated with physical inactivity among different racial-ethnic groups of U.S. middle-aged and older-aged women. *Health Psychol.* Jul 2000;19(4): 354-364.
15. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity. *Am J Prev Med.* Apr 2002;22(3):188-199.
16. Moudon AV, Lee C. Walking and bicycling: an evaluation of environmental audit instruments. *Am J Health Promotion.* 2003;18(1):21-37.
17. Bedimo-Rung AL, Mowen AJ, Cohen DA. The significance of parks to physical activity and public health: a conceptual model. *Am J Prev Med.* 2005;28(2S2):159-168.
18. Last JM, ed. *A Dictionary of Epidemiology.* 4th ed. New York: Oxford University Press; 2001.
19. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York: John Wiley & Sons; 1981.
20. Pikora TJ, Bull FCL, 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.
21. 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 Promotion.* 2003;18(1):38-46.
22. Robards T. *Instructions for WLPZ Canopy/Surface Cover Sampling* 1999.
23. Robards T, Berbach WB, Cafferata PH, Valentine BE. A comparison of techniques for measuring canopy in watercourse and lake protection zones. *California Forestry Note.* 2000;115.