Journal of Urban Health: Bulletin of the New York Academy of Medicine, Vol. 81, No. 4,

doi:10.1093/jurban/jth151

# Reliability of Self-Reported Neighborhood Characteristics

Sandra E. Echeverria, Ana V. Diez-Roux, and Bruce G. Link

**ABSTRACT** *The majority of studies examining the relation between neighborhood environments and health have used census-based indicators to characterize neighborhoods. These studies have shown that neighborhood socioeconomic characteristics are associated with a range of health outcomes. Establishing if these associations reflect causal relations requires testing hypotheses regarding how specific features of neighborhoods are related to specific health outcomes. However, there is little information on the reliability of neighborhood measures. The purpose of this study was to estimate the reliability of a questionnaire measuring various self-reported measures of the neighborhood environment of possible relevance to cardiovascular disease. The study consisted of a face-to-face and telephone interview administered twice to 48 participants over a 2-week period. The face-to-face and telephone portions of the interview lasted an average of 5 and 11 minutes, respectively. The questionnaire was piloted among a largely Latino and African American study sample recruited from a public hospital setting in New York City. Scales were used to assess six neighborhood domains: aesthetic quality, walking/exercise environment, safety from crime, violence, access to healthy foods, and social cohesion. Cronbach's $\alpha$'s ranged from .77 to .94 for the scales corresponding to these domains, with test–retest correlations ranging from 0.78 to 0.91. In addition, neighborhood indices for presence of recreational facilities, quality of recreational facilities, neighborhood participation, and neighborhood problems were examined. Test–retest reliability measures for these indices ranged from 0.73 to 0.91. The results from this study suggested that self-reported neighborhood characteristics can be reliably measured.*

**KEYWORDS** *Cardiovascular disease, Neighborhoods, Pathways, Reliability, Self-report.*

## INTRODUCTION

The majority of studies examining the relation between neighborhood conditions and health have used census-based indicators, typically constructed by aggregating the socioeconomic characteristic of neighborhood residents. Census-based data are useful because they allow researchers to characterize neighborhoods in a systematic fashion across large areas and may be the only measures available in large data sets.

Ms. Echeverria is a doctoral student at Columbia University, Mailman School of Public Health, Department of Epidemiology; Dr. Diez-Roux is Associate Professor of Epidemiology, University of Michigan School of Public Health, Department of Epidemiology; Dr. Link is Professor of Epidemiology and Sociomedical Sciences, Columbia University, Mailman School of Public Health, Departments of Epidemiology and Sociomedical Sciences.

Correspondence: Sandra E. Echeverria, 622 West 168th Street, PH-9, Room 105, New York, NY 10032. (E-mail: see9@columbia.edu)

Dr. Ana Diez-Roux and Sandra Echeverria were in part supported by the Columbia Center for the Health of Urban Minorities (CHUM). Additional support was provided to Sandra Echeverria by the Kellogg Foundation, Health Policy Doctoral Fellowship Program. Dr. Bruce G. Link was supported by a Robert Wood Johnson Health Policy Investigator Award.

However, the use of these aggregate measures also has important limitations.[1–3] A key limitation is that they are indirect proxies for features of neighborhoods that may be relevant to health outcomes. The use of indirect proxies makes it difficult to draw causal inferences regarding neighborhood health effects from the associations observed. What is needed are measures of the health-relevant neighborhood attributes that are necessary to test specific hypotheses regarding the processes through which residential environments may affect health.

Aside from census-based measures, a variety of approaches is available to characterize neighborhood attributes. These include the use of systematic observation of a local area in which trained observers rate neighborhoods on different attributes,[4] the use of geographic information systems to construct measures of the availability and accessibility of different types of resources,[5–7] and the administration of questionnaires to local residents to obtain self-reported measures of neighborhood conditions. Each approach provides different and complementary information. This study focused on the use of self-reported measures.

Although self-reported measures capturing neighborhood attributes have been examined in relation to health outcomes in several studies,[8,9] the reliability and validity of self-reported measures have not been systematically assessed. A first step in evaluating the utility of self-reported measures to investigate neighborhood effects is to examine the test–retest reliability of these measures and the internal consistency of any scales used.

