Journal of Physical Activity and Health 2006, 3, Suppl 1, S190-S207

Measuring Physical Environments of Parks and Playgrounds: EAPRS Instrument Development and Inter-Rater Reliability

Brian E. Saelens, Lawrence D. Frank, Christopher Auffrey, Robert C. Whitaker, Hillary L. Burdette, and Natalie Colabianchi

Background: Reliable and comprehensive measurement of physical activity settings is needed to examine environment-behavior relations. *Methods:* Surveyed park professionals (n = 34) and users (n = 29) identified park and playground elements (e.g., trail) and qualities (e.g., condition). Responses guided observational instrument development for environmental assessment of public recreation spaces (EAPRS). Item inter-rater reliability was evaluated following observations in 92 parks and playgrounds. Instrument revision and further reliability testing were conducted with observations in 21 parks and 20 playgrounds. *Results:* EAPRS evaluates trail/path, specific use (e.g., picnic), water-related, amenity (e.g., benches), and play elements, and their qualities. Most EAPRS items had good-excellent reliability, particularly presence/number items. Reliability improved from the original (n = 1088 items) to revised (n = 646 items) instrument for condition, coverage/shade, and openness/visibility items. Reliability was especially good for play features, but cleanliness items were generally unreliable. *Conclusions:* The EAPRS instrument provides comprehensive assessment of parks' and playgrounds' physical environment, with generally high reliability.

Key Words: recreation facilities, activity, trails, active living

A better understanding of environments in which physical activity occurs or could occur is critical to increasing physical activity. Public parks are among the most common settings in which individuals engage in physical activity.[1] Active adults are more likely to use community parks than their inactive counterparts.[2] Socio-demographic variables, including age, gender, and ethnicity, have been examined

---

Saelens is with the Cincinnati Children's Hospital Medical Center and University of Cincinnati College of Medicine, Cincinnati, OH. Frank is with Lawrence Frank & Co., Atlanta, GA. Auffrey is with the University of Cincinnati, Cincinnati, OH. Whitaker is with Mathematica Policy Research, Inc., Princeton, NJ. Burdette is with the Children's Hospital of Philadelphia and University of Pennsylvania School of Medicine, Philadelphia, PA. Colabianchi is with Case Western Reserve University, Cleveland, OH.

as predictors of park use, but these variables have only small predictive power for who and how often individuals use parks for active recreation.[3,4] Other potentially critical factors, including physical environment factors of the existence and quality of park elements (e.g., walking path) and the neighborhood and built environment context in which parks exist, have not been systematically evaluated as park use predictors. However, the presence of specific park and playground features are the factors that parents rate as important in play space choice for their children regardless of familial demographic factors.[5]

Residents' proximity to parks has been examined as an environmental indicator of park access,[1,6] with greater proximity related to greater use.[7,8] Proximity constitutes only one of many park and playground factors potentially relevant to use and physical activity within these settings. Perhaps based on the lack of instruments to quantify park and playground elements and quality, studies examining the relation between park use and physical activity have generally by default considered all parks and playgrounds to have the same elements and qualities, despite the awareness that they may differ substantially on these characteristics.

Prior to examining the relation between physical activity and environment, it is critical to develop psychometrically sound instruments to assess the physical environment within and surrounding behavior settings. For example, such measurement development has occurred in observational[9] and resident[10] assessment of neighborhood infrastructure for walking and cycling. There are numerous texts to guide park and recreation planners in park and playground design,[11-13] but few tested instruments that attempt to quantify the physical environment of parks and playgrounds. Other investigators have recognized this lack of measurement and recently evaluated new instruments to assess recreation facility presence and quality.[14,15] However, these instruments are generally focused on only the recreation facilities within parks (and thus may not capture the other aspects of parks that may promote use of the available recreation facilities), provide only general and a few quality ratings across the park rather than for each element (e.g., the cleanliness of the park is evaluated, but not separately for each area or amenity in the park), have only been tested in a relatively small number of parks (< 50), have not been tested for reliability or validity to our knowledge by investigators not associated with the development of the instrument. The aim of the present study was to develop and evaluate the inter-rater reliability of a comprehensive instrument that would allow observers to characterize the physical environments within public parks and playgrounds. Further aims included modifying the instrument based on the initial reliability testing and re-testing reliability in a subset of the same parks and playgrounds with different raters and also in a set of different playgrounds by raters trained by another investigator in the use of the instrument.

Methods

To develop and ensure the instrument included all possible park and playground elements, park and recreation professionals and frequent park users were recruited to complete two open-ended surveys about the physical environment characteristics generally found in public recreation spaces. Survey responses guided instrument creation and inter-rater reliability of the original instrument was evaluated through initial park and playground observation. The instrument was revised based on this information and inter-rater reliability was evaluated again in local parks and playgrounds and at another site by a different group of investigators and raters. Survey respondents provided written consent for participation. This study was approved by the Institutional Review Board of the Cincinnati Children's Hospital Medical Center.

