NIH Public Access
Author Manuscript
J Phys Act Health. Author manuscript; available in PMC 2013 June 21.

Published in final edited form as:
J Phys Act Health. 2011 November ; 8(8): 1117–1123.

How much observation is enough? Refining the administration of SOPARC

Deborah A Cohen[(a)], **Claude Setodji**[(b)], **Kelly R Evenson**[(c)], **Philip Ward**[(d)], **Sandra Lapham**[(e)], **Amy Hillier**[(f)], and **Thomas L. McKenzie**[(g)]

[(a)]RAND Corporation, Santa Monica, CA

[(b)]Dept of Statistics, RAND Corporation, Santa Monica, CA

[(c)]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, NC

[(d)]School of Physical Activity and Educational Services, The Ohio State University Columbus, OH

[(e)]Behavioral Health Research Center of the Southwest, Albuquerque, NM

[(f)]Dept of City and Regional Planning, University of Pennsylvania, Philadelphia, PA

[(g)]Institute for Behavioral Community Health, San Diego State University, San Diego, CA

Abstract

Background—The Systematic Observation of Play and Recreation in Communities (SOPARC) was designed to estimate the number and characteristics of people using neighborhood parks by assessing them 4 times/day, 7 days/week. We tested whether this schedule was adequate and determined the minimum number of observations necessary to provide a robust estimate of park user characteristics and their physical activity levels.

Methods—We conducted observations every hour for 14 hours per day during one summer and one autumn week in 10 urban neighborhood parks: 2 each in Los Angeles, CA, Albuquerque, NM, Columbus, OH, Durham, NC, and Philadelphia, PA. We counted park users by gender, age group, apparent race/ethnicity, and activity level. We used a standardized Cronbach's alpha and intra-class correlation coefficients to test the reliability of using fewer observations.

Results—We observed 76,632 individuals, an average of 547/day (range 155 – 786). Inter-observer reliability ranged from 0.80 to 0.99. Obtaining a robust estimate of park user characteristics and their physical activity required a schedule of 4 days/week, 4 times/day.

Conclusion—An abbreviated schedule of SOPARC was sufficient for estimating park use, park user characteristics, and physical activity. Applying these observation methods can augment physical activity surveillance.

Keywords

Physical activity; parks; measurement; direct observation; surveillance

Introduction

Physical activity is one of the most important health behaviors given its strong and consistent association with longevity and many chronic diseases[1]. Population-level surveillance of physical activity has relied primarily on self-reports garnered from large national surveys such as the Behavioral Risk Factor Surveillance System (BRFSS)[2]. Recently, surveillance efforts have been augmented by the use of more objective methods, including individuals participating in the National Health and Nutrition Examination Survey

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(NHANES) wearing accelerometers[3]. However, the wide discrepancy in the findings from objective and subjective surveillance methods (e.g., 51% of NHANES participants reporting meeting physical activity guidelines compared to <5% by objective measures[3]) gives rise to concerns as to whether population-level physical activity is best assessed by measures at the individual level.

Systematic observation is a direct method of assessing physical activity and has been primarily used in school settings to assess the quality of physical education instruction[4]. It has also served as a tool to assess children's physical activity levels in free play environments and has aided in understanding how the intensity of physical activity at recess is associated with environmental factors[5, 6]. Direct systematic observation is promising as a population measure because it can assess activity in large numbers of people in a short period of time without placing a burden on participants. Whether it can be used as a surveillance tool depends upon the ability of a limited number of sample observations to represent the population and their physical activity levels that would occur in the settings over time.

Population-level measures of physical activity are important not just for surveillance and as a means to assess the impact of community level interventions, but also to understand the association of physical activity with the built and social environments. Historically, many adults engaged in physical activity as part of their jobs, but as work in today's society becomes increasingly sedentary, physical activity, when performed, more often occurs during leisure time. Venues for leisure-time physical activity include streets, health and fitness clubs, parks, and homes. In some surveys, community members have identified parks as the place they most commonly exercise[7] and an estimated 70% of the US population lives within 2 miles of a public park[8], There is increasing interest in how parks can support many modes and levels of physical activity[9, 10].