Self-reported measures can be used in two ways. A first approach is to examine the relationship between individual-level perceptions of neighborhoods to individual-level health outcomes. For example, are individual perceptions of neighborhood safety related to individual-level mental health outcomes? This approach may be especially appropriate if hypothesized that it is perceived neighborhood attributes that matter. A limitation, however, is that associations may reflect same-source bias (e.g., a person's mental health may affect the likelihood that they report their neighborhood as unsafe).

A second use of self-reported measures is to aggregate across respondents in a neighborhood to construct a measure for that neighborhood. The underlying assumption is that this aggregation process over individuals' perceptions will result in a more valid measure of the objective attribute. Special methods that take into account inter-individual differences in responses can be used to construct the aggregate measure.[10,11] This second use of self-reported measures is what motivated this analysis, but the reliability results we report are relevant to either use of self-reported measures.

There has been little work on the development of instruments (or scales) that measure theoretically meaningful constructs hypothesized as related to specific health outcomes. One of the health outcomes that has recently been linked to neighborhood socioeconomic conditions is cardiovascular disease.[12–15] Important questions remain, however, regarding whether this association reflects a causal process. For example, does an impoverished neighborhood increase the risk of cardiovascular disease because poor neighborhoods limit the availability of healthy foods to its residents, or is this increased risk related to the general safety of the neighborhood, which can limit the use of any available public spaces for physical activity or socialization? Developing measurement instruments that characterize features of neighborhoods potentially relevant to cardiovascular outcomes would allow the testing of specific hypotheses with empirical data.

In sum, the absence of instruments and the use of measures of questionable reliability represent major limitations in research on neighborhood health effects. In this study, we assessed the reliability of domains intended to measure theoretically meaningful constructs for examining neighborhood effects on cardiovascular disease.

Specifically, we tested the reliability of various self-reported neighborhood characteristics among participants residing in neighborhoods located in a large urban area.

## METHODS

### Study Population

A face-to-face and telephone questionnaire were administered to a largely Latino and African American study sample of 48 persons recruited through fliers in a public hospital setting in New York City. The majority of participants resided in the Washington Heights area of New York City, an area in which the 2000 US Census estimated that 27% of families lived below the federal poverty line, and only 10% of residents 25 years and older had earned a bachelor's degree.[16] The use of the face-to-face and telephone administration for different parts of the questionnaire was based on the desire to mimic the type of administration to be used in a larger study for which the questionnaire was being developed. Because this was a pilot investigation of the reliability of self-reported measures, we did not restrict study participants to a particular neighborhood.

To participate, study subjects had to be 18 years of age or older, be able to conduct interviews in the English language, and be available for interviews in person and over the telephone. The study's purpose and procedures were explained using standard informed consent procedures, and each participant received a copy of the consent form approved by an institutional review board. Respondents were paid for their participation.

### Questionnaire Administration

The face-to-face, or in-person, questionnaire was the first of the two interviews conducted and was immediately followed by the telephone interview (usually 1–2 days after the in-person interview). After the initial round of surveys (time 1), participants were scheduled for the second round of interviewing after a 2-week period (time 2). The same questionnaires were applied at time 1 and time 2. One person did not complete the second round of interviews and was excluded from all analyses.

The telephone questionnaire focused on the assessment of neighborhood-level constructs in preparation for a large-scale residential telephone survey (hence the telephone administration). The in-person questionnaire focused on the measurement of the extent to which persons performed certain types of activities (such as food shopping) within their local area and the amount of time spent in their neighborhoods. This type of information is useful for stratification in analyses relating specific neighborhood attributes to health or health-related behaviors. The in-person questionnaire was piloted in preparation for in-person administration to participants in a large longitudinal study of cardiovascular risk factors. Following procedures employed in other studies,[17] participants were asked to consider their "neighborhood" as the area within a 20-minute walk or about a mile from their home for all of the questions tested in the instrument. Two interviewers were trained and certified by one of the authors before administering questionnaires.