Participants

Park and Recreation Professionals. Professionals (n = 306) from the Ohio Parks and Recreation Association (OPRA) were mailed a study introduction letter, with 44 professionals (14.4% recruitment rate) completing the first survey and 34 of these professionals completing the second survey (77.3% retention rate).

Frequent Park Users. To obtain a sampling frame consisting of residents who might be more likely to be park users, census block groups within Hamilton County (part of the Greater Cincinnati area) were evaluated for net residential density, average household size, and the proportion of households with children under the age of five. Forty-five block groups with higher net residential density (4.8 to 7.6 dwelling units per acre) and comparatively larger household size (> 2.5 persons) and high proportion of households with young children (18 to 67% of households with at least one child < 5 y old) were identified. Nine of these block groups with the greatest geographic diversity (distance between different block groups) were selected. Address information was obtained from a marketing company and 1000 residents total from these nine block groups were sent a study introduction letter. Interested residents (n = 121) returned postcards with phone contact and park use information (12.1% recruitment rate). Based on the criteria of having visited parks 25 or more times during the past year and visiting three or more different parks, 50 residents were classified as frequent park users. These frequent park users were asked to sign a consent form and complete two written surveys. Forty frequent park users (80.0% response rate) completed the first survey and 29 of these residents (72.5% retention rate) also completed the second survey. Demographic and park visitation information for the professionals and frequent park users completing both surveys is provided in Table 1.

Procedures

Observation Instrument Development. The first written survey completed by professionals and park users included open-ended queries about which environmental elements were 1) needed to do various park-related activities (e.g., "What would be important to have in a park in order for people to walk or exercise in the park?"), 2) expected to be different in small compared to large parks (e.g., "What elements are typically found in a neighborhood park"?), 3) expected to encourage or discourage park use, and 4) needed for park-related activities specific to various child age groups. Recognizing that responses would overlap, this format was used to elicit a comprehensive set of physical elements. Responses within and across respondents were consolidated (and duplicates eliminated) to create a comprehensive list of physical elements of parks and playgrounds.

Table 1 Demographic Information for Park and Recreation Professionals and Frequent Park Users

	Park & recreation professionals (*N* = 34)	Frequent park users (*N* = 29)
Gender (% male)	70.6	17.2
Age (y, SD)	42.8 (8.5)	37.0 (8.1)
Completed college/university (%)	94.1	62.1
Race (%)		
Caucasian	94.1	96.6
African American	2.9	3.4
Asian	2.9	0
Experience in park/recreation field (y, SD)	17.8 (8.8)	N/A
Child under 5 y old present in household (%)	23.5	72.4
Number of park visits in past year for work (SD)	214.5 (140.2)	N/A
Number of park visits in past year during leisure (SD)	42.3 (42.0)	24.8 (16.7)
Number of different parks visited in past year (SD)	21.4 (13.1)	5.6 (2.5)

The second written survey completed by professionals and park users included open-ended queries to define 1) the constituent elements of the larger physical elements identified in the first survey (e.g., "What are some aspects or elements of swings?") and 2) qualities of the larger and constituent elements (e.g., "What are some trail qualities that contribute to its safety?"). Responses within and across respondents were compiled with (and duplicates removed) first survey responses to create a comprehensive park and playground environment instrument. This instrument contains 1088 items and is titled the Environmental Assessment of Public Recreation Spaces (EAPRS; see Figure 1 for sample section). Elements were rated as present/absent and counted where applicable (see Table 2 for element list). Sub-elements were considered to be elements that were part of larger elements and were rated as present/absent and were often countable (e.g., fitness stations on a trail, lighting on a trail). In contrast, qualities were considered observable, but not necessarily countable (e.g., condition of a piece of an athletic field, content on trail signs). Qualities were rated as present/absent or using 5-point Likert-type scales (see Table 3 for response scales). Some qualities items on the instrument were common to many different element categories and elements, but some were specific to an element category or element (e.g., continuity of a trail). The complete original and revised (see below) EAPRS instrument and instructions are available at http://www.cincinnatichildrens.org/research/div/psychology/fs/fac/brian-saelens.htm.

A. Trails

1. Paved - existence and surface

Aspect	Rating	Scaling	Considerations
Paved trail presence	Yes No		If none present, skip section
Number	Number ____		
Material	Asphalt Concrete		Circle predominant material
Condition	1 2 3	PEX	Standard condition definition; also consider holes, cracks, tree branches under surface
Width	1 2 3		<2 ft (1 adult); 2-5 ft (2 adults; side-walk width); >5 ft (>3 adults)
Cleanliness	1 2 3	NATE	standard cleanliness, plus consider mud, rocks, twigs
Flatness	1 2 3	NATE	
Continuity	Yes No	NATE	Consider change in surface material
Clear from obstruction	1 2 3	NATE	Standard obstruction definition
Coverage/ shade	1 2 3	PER	Consider the entire length of the trail, but not width
Dividing-line stripe	Yes No		

Figure 1—Sample revised EAPRS instrument items for park and playground environment assessment. The complete instrument and ratings instructions are available at http://www.cincinnatichildrens.org/research/div/psychology/fs/fac/brian-saelens.htm. PEX scaling 1=poor, 2=fair, 3=excellent; NATE scaling 1=not at all, 2=somewhat, 3=mostly to extremely; PER scaling 1=0-33%, 2=34-66%, 3=67-100%.