Since increasing physical activity is an important public health goal,[11] it is critical to have tools that not only help identify the role of environmental factors, but also to measure whether physical activity goals are being achieved. To that end, the System for Observing Play and Recreation in Communities (SOPARC) was designed to capture both park and park user variables in urban and suburban neighborhood parks. It documents throughout the day and across an entire week the level to which park areas are being used, including the number of park users, and reliable judgments about their age group, apparent race/ethnicity, and level of physical activity. Because physical activity behavior is dependent upon contextual factors [12], SOPARC also measures an area's accessibility, usability, and whether or not activities are organized, equipped, and supervised. In the original implementation of SOPARC, park areas were observed 4 times per day during each day of the week for seven days [13].

A limitation to using systematic observation is the cost in both time and money. It is important to have sufficient samples of days and times during the day to be able to obtain valid estimates. For the most precise measure of park use, observations should ideally be continuous, with samples taken every minute throughout every day of the week and during different seasons of the year. Selecting too few observations could potentially be inaccurate and may introduce very large variability, thus misrepresenting usage through the day, between weekdays and weekends, between holidays and non-holidays, and even between weeks and seasons. Selecting too many observations increases significantly the costs of conducting such research. However, given limited resources, a balance between cost and precision must be attained, allowing researchers to sample only some days and some hours during the day. How should observations be optimally scheduled for reliable estimation of park usage and physical activity? Can limited observations robustly estimate the diversity of use among genders, age groups, and race/ethnic populations? In this study, we address these questions by comparing various potential samples of observations to the total aggregate sample of park usage measured during hourly observations over the 7 days of the week in a diverse sample of urban parks.

Methods

SOPARC provides observational user data and contextual information on the setting in which the physical activity occurs. During a scan (i.e., an observation sweep moving from left to right) the physical activity of each individual is coded as sedentary (i.e., lying down, sitting, or standing), walking, or vigorous. These activity codes have been validated using heart rate monitoring and by accelerometry in physical education classes with children and youth in kindergarten through twelfth grade[14, 15]. The activity codes are also consistent with published energy expenditures for adults [16]. People are counted only if they are in a specified target area at the time of the scan and their level of activity at the moment of observation is recorded. Separate scans are made for females and males. The predominant type of activity in each area that females and males are engaging in also is recorded, drawing from a pre-developed list of activities (e.g., basketball, football).

Two neighborhood parks from each of five cities which provided geographic and demographic diversity were selected for observation (Los Angeles, CA, Albuquerque, NM, Durham, NC, Columbus, OH, and Philadelphia, PA). Neighborhood parks are the recreational and social focus of communities and considered by city planners to be recreational and social centers and comprise the basic unit of the park system[17]. Parks were chosen from diverse income areas and observed during 2008. Park characteristics and seasons of observation are described in Table 1. Socio-demographic characteristics of the ½ mile area surrounding the park were obtained from block group data from the US 2000 Census[18]. The 10 parks were mapped into 274 discrete target areas, which included such facilities as basketball courts, tennis courts, and playgrounds. Having target areas facilitated observations and allowed for the documentation of the how different park areas were used.

Observer Training and Reliability Assessment—Coordinators and lead observers from each of the five sites met at a central location for initial training; and they, in turn, trained local data collectors. Observers became familiar with the operational definitions, instrument notation, coding conventions, and discriminating among various physical activity and contextual characteristics. They then practiced coding and received feedback on their scoring using examples contained in a SOPARC training DVD. Afterwards observers reviewed the materials and practiced their skills in diverse field exercises conducted in local parks. Finally, observers demonstrated their accuracy in coding by completing assessments of one round of a park scan in 1 hour or less and by completing an assessment of 5 challenging situations in a park in which in 5 separate target or sub-target areas had at least 10 people to be counted. Certification was conferred if staff had an accuracy rate of at least 90% with respect to the number of people counted and 85% for all other categories, except race/ethnicity and challenging situations (e.g., >5 people engaged in vigorous activity) where 80% reliability was accepted.