*Telephone Questionnaire* A literature review was conducted using PubMed, and citations from recent review articles were used to identify studies that used self-reported measures of the neighborhood environment of particular relevance to cardiovascular outcomes. Using these studies as a guide, we defined several domains of

interest a priori because of their theoretical relevance to cardiovascular health. After all of the items within each domain had been listed, they were scrutinized to eliminate any that appeared inappropriately worded or explicitly repetitive with others on the list. Problematic items were eliminated or reworded after a preliminary round of reviews by one of the authors (A. D. R.), a reviewer with expertise in questionnaire design, and the interviewers. In addition, we conducted a preliminary round of testing on 20 volunteers (none of whom participated in the present investigation) to determine the timing of the instrument and ordering of items.

After the preliminary round of testing, the final items included in the questionnaire tested the reliability of six scales and four indices, shown in Tables 1 and 2. The six scales were an aesthetic quality scale consisting of 7 items; a walking/exercise environment scale consisting of 11 items; an access to healthy foods scale consisting of 6 items; two safety/crime scales (a safety from crime scale consisting of 3 items and a violence scale consisting of 4 items); and a social cohesion scale consisting of 5 items. We defined these scales as measures of a single theoretical construct pertaining to the neighborhood environment. Thus, internal reliability (Cronbach's $\alpha$) was calculated for each scale. When an item either did not change or decreased the value of $\alpha$ at both time periods, the item was dropped from the scale. Excluding the violence scale, the remaining scales included items with response categories ranging from 1 to 5, for which 1 indicated strongly agree; 2 agree; 3 neutral (neither agree nor disagree); 4 disagree; and 5 strongly disagree. For the violence scale, the response options ranged from 1 to 4 (1=often, 2=sometimes, 3=rarely, and 4=never). Items were reverse coded if necessary so that increasing score reflected increasing disadvantage. For example, in the violence scale, all of the items were reverse coded so that increasing score reflects increasing violence. For each scale, we summed across responses for each item to create a scale score.

In addition, we identified four indices to assess the presence and quality of recreational facilities (recreational facilities indices), participation in neighborhood activities (neighborhood participation index), and potentially stressful neighborhood problems (neighborhood problems index). We defined these indices to summarize a series of neighborhood features. The recreation facilities index included 8 items on the presence of recreational facilities in the neighborhood (yes/no), followed by a rating of these facilities as being in excellent (coded as 1), good (coded as 2), fair (coded as 3), or poor (coded as 4) condition, with a higher score representing poorer quality. Two scores were created for recreational facilities: an availability score and a quality score. The availability score was constructed by assigning 1 point for each facility to which the participant responded "yes" and summing across all facilities. The quality score was constructed by summing the quality responses (scored as noted above) for each of the recreational items for which the participant reported quality.

The neighborhood participation index measured a person's participation (yes/no) in civic and political activities with their neighbors and included 12 items. The neighborhood participation index was created by summing the total number of yes responses, with a maximum score of 12 indicating participation in all 12 activities queried. The neighborhood problems index included 18 items measuring neighborhood characteristics such as presence of trash/litter, unkempt properties, and violence. This index included response categories ranging from 1 to 3, with 1 indicating that the neighborhood attribute was not a problem, 2 it was somewhat of a problem, and 3 it was a big problem. Responses across the 18 problems queried were summed to construct the neighborhood problems index score. Thus, the higher the score on this index was, the more problems the neighborhood was perceived to have.