Initial Park and Playground Selection. A representative sample of parks and elementary school-based playgrounds within Hamilton County was sought to conduct observations. The Cincinnati area Geographic Information System (CAGIS) database (http://www.cagis.org/index.html) contains 1569 distinct land parcels with an "open space" designation. To specify the public parks, the following inclusion criteria were applied: 1) public and open to anyone (e.g., private clubs and community member-only facilities were excluded), 2) owned and maintained by a non-profit organization such as a city or village (e.g., private commercial facilities were excluded), 3) no other designated primary function (e.g., cemetery, church) or exclusive recreational function (e.g., golf course). This resulted in 596 eligible parks. CAGIS includes 128 eligible public elementary schools, all eligible for observation.

Parks were classified into urban (> 12 households per residential acre), urban periphery (3.0 to 11.99 households per residential acre), or suburban (≤ 2.99 households per residential acre) location. Location was defined as the block group in which the park's centroid was located. Parks were also classified into small (< 5 acres; ~ 50% of parks), moderate (5 to 50 acres; ~ 39% of parks), and large parks

Table 2 Elements on the Park and Playground Instrument

Trail/path	Designated and specific use areas	Water areas	Other amenities and facilities	Playground equipment and fields and courts
Paved trail	Open space	Pond/lake	Grills/fire pits	Playset
Unpaved trail	Meadow	Stream/creek	Concessions	Ground surface
Path	Wooded area	Wetland	Vending machines	Things to swing/hang from
	Picnic area	Fishing area	Restrooms	Things to slide down
	Camping site	Boat loading/ beach	Historical markers	Things to climb on/up/ through
	Shelter/pavilion/ gazebo	Swimming pool	Visitor's center	Things to stand or walk on
	Entertainment venue/stage	Fountain	Recreation center	Other equipment (e.g., game panel)
	Area surrounding park	Waterfall	Benches, tables	Climbers
	Wildlife or pet area		Seat walls, bleachers	Swings
	Parking lot		Landscaping	Sandbox or other digging area
			Views inside & outside park	Blacktop games
			Sculpture or other art	Merry-go-round and teeter totter
			Trash cans	Spring toys
			Entrances	Imaginary play structure
			Bike racks	Specific fields (e.g., soccer)
			Sidewalks into park	Specific courts (e.g., basketball)
			Roadways through park	Skate and skateboard area
			Signage, maps, event postings	
			Telephones	

(> 50 acres; ~ 11% of parks). Within each location type (urban, urban periphery, suburban), parks were randomly selected to approximate the distribution of park sizes within the county. Observations were conducted in 80 parks and 12 randomly selected elementary school playgrounds (see Table 4 for park and playground locations and sizes).

Table 3 Qualities and Corresponding Response Categories on the Park and Playground Instrument

Quality	Initial response categories	Revised response categories
Cleanliness/ aesthetics	*Not at all, A little, Somewhat, Mostly, Extremely*	*Not at all, Somewhat, Mostly to Extremely*
Condition	*Poor, Fair, Mediocre, Good, Excellent*	*Poor, Fair, Excellent*
Coverage/shade	*Yes, No*	*Yes, No* or *0-33%, 34-66%, 67-100%*
Dimension	*5-point Likert-type response scale that varies (e.g., trail width options were < 2 ft, 2-5 ft, 5-8 ft, 8-11 ft, >11 ft)*	*3-point Likert-type response scale that varies depending upon element rated (e.g., trail width options were <2 ft, 2-5 ft, >5 ft)*
Openness/ visibility	*Not at all, A little, Somewhat, Mostly, Extremely*	*Not at all, Somewhat, Mostly to Extremely*
Proximity	*<25 ft, 25-50 ft, 51-100 ft, 101-200 ft, > 200 ft*	*<25 ft, 25-100 ft, >100 ft*
Specific qualities	*5-point Likert-type response scale or Yes/No scale depending upon the quality assessed*	*5-point Likert-type response scale or Yes/No scale depending upon the quality assessed*

Table 4 Number (Average Size in Acres ± Standard Deviation) of Parks and School Playgrounds In Which Observations Were Conducted

	Urban	Urban periphery	Suburban
Initial observation			
Small parks (<5 acres)	14 (1.6 ± 1.5)	13 (1.6 ± 1.0)	11 (1.5 ± 1.6)
Moderate parks (5-50 acres)	11 (14.6 ± 8.6)	10 (14.6 ± 6.3)	11 (18.9 ± 9.3)
Large parks (>50 acres)	1 (187.1)	3 (90.4 ± 17)	6 (129.7 ± 81.7)
School playgrounds	9 (1.5 ± 1.0)		3 (10.8 ± 2.9)
Second observation			
Small parks (<5 acres)	1 (0.9)	1 (1.6)	--
Moderate parks (5-50 acres)	5 (15.8 ± 8.6)	6 (14.0 ± 6.7)	5 (21.4 ± 8.1)
Large parks (>50 acres)	--	2 (84.0 ± 17)	1 (138.0)
Other site observation			
School playgrounds	20 (N/A)	--	--