Using the SOPARC protocol[13], trained field staff rotated through the park areas in a systematic manner, coding each individual in a target area separately by gender for age group (child, teen, adult, or senior), race/ethnicity (White, African-American, Latino, or Asian and Other), and activity level (sedentary, walking, or vigorous). A rotation through all park target areas usually took less than one hour to complete. When a target area was crowded, sub-target scans were made by dividing the target areas into smaller observation areas.

We observed each target area every hour for 14 hours per day for two weeks during clement weather (i.e. no precipitation), including one week in the summer and one in the fall. Missed observations were made up during the following week in the same time period. An exception was in Los Angeles, where one park was observed for 2 consecutive weeks in the spring and one week in the fall and the second park was observed for only one week in the fall. In addition, during the data collection period we assessed inter-rater reliability for 15% of the 980 observations (N=147) by having two observers collect data simultaneously but independently[13].

Data analyses

Given that 14 hours of daily observation for 7 days was our "gold standard", we used a reliability estimation technique to test whether an abbreviated schedule (e.g., collection on Monday, Wednesday and Saturday with observations collected only twice a day), could produce comparable estimates of park usage. Within each target area in each park, using the abbreviated schedules, we computed the average hourly park usage to see if the abbreviated schedules would approximate the hourly park usage obtained from the gold standard.

Multiple methods have been designed for reliability estimation including test-retest reliability, inter-rater and intra-rater reliability as well as internal consistency [19] and generalizability theory [2021]. When the interest is in estimating the proportion of individual variances in park usage due to individual factors such as days of the week and times of the day, a variance component estimation using intra-class correlations (ICC) can assess the reliability of continuous measures. This ICC is based on one-way analysis of variance. We used the following cut-points to interpret their estimates: ICC values <0.4 signify poor agreement, values between 0.4 and 0.75 signify fair to good agreement, and values >0.75 signify excellent agreement beyond chance [22].

We also used a measure of internal consistency, the standardized Cronbach's alpha. Cronbach's alpha is a measure between 0 and 1, and examines the correlations between the "gold standard" or comparison schedule and the abbreviated schedules. We used the following criteria to evaluate the alpha coefficient: < 0.7 is poor, between 0.7 to <0.8 is acceptable, between 0.8 and <0.9 is good, and between 0.9 and 1 is excellent. Very high reliability (0.95 or higher) is indicative that the estimates may be entirely redundant [23, 24]. We used these thresholds for defining acceptable schedules to estimate park usage and physical activity.

To determine the number of randomly chosen days and the number of randomly chosen times necessary for the reliable estimation of park usage, we estimated the reliability of the hourly park usage for all possible abbreviated schedules: combinations of two, three, or four observation times per a day compared to the 14 times per day. Twice a day collection was defined as once in the morning (6:30 AM to 12:30 PM at 30 minutes after the hour) and once after noon (1:30 PM to 8:30 PM at 30 minutes after the hour); three times a day was defined as early morning (6:30 AM to 11:30 AM), midday (12:30 AM to 3:30 PM) and afternoon-evening (4:30 PM to 8:30 PM); and four times a day was defined as early morning (6:30 AM to 9:30 AM), early midday (10:30 AM to 1:30 PM), afternoon (2:30 PM to 5:30 PM) and evening (6:30 PM to 8:30 PM).

There are multiple ways to schedule observations; for example, there are 56 different combinations of one-hour time frames in the morning and the afternoon (56=7×8, 7 possible times in the morning and 8 possible times in the afternoon) and 35 different ways to choose three days in the week, leading to a total of 1,960 ways of choosing a 3 day, twice a day schedules for park observations. For each one of the 1,960 possible schedules, we computed a reliability estimate. The same procedure was conducted for all the other abbreviated schedules. Together there are over 46,000 possible observation schedules that could be chosen from 1 to 6 days per week and 2 to 4 observations per day (Table 2).

The same methods were used to estimate the reliability of all the characteristics measured—including by gender the number of users, age group, race/ethnicity, and activity levels. These reliability analyses were conducted at the park level as well as at the target area level, which is more accurate for estimating the number of people in different areas of a park hourly.