**TABLE 1. Items included in each scale and sources from which they were drawn***

| Aesthetic environment | Walking/exercise environment | Safety from crime | Access to healthy foods | Social cohesion (Sampson scale) | Violence in past 6 months |
|---|---|---|---|---|---|
| 1. My neighborhood is attractive[18] | 1. My neighborhood offers many opportunities to be physically active[21] | 1. I feel safe walking in my neighborhood during the evening[23] | 1. It is easy to purchase fresh fruits and vegetables in my neighborhood | 1. This is a close-knit or unified neighborhood[11] | 1. During the past six months, how often was there a fight in this neighborhood in which a weapon was used? [11] |
| 2. There is a lot of trash and litter on the street in my neighborhood[8] | 2. Local sports clubs and other providers in my neighborhood offer many opportunities to get exercise[21] | 2. My neighborhood is safe from crime[24] | 2. There is a large selection of fresh fruits and vegetables available in my neighborhood | 2. People around here are willing to help their neighbors[11] | 2. Any gang fights? [11] |
| 3. There are interesting things to do in my neighborhood[5] | 3. It is pleasant to walk in my neighborhood[18] | 3. Violence is a problem in my neighborhood[8] | 3. The fresh produce in my neighborhood is of high quality | 3. People in this neighborhood generally don't get along with each other[11] | 3. A sexual assault or rape? [11] |
| 4. There is enjoyable scenery in my neighborhood[19,20] | 4. There are enough trees in my neighborhood to provide shade[22] | | 4. It is easy to purchase low-fat products (such as low-fat milk or lean meats) in my neighborhood | 4. People in this neighborhood can be trusted[11] | 4. A robbery or mugging? [11] |
| 5. There is a lot of noise in my neighborhood[8] | 5. My neighborhood has heavy traffic[5] | | 5. There is a large selection of low fat products available in my neighborhood | 5. People in this neighborhood do not share the same values[11] | |

**TABLE 1.** ***Continued***

| Aesthetic environment | Walking/exercise environment | Safety from crime | Access to healthy foods | Social cohesion (Sampson scale) | Violence in past 6 months |
|---|---|---|---|---|---|
| 6. In my neighborhood the buildings and homes are well maintained[5] | 6. There are busy roads to cross when out for walks in my neighborhood[5] | | 6. The low-fat products in my neighborhood are of high quality | | |
| 7. The buildings and houses in my neighborhood are interesting[18] | 7. In my neighborhood it is easy to walk to places | | | | |
| | 8. There are stores within walking distance of my home | | | | |
| | 9. In my neighborhood, the streets and sidewalks are in good condition[18] | | | | |
| | 10. I often see other people walking in my neighborhood[5] | | | | |
| | 11. I often see other people exercise (for example, jog, bicycle, play sports) in my neighborhood[19] | | | | |

*Items are not necessarily in the order that they appeared in the questionnaire. The first five scales opened with the following: "For each of the statements that I will read, please tell me whether you agree by choosing the best option. In answering these questions, please think of your neighborhood as the area within about a 20-minute walk from your home." The violence scale began with: "Now I am going to describe some events that may or may not have happened in this neighborhood. For each phrase, please tell me how often it has happened in this neighborhood during the past 6 months. Please base your answer on what you know about your neighborhood generally, and not only on your personal experience."

**TABLE 2. Items included in each index and sources from which indices were drawn***

| Presence of recreational facilities index[19,23] | Activities with neighbors index[25] | Neighborhood problems index[8,26–28] |
|---|---|---|
| 1. Public park | 1. A neighborhood association like a block association, a homeowner or tenant association, or a crime watch group | 1. Trash or litter in the streets |
| 2. Public sports field, basketball court, or tennis court | 2. Religious groups or charitable organizations | 2. Noise from traffic, other homes, airplanes, or businesses |
| 3. Public pool or beach | 3. Parent-teacher associations or other school support or service groups | 3. Smells or fumes |
| 4. Schools, colleges, or community centers with recreational facilities that are free and open to the public | 4. Youth organizations such a youth sports league or the scouts | 4. Lack of safety for walking around after dark |
| 5. Gyms, health/fitness clubs, or pools that you have to join and pay for | 5. Clubs or associations for senior citizens or older people | 5. Lack of places to go for entertainment (restaurants, movie theaters, cafes, bars) |
| 6. YMCAs or YWCAs | 6. A labor union | 6. Poor traffic and road safety |
| 7. Bicycle path in the street or park | 7. A professional, trade, farm, or business association | 7. Lack of places to shop |
| 8. Are there sidewalks in your neighborhood? | 8. Adult sports clubs or leagues or an outdoor activity club | 8. Vandalism, like people breaking windows or spray painting buildings |
| | 9. A literary, art, discussion, or study group or a musical, dancing, or singing group | 9. Vacant housing |
| | 10. Any other hobby club or society | 10. Vacant lots with trash or junk |
| | 11. Ethnic, nationality, or civil rights organizations | 11. Assaults, muggings, or burglaries |
| | 12. Other public interest groups, political groups, or party committees | 12. Lack of trees or green spaces |