Initial Park and Playground Observation. All observations were conducted during daylight hours by the same two raters working independently. Raters did not conduct observations on the same day as each other, but the two observations for any given park or school playground occurred within 4 wk of each other. Raters were trained in the use of the instrument by review of the instrument instructions and by in vivo training provided by the first author (B.E.S.) in three parks (not included in analysis). Raters had no prior experience or training in parks and recreation or environmental assessment. Minor modifications to the instructions and further review of the instructions were made after the first 10 parks were rated by both raters (included in analysis). The observations averaged 67.3 min, ranging from 10 to 258 min.

Instrument Modification and Second Inter-Rater Reliability Testing. Based on the findings from the initial observation and raters' comments, modifications were made to the instrument. Modifications to items with low reliability included changing the instructions/definitions to provide more clarity and re-scaling ordinal items to a 3-point Likert-type response format (see Table 3 for the revised response scales). Items that assessed presence/absence of a larger element and its sub-elements and qualities were eliminated if these larger elements were not observed three or more times. This included "camping site," "wetland," "pet area," "fishing area," "waterfall," "boat loading/beach," "concessions," "visitor's center," "viewing areas for views within the park," "sandbox or other digging area," "merry-go-round and teeter totter," and "skate and skateboard area". The revised instrument contains 646 items. The revised instrument was used by two new independent raters to observe 21 parks in the Greater Cincinnati area selected from the parks rated as part of the initial observation. Training for these raters was similar to the training for the initial raters.

Other Site Observation Inter-Rater Reliability Testing. The primary author (B.E.S.) trained the last author (N.C.) in the use of the revised instrument by in vivo instrument review in five playgrounds in the Cincinnati area. The last author then trained two raters in instrument use by review and in vivo use at local Cleveland playgrounds. After each element in a section had been assessed in this training phase, ratings were compared and discrepancies reconciled. Twenty playgrounds in Cleveland were then evaluated by two raters independently, but on the same day, on weekdays during school hours. Open spaces immediately adjacent to the playground were included in the assessment; non-adjacent open spaces (i.e., separate parking lots) were not included unless the area was designed for activity (i.e., blacktop games).

Data Analysis

Inter-rater reliability was assessed at the individual item level. Items with dichotomous values (yes/no) were assessed using the kappa statistic, with the criteria for kappa values established as "good to excellent" (≥ 0.60), "moderate" (0.41-0.60), or "poor" (≤ 0.40) based on previously proposed classification systems.[16] Since kappa calculation relies on the existence of some variability in ratings made within each rater, percent agreement was examined when there was no response variability among one or both raters, with the percent agreement criteria established as "good to excellent" (≥ 75%), "moderate" (60-74%), or "poor" (< 60%).

Two-way random effects single rater intraclass correlation coefficients (ICC) were calculated for items with ordinal (e.g., Likert-type ratings) or continuous (e.g., number of a certain element) responses. The ICC, rather than other metrics such as Pearson product moment correlations, was selected because it is sensitive to both association and absolute differences between raters.[17, 18] The two-way random effects model was selected because raters were considered a random sample of all potential and future raters and all parks and playgrounds were rated for each item by raters.[17, 18] The single measure ICC was selected because the unit of analysis is the individual rater, rather than an average across raters. The magnitude of ICCs was classified similarly to kappa values.

Kappa and ICC values can be low and suggest poor agreement when there is little variability by one or both raters. To account for this potential misclassification of items as having low inter-rater reliability in the presence of little or no response variability, percent agreement was examined for moderate and poor ICC and kappa values. For example, if an item was rated a "5" by both raters in all but one of 10 parks and rated a "5" and a "4" respectively by raters in the 10th park, the ICC would be 0.00, but the percent agreement would be 90%. Dichotomous response 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 high reliability.

Items for which there were less than three available ratings (n = 288 items for the initial instrument; n = 140 for the revised instrument) were considered insufficiently sampled and inter-rater reliability was not calculated (e.g., there were no beach areas observed by either rater, so reliability for that element and related elements and qualities was not calculated). This resulted in n = 800 items evaluated for reliability for the initial observations and n = 506 items evaluated for reliability for the second observation using the revised instrument. The other site observations, which used only the play equipment sections of the revised instrument, had 103 items available for reliability testing. To avoid inflation of reliability estimates, analyses were conducted hierarchically, such that if an element category or larger element was not rated as present, their constituent elements and qualities could not be rated and thus were not considered in the reliability analysis. For example, if a paved trail was not rated as present, items related to places to sit along the paved trail were assigned missing values (and thus not considered in analyses) rather than rated as not present. For ease of reporting, elements and sub-elements were considered "presence/number" items and quality items were grouped into one of seven categories (see Tables 5 to 8). Item-specific reliability values for the original and revised instrument are available at the website listed above.