Results

The 10 parks ranged in size from 4.7 to 39.5 acres and they had an average of 31 target areas (range 19–53). Population density within a one-half mile radius around the park ranged from 8,039 to 28,111 with an average of 17,258 (Table 1). Four parks had more than 15% of households in poverty and six had a majority white population. Facilities and amenities varied greatly: they all had playground areas, but only 9 had outdoor basketball courts, 8 had tennis courts, 7 had multi-purpose fields and baseball fields, and 6 had picnic areas. Five had an indoor gymnasium, and two had swimming pools which were open in the summer only. Six had full time recreational staff, and 4 of the eight parks observed in the summer held day camps for children during the observation period.

Overall (i.e., for 14 hours/day for 7 days per week during two different weeks), we observed a total of 76,632 individuals. Table 1 shows the range of park users, with 1091 people being counted during one fall week in the least busy park in NC and 7,858 being counted during a summer week in the busiest park in OH. On average, 547 people were counted in a park per day (range 155–786), with the differences in individual park use between summer and fall varying from no meaningful change to nearly a two-fold difference. Park use varied from day-to-day and hour-to-hour. On average we observed 39 people in a park every hour (range 12 to 76 from park to park). These numbers comprised an estimated 21 men (range 9 to 41) and 19 women (range 4 to 35).

Inter-rater Reliability

Agreements for park user characteristics between two independent observers were high--averaging 94.2%, with a range from 85% to 99%. When examining only the instances when the target areas were not empty, average agreement on specific park user characteristics was 87% for the total number of individuals, 82% for race/ethnicity, 82% for age group and 80% for activity level, a high level of concordance.

Variation in Use over Hours of the Day and Days of the Week

In general, parks were busiest in the evenings and weekends and used the least in the mornings. When observed on an hourly basis, an average of 76% of the target areas was empty. The percentage of empty target areas was highest in Southern Boundaries, NC (89%) and lowest in Queen Anne, CA (57%). The number and characteristics of persons using parks differed not only between weekdays and weekends, but also between Saturdays and Sundays. The most people were observed on Saturdays (n=12,635), with the fewest on Fridays (n=9,324) and Tuesdays (n=9,399). Playgrounds were the location where the largest number of park users was observed; the fewest were observed in classrooms.

Table 3 reports the average reliability of park use estimates for all possible combinations of days a week and number of times a day. When the interest is in the accuracy of observations in each target area, scheduling data collection on 4 (randomly chosen) days a week and 4 times a day produced excellent consistency (i.e., close to perfect replication of the full day, 7 days observations) of the estimation of the total park usage for all characteristics studied by target area except for the number of people doing vigorous activity. The reliability for estimating the number of people vigorously engaged (alpha=0.86, ICC=0.73) produced a "good" consistency threshold and agreement. The 4 days a week, 4 times a day schedule, produced an average (over all users and not divided by activity type) reliability for the total number of users that was excellent (alpha=0.90 and ICC=0.81). Moreover, with the 6,720 possible ways of choosing 4 observations on each of 4 days per week, the 95% confidence intervals around these average estimates were narrow.

A 3-day schedule observing four times/day might be chosen if the outcome of interest was the count of the number of people using the different target areas (average alpha=0.88 and ICC=0.76) and might still produce to close to excellent reliability even when activity types are of interest. On the other hand, when specific target areas are not of importance, the schedules required to estimate total park use are less demanding (Table 4). Scheduling observations 4 times a day for 2 days, or 3 times a day for 3 days, or twice a day for 5 days allowed for accurate estimation of the sum of all target areas over 14 hours.