**TABLE 2.** ***Continued***

| Presence of recreational facilities index[19,23] | Activities with neighbors index[25] | Neighborhood problems index[8,26–28] |
|---|---|---|
| | | 13. People who don't keep up their property or yards |
| | | 14. No sidewalks or sidewalks in bad condition |
| | | 15. Problems with public services such as street lighting, garbage pickup, and police |
| | | 16. Lack of public transportation |
| | | 17. People fighting or arguing |
| | | 18. People selling illegal drugs |

*For each index, the items are listed in the order that they appeared in the questionnaire. The items in the presence of recreational facilities index had the following introduction: "I would like to ask you about things available in your neighborhood. Please tell me if there are any of the following within a 20-minute walk from your neighborhood, and if so, the condition in which they are in." The activities with neighbors index opened with, "I am going to read you a list of organizations. Please tell me if you regularly join in the activities of these organizations with people in your neighborhood." The neighborhood problems index was introduced as "For each of the following items, please tell me whether it is currently not a problem in your neighborhood, somewhat of a problem, or a big problem."

In addition to the item components of scales and indices, the telephone questionnaire also included three additional single-item questions. The first of these questions was "How safe from crime do you consider your neighborhood to be?" (response options were extremely safe, quite safe, slightly safe, and not at all safe), adapted from the 1996 Behavioral Risk Factor Surveillance System survey.[24] Another question came from the Social Capital Community Benchmark Survey[25] and asked participants how they would rate their neighborhood as a place to live (defined here as a general neighborhood quality measure), with response options ranging from excellent, good, only fair, or poor. The last question was adapted from the Project on Human Development in Chicago Neighborhoods[11] and asked participants to compare their neighborhood to others in their county (possible responses were much better, better, same, worse, much worse).

*In-Person Questionnaire* The in-person questionnaire was a much shorter instrument that included questions on the types of stores and locations where the person's household usually did its food shopping, on the consumption of fast food, on the types of places and locations the respondent used most frequently to get exercise, and on the time the participant spent in his or her neighborhood. The questions were analyzed as single-item questions and were not included in any of the scales or indices. Questions were loosely adapted from the Oakland consumer survey[29] or created by the authors.

### Analysis

The distribution of responses was examined for each item. Means and standard deviations for scales and indices were estimated for each administration time. The internal consistency of the scales was estimated using Cronbach's $\alpha$ coefficient. Test–retest reliability for scales and indices was estimated using intraclass correlation coefficients, variant ICC (1,1), described by Shrout and Fleiss.[30] Agreement in responses to in-person questionnaire and selected items of the phone questionnaire was assessed using $\kappa$ and weighted $\kappa$, with weights estimated based on a form similar to that of Cicchetti and Allison.[31] Spearman correlation coefficients were used to assess convergent validity among scales, indices, and the general neighborhood quality question.[10] The Spearman correlation coefficient was used because one of the variables (general neighborhood quality) was an ordinal variable with just four levels and clearly was not normally distributed.

## RESULTS

The mean age of the 48 respondents was 38 years; 75% of the sample was female, and participants had lived an average of 13 years in their neighborhood. Overall, the in-person interview took less than 5 minutes to conduct, and the telephone interview lasted approximately 11 minutes (Table 3). Because the interviews could be administered in less than 15 minutes, approximately one third of the sample opted to be interviewed during their working hours (data not shown).