Results

Based on the initial instrument and observations, approximately 69.6% of the 800 items tested had reliability values in the good-excellent range or high percent agreements. The revised instrument had 65.6% of the 506 items tested with similarly high reliability or percent agreement values. A summary of the item inter-rater reliability values is provided in Tables 5 to 8.

Table 5 Summary of Inter-Rater Reliability of Paved and Unpaved Trail and Path Items

	Initial observations					Second observations				
	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]
Paved trail										
Presence/number	23	17	6	0	6	24	13	2	9	3
Cleanliness/aesthetics	6	0	2	4	2	5	1	0	4	0
Condition	6	1	1	4	0	6	3	1	2	2
Coverage/shade	2	1	0	1	0	2	1	1	0	1
Dimension	4	3	0	1	1	4	0	1	3	2
Openness/visibility	2	0	0	2	0	2	1	0	1	0
Proximity	7	5	0	2	1	5	0	2	3	2
Specific qualities	21	13	4	4	2	4	3	0	1	1
Total	71	40	13	18	12	52	22	7	23	11
Unpaved trail										
Presence/number	18	10	5	3	4	21	13	2	6	2
Cleanliness/aesthetics	4	0	0	4	0	5	1	0	4	3
Condition	5	1	0	4	1	6	1	1	4	3
Coverage/shade	2	2	0	0	--	2	0	0	2	2
Dimension	4	2	1	1	1	4	1	1	2	1
Openness/visibility	1	0	1	0	0	2	1	0	1	0
Proximity	7	4	0	3	0	5	1	0	4	0
Specific qualities	11	9	1	1	0	8	5	3	0	3
Total	52	28	8	16	6	53	22	7	23	14
Path										
Presence/number	4	3	1	0	0	4	0	0	4	1
Cleanliness/aesthetics	1	0	0	1	1	1	0	0	1	0
Condition	4	0	0	4	4	3	0	0	3	2
Coverage/shade	1	0	0	1	1	1	0	0	1	0
Dimension	1	0	0	1	1	1	0	0	1	0
Proximity	3	0	1	2	2	3	0	0	3	0
Specific qualities	4	2	1	1	0	3	2	1	0	0
Total	18	5	3	10	4	16	2	1	13	3

Note. G-E[a] = good to excellent reliability (≥ 0.60); Mod[b] = moderate reliability (≥ 0.40-0.59); Poor[c] = poor reliability (< 0.40); PA[d] = among the items considered "moderate" or "poor" by kappa or ICC, the number of items with percent agreement for dichotomous response items ≥ 75% and for ordinal or continuous response items ≥ 60%.

Table 6 Summary Inter-Rater Reliability of Designated and Specific Use Area and Water Area and Feature Items

	Initial observations					Second observations				
	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]
Designated and specific use areas										
Presence/number	55	38	9	8	8	37	17	4	16	4
Cleanliness	10	1	4	5	1	7	0	1	6	4
Condition	8	1	4	3	1	6	3	0	3	2
Coverage/shade	3	1	1	1	0	1	1	0	0	--
Dimension	6	4	1	1	1	4	1	0	3	0
Openness/ visibility	2	1	0	1	0	1	0	1	0	1
Proximity	14	6	4	4	3	7	1	1	5	2
Specific qualities	33	13	8	12	10	16	9	4	3	5
Total	131	65	31	35	24	79	32	11	36	18
Water areas										
Presence/number	28	17	2	9	3	11	8	2	1	2
Cleanliness	6	2	1	3	2	1	0	1	0	1
Condition	6	3	1	2	1	1	0	1	0	0
Coverage/shade	1	1	0	0	--	0	0	--	--	--
Dimension	7	3	2	2	1	2	1	1	0	1
Proximity	4	2	0	2	2	0	0	--	--	--
Specific qualities	13	8	4	1	2	7	4	2	1	1
Total	65	36	10	19	11	22	13	7	2	5

Note. G-E[a] = good to excellent reliability (≥ 0.60); Mod[b] = moderate reliability (≥ 0.40-0.59); Poor[c] = poor reliability (< 0.40); PA[d] = among the items considered "moderate" or "poor" by kappa or ICC, the number of items with percent agreement for dichotomous response items ≥ 75% and for ordinal or continuous response items ≥ 60%.

Paved and Unpaved Trail and Path Items

A summary of reliability values for paved and unpaved trail and path items is provided in Table 5. In the initial observation, presence/number ratings for paved and unpaved trail and path items were highly reliable, with 42 of 45 (93.3%) falling in the moderate range or above. Further, 10 of the 15 presence/number items (66.7%) in the poor or moderate range for kappa or ICC had high levels of percent agreement. The reliability for the dimension, proximity, and other quality items

Table 7 Summary Inter-Rater Reliability of Other Park Amenities and Facilities

	Initial observations					Second observations				
	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]
Presence/number	98	70	13	15	14	56	29	13	14	12
Cleanliness	34	6	8	20	8	22	6	1	15	4
Condition	27	4	7	16	6	24	4	0	20	10
Coverage/shade	6	1	1	4	1	5	5	0	0	--
Dimension	8	0	1	7	4	8	2	1	5	3
Openness/ visibility	5	2	0	3	0	1	0	0	1	1
Proximity	12	4	2	6	3	5	0	0	5	3
Specific qualities	81	41	9	31	14	34	21	4	9	1
Total	271	128	41	102	50	155	67	19	69	34