Discussion

Efforts to objectively measure park use previously have primarily been limited to large parks in which counts were based upon the number of pedestrians, bicycles, and vehicles traveling through specific entrances and often measured using mechanical counting technology operating 24 hours/day, 7 days a week or by counting entrance fees.[25] While SOPARC provides a snapshot of an hourly count and duplicates counts of park users who stay more than one hour, we found that selecting as few as 12–16 hours a week of measurement appears sufficiently robust for estimating 96 hours of park use and physical activity over a week's time. Moreover, SOPARC is designed to capture many details that have not been noted in studies of larger parks, including personal characteristics of users, the intensity of their activity, the types of activities they engage in, and the specific sites in the parks where they spend time. It also measures routine use of local neighborhood parks which have many entrances, making the use of mechanical counters infeasible.

The high inter-rater reliabilities that resulted from a few days of training and practice for staff recruited locally suggest that this method is one that could readily be used in most settings. The potential uses of SOPARC data are many. If SOPARC were used in parks throughout the country it could ultimately be possible to define the characteristics of successful parks and to develop standards and benchmarks for population physical activity in different settings. There is potential to address health disparities related to physical activity if the health disparities are shown to be associated with park features or park programming that can be changed.

In addition to serving as a surveillance and evaluation tool, the data from SOPARC provide an indication as to when and where marketing and park programming might be introduced to increase park use, or to potentially shift some of the use from the busiest times (or locations) to less used times (or locations). The usage patterns indicate particular facilities and features that either enhance or potentially detract from park use. This information would be useful to park designers and landscape architects to inform renovations and new designs.

Limitations

A limitation of the study is that we assessed most parks during only two different seasons, making it impossible to infer annual usage rates across all seasons of the year. We also assessed parks only during clement weather, rescheduled observations if it rained, and did not made adjustments for diminished park use associated with poor weather conditions. The sample of parks, while diverse, was a convenience sample and included only neighborhood parks of a limited size and none in rural geographic areas. We had just one observation team per park, so limited parks to those that could be rated within a one hour period. Finally, SOPARC is only a snapshot of park use at particular points in a day and week; while an abbreviated scheduled sufficiently represents a more intensive assessment, we still do not know how accurately the intensive assessment represents the actual daily park use.

Conclusion

Cities have built a large infrastructure of parks and play spaces. Considering the high frequency with which we observed park areas in our sample to be vacant, parks across the country appear to be underutilized. Thus, parks are ripe for interventions which SOPARC can inform as well as evaluate. SOPARC methodology can be used for merely 4 times per day, 3 to 4 days per week to provide a robust assessment of urban and suburban neighborhood park use.

Acknowledgments

Funding Source

This study was funded by the National Institutes of Health (NIH), National Heart Lung and Blood Institute # R01HL092569. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

1. USDHHS. Physical activity and health: a report of the Surgeon General. USDHHS; 1996.
2. CDC. [Accessed April 13, 2005] BRFSS: Risk factor for meeting recommended guidelines for vigorous physical activity. http://apps.nccd.cdc.gov/brfss/list.asp?cat=PA&yr=2003&qkey=4419&state=All
3. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. Jan; 2008 40(1):181–188. [PubMed: 18091006]
4. McKenzie TL, Sallis JF, Nader PR. SOFIT: System for observing fitness instruction time. Journal of Teaching in Physical Education. 1991; 11:195–205.
5. Sallis JF, Conway TL, Prochaska JJ, McKenzie TL, Marshall S, Brown M. School environments are associated with youth physical activity. American Journal of Public Health. 2001; (91):618–620. [PubMed: 11291375]
6. McKenzie TL, Marshall SJ, Sallis JF, Conway TL. Leisure-time physical activity in school environments: An observational study using SOPLAY. Preventive Medicine. 2000; (30):70–77. [PubMed: 10642462]
7. Cohen DA, McKenzie TL, Sehgal A, Williamson S, Golinelli D, Lurie N. Contribution of public parks to physical activity. Am J Public Health. Mar; 2007 97(3):509–514. [PubMed: 17267728]
8. Godbey, GC.; Graefe, A.; James, SW. The benefits of local recreation and park services: a nationwide study of the perceptions of the American public. Ashburn VA: National Recreation and Park Association; 1992.
9. Kaczynski AT, Potwarka LR, Saelens BE. Association of park size, distance, and features with physical activity in neighborhood parks. Am J Public Health. Aug; 2008 98(8):1451–1456. [PubMed: 18556600]
10. Cohen DA, Sehgal A, Williamson S, Marsh T, Golinelli D, McKenzie TL. New recreational facilities for the young and the old in Los Angeles: policy and programming implications. Journal Of Public Health Policy. 2009; 30 (Suppl 1):S248–263. [PubMed: 19190577]
11. USDHHS. [Accessed May 22, 3006, 2006] Healthy People. 2010. http://www.healthypeople.gov/
12. Barker, R. Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior. Stanford, CA: Stanford University Press; 1968. Behavior Settings: Defining Attributes and Varying Properties.