Note. G-E[a] = good to excellent reliability (≥ 0.60); Mod[b] = moderate reliability (≥ 0.40-0.59); Poor[c] = poor reliability (< 0.40); PA[d] = among the items considered "moderate" or "poor" by kappa or ICC, the number of items with percent agreement for dichotomous response items ≥ 75% and for ordinal or continuous response items ≥ 60%.

were also generally in the moderate to good-excellent range or had high percent agreement. In contrast, the cleanliness, condition, and openness/visibility items had mostly low reliability values.

The second set of observations, using the revised instrument, resulted in some improvements in the reliability of paved and unpaved trail condition, coverage/shade, and openness/visibility items, but the cleanliness/aesthetics items remained largely unreliable. There was little improvement in reliability for any of the path items using the revised instrument, and the dimension and proximity items appeared to have decreased reliability using the revised instrument.

Designated and Specific Use and Water Area Items

A summary of reliability values for designated and specific use and water area items is provided in Table 6. Items on the original instrument assessing designated and specific use areas showed generally high reliability for presence/number, coverage/shade, and dimension. Reliability was again low for cleanliness and condition items, with mixed reliability for openness/visibility, proximity, and specific qualities items. The revised instrument resulted in some improvements in condition and other quality item reliability, but continued low reliability for cleanliness items and decreased or no change in reliability for dimension or proximity items from the original instrument. Water area items evidenced generally high reliability for the presence/number items and the various qualities. The revised instrument items demonstrated similar high reliability for water areas across item content, although there were limited items evaluated due to low frequency of water areas in the parks rated in the second observation.

Table 8 Summary Inter-Rater Reliability of Play Equipment and Field and Court Items

	Initial Observations					Second Observations					Other Site Observations				
	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]	No. of items	G-E	Mod[b]	Poor[c]	PA[d]	No. of items	G-E[a]	Mod[b]	Poor[c]	PA[d]
Presence/number	76	59	11	6	14	54	39	6	9	4	44	41	2	1	3
Cleanliness	20	5	6	9	5	14	4	1	9	2	12	7	3	2	3
Condition	19	3	7	9	3	12	2	0	10	6	11	5	1	5	4
Coverage/shade	5	1	0	4	2	3	2	0	1	0	3	2	0	1	0
Dimension	12	2	1	9	5	9	4	1	4	4	7	1	2	4	5
Openness/visibility	5	1	1	3	3	2	1	0	1	1	2	1	0	1	1
Proximity/location	9	2	6	1	2	8	2	0	6	0	4	1	2	1	2
Specific qualities	46	28	4	14	13	27	13	6	8	5	20	14	3	3	3
Total	192	101	36	55	47	129	67	14	48	22	103	72	13	18	21

Note. G-E[a] = good to excellent reliability (≥ 0.60); Mod[b] = moderate reliability (≥ 0.40-0.59); Poor[c] = poor reliability (< 0.40); PA[d] = among the items considered 'moderate' or 'poor' by kappa or ICC, the number of items with percent agreement for dichotomous response items ≥75% and for ordinal or continuous response items ≥ 60%.

Amenities and Facilities

A summary of reliability values for other park amenities and facilities items is provided in Table 7. Presence/number items on the original instrument showed generally high reliability for park amenities and facilities, as did specific qualities items. The cleanliness, condition, coverage/shade, dimension, and openness/visibility items generally had low reliability, with proximity/location items having mixed reliability values. The reliability for the condition, coverage/shade, and dimension items was improved using the revised instrument, with little change for the presence/number, cleanliness, proximity, and specific qualities items.

Play Equipment and Field and Court Items

A summary of reliability values for play equipment and field and court items is provided in Table 8. Presence/number items again demonstrated generally high reliability for play equipment and field and court items, as did openness/visibility and specific qualities items on the original instrument. Cleanliness, condition, and proximity/location items evidenced low reliability on the original instrument and observations. The reliability of the condition, dimension, and openness/visibility items generally improved in the second observation with the revised instrument, but cleanliness item reliability remained low and proximity/location, presence/number, and specific qualities item reliabilities were lower than for the original instrument and observations.

The other site observations using the revised instrument demonstrated generally high reliability across all presence/number and quality items that were part of play equipment and field and court areas. Indeed, 93 of 103 of the items (90.3%) were in the good-excellent reliability range or had high percent agreement values.

Reliability Within Item Domains

Overall, high reliability was obtained for the presence/number and specific qualities items across the various park areas and features, with 87.1% (71.0%) and 75.1% (75.8%) of these items on the original (and revised) instruments respectively having good-excellent reliability or high percent agreement. Condition, coverage/shade, and openness/visibility items had only 34.7%, 50%, and 46.7% of their items respectively on the original instrument meet this criterion. However, these percentages increased to 65.5%, 85.7%, and 75.0% respectively for the revised instrument and second observation. The dimension items retained a generally high level of reliability across the original and revised instrument, with 66.7% and 62.5% of these items respectively having good-excellent reliability or high percent agreement values. In contrast, proximity item reliability values dropped from 64.3% meeting this criterion on the original instrument to 33.3% meeting this criterion on the revised instrument. Cleanliness/aesthetic items had generally low reliability on both the original and revised instrument, with only 39.5% and 47.3% of these items respectively meeting this criterion. For the other site observations conducted on school playgrounds using the revised instrument, reliability was routinely high (> 66.6% of items within each domain having good-excellent reliability or high percent agreement) within and across the different domains.

Discussion

There is a growing need to have available reliable and valid instruments to assess physical environmental factors that may be impacting physical activity and other health behaviors.[19,20] Given that parks and playgrounds represent a frequent setting or at least opportunity for physical activity,[1,21,22] it is important to reliably characterize the physical environment that exists within parks and playgrounds in order to augment the existing environmental metrics of park and playground proximity or accessibility. This characterization of physical environments within parks and playgrounds will allow for more systematic evaluation of the relations among environment, park use, and individuals' physical activity[23] and may help further elucidate relations between parks and other health variables (e.g., obesity).[24]

Given the early status of the empirical literature regarding public recreation spaces and physical activity, it is unknown which aspects of parks' or playgrounds' physical environments are most relevant to physical activity and overall health. Therefore, we sought to develop a comprehensive instrument that would characterize aspects of these spaces that are specifically designed for active use (e.g., playground equipment, trails) and their supporting elements (e.g., drinking fountains), but to also evaluate other aspects that may or may not ultimately influence public recreation space use and physical activity. The open-ended qualitative surveying of park and recreational professionals and frequent park users appears to have accomplished this task, resulting in a considerable number of elements, sub-elements, and qualities of public recreation space environments and thus contributing to solid content validity of the developed instrument. Comprehensiveness of a park environment instrument may be particularly important in evaluating the relation between park environment and physical activity across the life span. For example, different physical environment factors may be more or less attractive and engaging to children versus middle aged adults versus older adults.[25]

The inter-rater reliability testing revealed generally high reliability for ratings for the EAPRS instrument. The most consistently high reliability was obtained for items assessing presence and number of elements and sub-elements, especially for the paved trail and play equipment and fields and courts items. The specific qualities items (e.g., continuity of a trail, drainage of slide landing area) were also generally rated reliably. The qualities rated across various element categories, including condition, coverage/shade, and openness/visibility items had modest inter-rater reliability values with the initial observation and the original instrument, but the revised instrument on which some definitions and response scales were modified resulted in significant improvements in the reliability of items in these quality categories. Items that assessed element dimensions (e.g., height of slides) also demonstrated good reliability. Proximity items had generally good reliability on the original instrument that worsened on the revised instrument, suggesting perhaps that the switch from a 5-point to 3-point Likert-type scale for these items was misguided. With the exception of the observations conducted at the other site (which were done by a different investigator and raters using the revised instrument, but only done on school playgrounds), the cleanliness/aesthetics items fared poorly in reliability between raters using either the original and revised instrument.

It is perhaps not surprising the presence/number, dimension, and proximity items were more consistently reliable, as there is less subjective influence on these ratings. More subjective items were on the whole less consistently rated (e.g., cleanliness), although improvements in the definitions and scaling from the original to revised instrument appeared to remedy some of these problems. The temporal nature of some of the elements' qualities could also have influenced reliability. It is noteworthy that the other site school playground observations were conducted by raters on the same day, whereas the initial and second observations were conducted by raters up to 4 wk apart. It is not unexpected that items with more temporal variability such as cleanliness would not be rated consistently over time. This suggests that multiple assessments of these dynamic characteristics may be important to conduct in order to generate a reliable indicator. It is premature to simply eliminate these more subjective and temporal items, as they may be very important determinants of park use and the amount and frequency of physical activity within parks and playgrounds.

One of the limitations of some of the instrument items is the lack of variability in rater response. For example, some of the improvements from the original to the revised instrument in condition items were the higher rate of items with high percent agreement values between raters. This remains a difficult psychometric problem, with the opposing solutions of either accepting low variability in that item or specifying a more fine-grained distinction between item responses to increase reliability that imposes an artificial distinction with little relevance for users (e.g., rating play equipment low on cleanliness if any dirt is on the equipment). Some individual items within domains, that overall showed good reliability, need more development and resulting better instructions. Further, the development sample included individuals who all work and reside in the Midwestern region of the US and the instrument has also only been tested to date within this region. The instrument may require adaptation for other regions, in order to include all park environment elements and qualities. Whereas it is proposed that park and recreation professional and frequent users would be invested in and aware of park environments, others such as non-users may identify elements that prevent them from park use and are not present on the EAPRS instrument. Also, there may be elements and qualities that these park professionals and users did not identify that are part of the intricate process of park planning and design or that are too complex on which to conduct single item ratings.[12] As the EAPRS instrument was designed based solely on professionals' and users' responses and not prior park design literature,[12] such factors may not be part of the instrument, but it was assumed that the responses by park and recreation professionals involved in the development of the instrument would be influenced by this literature. Raters or investigators were also all trained by the first author (B.E.S.), although it is planned with further development and testing that the instrument will be a stand alone instrument that other investigators can use "off the shelf." At this point, it is recommended that investigators test and monitor the reliability of their raters' use of the EAPRS instrument in their own settings.