13. McKenzie TL, Cohen DA, Sehgal A, Williamson S, Golinelli D. System for Observing Parks and Recreation in Communities (SOPARC): Reliability and feasibility measures. Jl of Physical Activity and Health. 2006; 3 (Suppl 1):S208–S222.
14. Sallis JF, McKenzie TL, Conway TL, et al. Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. American journal of preventive medicine. Apr; 2003 24(3):209–217. [PubMed: 12657338]
15. McKenzie TL, Sallis JF, Nader PR, et al. BEACHES: an observational system for assessing children's eating and physical activity behaviors and associated events. J Appl Behav Anal. Spring;1991 24(1):141–151. [PubMed: 2055797]
16. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. Medicine and science in sports and exercise. Sep; 2000 32(9 Suppl):S498–504. [PubMed: 10993420]
17. Mertes, JD.; Hall, JR. Park, Recreation, Open Space and Greenway Guidelines. Ashburn, VA: National Recreation and Park Association, American Academy for Park and Recreation Administration; 1996.
18. USCensus. Census 2000 Summary File 2 (SF 2) 100-Percent Data. 2004. http://factfinder.census.gov/servlet/DatasetMainPageServlet?_ds_name=DEC_2000_SF3_U&_program=DEC&_lang=en
19. Hinton, PR. Statistics explained. London: Routledge; 2004. Vol 2nd ed
20. Brennan, RL. Generalizability Theory. New York: Springer-Verlag; 2001.
21. Cronbach LJ, Nageswari R, Gleser GC. Theory of generalizability: A liberation of reliability theory. The British Journal of Statistical Psychology. 1963; 16:137–163.
22. Giles-Corti B, Timperio A, Cutt H, et al. Development of a reliable measure of walking within and outside the local neighborhood: RESIDE's Neighborhood Physical Activity Questionnaire. Preventive Medicine. 2006; 42(6):455–459. [PubMed: 16574208]
23. Nunnally, JC.; Bernstein, IH. Psychometric theory. 3. New York: McGraw-Hill; 1994.
24. George, G.; Mallery, P. SPSS for windows step by step: A simple guide and reference, 11.0 update. Boston, MA: Allyn & Bacon; 2003.
25. Mowen, AJ. 2001 Park District Visitation Report: Statistics on Visitors Entering Reservations, Recreational Visits, and Program/Facility Attendance. Cleveland Metroparks; 2001.

Table 1

Park and Population Characteristics

Park, State	Acres	Population within 1/2 mile radius[1]	% white population[1]	% in poverty[1]	# Park Users observed Summer	% park areas empty Summer	# Park Users observed Fall	% park areas empty Fall
Queen Anne, CA * (week 1) Queen Anne, CA (week2)	4.7	25,280	9.5	20.5	6157 5604	57.4 64.7	3271	60.7
Studio City, CA	9.8	18,752	81.6	5.7	n/a	65.4	7469	71.7
Hoffman, NM	12.8	20,573	65.4	9.7	2162	81.1	1436	68.2
Tiquex, NM	12.8	8,039	41.1	21.3	3684	68.7	2590	73.2
Barnett, OH	5.8	11,580	46.0	13.0	7858	76.5	4868	82.6
Carriage Place, OH	13.0	10,500	81.4	6.2	2288	80.9	2128	82.2
Markward, PA	13.0	27,283	57.4	17.5	5756	66.7	6001	75.7
Myers, PA	8.0	28,111	6.6	31.1	5375	74.1	2137	86.7
Southern Boundaries, NC	29.7	8,992	51.0	4.0	1144	88.9	1091	88.7
Pineywood, NC	39.5	13,472	58.0	0.5	3270	80.7	2343	83.7

*observed in Spring

Note

#Park users is the total number counted 14 hours/day, 7 weekdays

[1]From 2000 US Census

Table 2

Number of combinations of park observations averaged for each cell

Number of times per day	Number of days of observation					
	1	**2**	**3**	**4**	**5**	**6**
2	392	1176	1960	1960	1176	392
3	840	2520	4200	4200	2520	840
4	1344	4032	6720	6720	4032	1344

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Reliability estimates using alpha (intraclass correlation coefficient or ICC) averaged over combinations of day and time of day for estimates of park usage within target areas. (

Based upon 76,632 observed park users)

Outcome measured	Number of times per day	Number of days of observation					
		1	**2**	**3**	**4**	**5**	**6**
Number of Users	2	*** (**)	0.75 (0.54)	0.79 (0.61)	0.82 (0.66)	0.84 (0.69)	0.85 (0.71)
	3	0.70 (0.48)	0.80 (0.63)	0.84 (0.71)	0.86 (0.75)	0.88 (**0.78**)	0.89 (**0.79**)
	4	0.75 (0.54)	0.84 (0.69)	0.88 (**0.76**)	**0.90 (0.81)**	**0.92 (0.83)**	**0.93 (0.85)**
Sedentary	2	*** (**)	0.73 (0.50)	0.78 (0.58)	0.82 (0.63)	0.84 (0.67)	0.85 (0.70)
	3	*** (0.44)	0.78 (0.60)	0.83 (0.68)	0.86 (0.73)	0.88 (**0.77**)	0.90 (**0.79**)
	4	0.73 (0.51)	0.83 (0.67)	0.87 (0.75)	**0.90 (0.80)**	**0.92 (0.83)**	**0.93 (0.85)**
Walking	2	*** (**)	*** (0.46)	0.74 (0.55)	0.77 (0.60)	0.79 (0.64)	0.81 (0.67)
	3	*** (**)	0.74 (0.56)	0.79 (0.65)	0.82 (0.71)	0.84 (0.74)	0.86 (**0.77**)
	4	*** (0.47)	0.80 (0.64)	0.85 (0.73)	0.88 (**0.78**)	0.89 (**0.81**)	**0.91 (0.83)**
Vigorous	2	*** (**)	0.72 (0.51)	0.76 (0.57)	0.78 (0.60)	0.80 (0.62)	0.80 (0.63)
	3	*** (0.46)	0.77 (0.59)	0.80 (0.65)	0.82 (0.68)	0.83 (0.69)	0.83 (0.70)
	4	0.72 (0.50)	0.81 (0.64)	0.84 (0.70)	0.86 (0.73)	0.87 (0.75)	0.87 (**0.76**)

*** Cronbach's alpha < 0.7 (poor agreement)

** ICC estimates < 0.4 (poor agreement)

Bold numbers signify excellent reliability for Cronbach alpha, excellent agreement beyond chance for ICC.

Table 4

Average Cronbach's alpha (intraclass correlation coefficient or ICC) for different estimates of the whole park usage.

Average alpha (ICC) for different observation schedules							
	Number of times per day	Number of days					
		1	**2**	**3**	**4**	**5**	**6**
Reliability to Estimate All Users	2	0.81 (0.57)	0.86 (0.66)	0.88 (0.71)	0.90 (0.74)	**0.90 (0.75)**	**0.91 (0.76)**
	3	0.83 (0.64)	0.88 (0.74)	**0.91 (0.79)**	**0.93 (0.82)**	**0.94 (0.84)**	**0.94 (0.85)**
	4	0.86 (0.68)	**0.92 (0.80)**	**0.94 (0.85)**	**0.95 (0.87)**	**0.96 (0.89)**	**0.97 (0.90)**

Bold numbers signify excellent reliability for Cronbach alpha, excellent agreement beyond chance for ICC.