A comprehensive assessment of the physical environment of parks and playgrounds is potentially important for establishing predictors of park use and perhaps use of specific areas within a park or playground. In addition, other environmental factors, perhaps especially the social context of parks may be important determinants of use and physical activity within public recreation spaces.[26] The context in which parks exist (e.g., safety of neighborhood, accessibility of park from neighborhood) may be an important facilitator or barrier to park use.[27] The EAPRS instrument is a comprehensive direct observation instrument that provides a reliable assessment of the physical environment of public parks and playgrounds.

Acknowledgments

This project was funded by the Active Living Research Program, a national program of The Robert Wood Johnson Foundation. The authors would like to thank the park and recreation professionals and park users who helped with the instrument development, as well as Michelle Park, the Executive Director of the Ohio Parks and Recreation Association.

References

1. Giles-Corti B, Donovan RJ. The relative influence of individual, social and physical environment determinants of physical activity. *Soc Sci Med.* 2002;54:1793-1812.
2. Addy CL, Wilson DK, Kirtland KA, Ainsworth BE, Sharpe P, Kimsey D. Associations of perceived social and physical environmental supports with physical activity and walking behavior. *Am J Pub Health.* 2004;94:440-443.
3. Payne LL, Mowen AJ, Orsega-Smith E. An examination of park preferences and behaviors among urban residents: the role of residential location, race, and age. *Leisure Studies.* 2002;24:181-198.
4. Scott D, Jackson EL. Factors that limit and strategies that might encourage people's use of public parks. *J Park Recreation Admin.* 1996;14:1-17.
5. Sallis JF, McKenzie TL, Elder JP, Broyles SL, Nader PR. Factors parents use in selecting play spaces for young children. *Arch Pediatr Adolesc Med.* 1997;151:414-417.
6. 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.* 2000;31:15-22.
7. Mowen AJ, Confer JJ. The relationship between perceptions, distance, and socio-demographic characteristics upon public use of an urban park "in-fill". *J Park Recreation Admin.* 2003;21:58-74.
8. 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.* 2001;32:191-200.
9. 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.* 2003;23:187-194.
10. Saelens BE, Sallis JF, Black JB, Chen D. Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health.* Sep 2003;93:1552-1558.
11. Mertes JD, Hall JR. *Park, recreation, open space and greenway guidelines*. Ashburn, VA: National Recreation and Park Assn; 1996.
12. Hultsman J, Cottrell RL, Hultsman WZ. *Planning parks for people*. 2d ed. State College, PA: Venture Publishing; 1998.
13. US Consumer Product Safety Commission. *Handbook for public playground safety.* Washington: 2002. p. 325.
14. Cavner MM, Kirtland KA, Evans MH, et al. Evaluating the quality of recreation facilities: development of an assessment tool. *J Park Recreation Admin.* 2004;22:96-114.
15. Broomhall M, Giles-Corti B, Lange A. Quality of public open space tool (POST). August 23, 2004 2004:Accessed on August 23, 2004 at http://www.populationhealth.uwa.edu.au/welcome/research/hpreg/hpreg/programs/parg/reside/pos_audit.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.

17. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. *Stat Med.* 1994;13:2465-2476.
18. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420-428.
19. Sallis JF, Bauman A, Pratt M. Environmental and policy interventions to promote physical activity. *Am J Prev Med.* 1998;15:379-397.
20. Sallis JF, McKenzie TL, Conway TL, et al. Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. *Am J Prev Med.* 2003;24:209-217.
21. Godbey GC, Caldwell LL, Floyd M, Payne LL. Contributions of leisure studies and recreation and park management research to the active living agenda. *Am J Prev Med.* 2005;28:150-158.
22. Moody JS, Prochaska JJ, Sallis JF, McKenzie TL, Brown M, Conway TL. Viability of parks and recreation centers as sites for youth physical activity promotion. *Health Promot Pract.* 2004;5:438-443.
23. 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;28:159-168.
24. Burdette HL, Whitaker RC. Neighborhood playgrounds, fast food restaurants, and crime: relationships to overweight in low-income preschool children. *Prev Med.* 2004;38:57-63.
25. Raymore L, Scott D. The characteristics and activities of older adult visitors to a metropolitan park district. *J Park Recreation Admin.* 1998;16:1-21.
26. Krenichyn K. Women and physical activity in an urban park: Enrichment and support through an ethic of care. *J Environ Psychol.* 2004;24:117-130.
27. Timperio A, Crawford D, Telford A, Salomon J. Perceptions about the local neighborhood and walking and cycling among children. *Prev Med.* 2004;38:39-47.