### Phone Questionnaire

As discussed in the Methods section, we first conducted a pretest of the questionnaire items we developed or identified through the literature review. One item ("my neighborhood is friendly") originally included in the aesthetic quality scale was dropped because deleting this item did not significantly change Cronbach's $\alpha$ for the scale. In the walking/exercise environment scale, deleting the item "There are

**TABLE 3. Demographic characteristics of study sample and mean time of interviews (N = 48)**

| | Mean | SD |
|---|---|---|
| Age | 38.4 | 12.2 |
| Gender: female (%) | 75% | — |
| Number of years living in neighborhood | 12.6 | 12.0 |
| Length of phone interview (minutes): time 1 | 12.3 | 4.0 |
| Length of phone interview (minutes): time 2 | 10.5 | 2.2 |
| Length of in-person interview (minutes): time 1 | 5.3 | 2.6 |
| Length of in-person interview (minutes): time 2 | 4.2 | 2.1 |

stores within walking distance of my home" did not substantially modify the α, but the item was retained. The rationale for keeping this item was that there was likely little variability in this item because of the New York City sample, and in other areas, the item might contribute to reliability. An item on access to fast foods and two affordability items were dropped from the access to healthy foods scale because deleting them increased the α of the scale. The item "I feel safe walking in my neighborhood during the day" was also dropped from the safety from crime scale. The item "This is a close-knit neighborhood" in the social cohesion scale was identified as problematic after the preliminary round of testing indicated that the wording was unclear to many participants. The question was modified to read, "This is a close-knit or unified neighborhood."

Means and standard deviations for the final scale scores are shown in Table 4. All of the scales tested achieved Cronbach's α of .77 or greater, ranging from .77 at time 1 for the safety from crime scale to .94 at time 2 for the access to healthy foods scale, indicating very good internal consistency of the measures (Table 4). Moreover, test–retest reliability was equally strong across all of the scales, with correlation coefficients greater than or equal to 0.78 for all scales and 0.88 or greater for four of the six scales.

Results for indices are shown in Table 5. The mean number of facilities in the neighborhood was 6 at time 1 and 6.5 at time 2, with a mean quality score of 13.0 at time 1 and 14.3 at time 2, indicating an average per facility score of slightly over 2. Results for the neighborhood participation index suggest that the study subjects rarely participated in group activities with their neighbors (mean = 1.9 and SD = 2.2 at time 1; mean = 1.5 and SD = 2.2 at time 2). Only 2 of the 48 subjects interviewed participated in more than 3 community activities (data not shown). The mean value of the neighborhood problem index was 31.8 with a standard deviation of 9.1 at time 1, with similar results at time 2. The items most frequently reported to be a big problem in the neighborhood were noise from traffic, other homes, airplanes, or businesses (time 1 = 50%, time 2 = 46%) and people selling illegal drugs (time 1 = 38%, time 2 = 43%). Test–retest reliability was highest for the neighborhood problems index (ICC = 0.91), followed by the recreational facilities indices (ICC = 0.85), and lowest for the neighborhood participation index (ICC = 0.73).

The responses to the three individual questions included in the telephone questionnaire are shown in Table 6. Although approximately 60% of the respondents felt that their neighborhoods were either only slightly safe or not safe at all from crime, 60% or more of the participants rated their neighborhoods as an excellent or good place to live (general neighborhood quality measure). In addition, fewer than 25% of the participants reported that their neighborhood was worse or much worse

**TABLE 4. Mean levels at both administration times and estimates of internal consistency and test–retest reliability for neighborhood scales**

| Construct | Number of items per scale | Score range* | Mean/SD (n), time 1 | Mean/SD (n), time 2 | Cronbach's α, time 1 | Cronbach's α, time 2 | Test–retest ICC (95% confidence interval) |
|---|---|---|---|---|---|---|---|
| Aesthetic quality | 7 | 7–35 | 20.7/6.7 (47) | 19.8/6.2 (48) | .89 | .88 | 0.91 (0.85–0.95) |
| Walking/ exercise environment | 11 | 11–55 | 29.2/6.6 (46) | 28.5/6.2 (48) | .78 | .78 | 0.88 (0.79–0.93) |
| Access to healthy foods† | 6 | 6–30 | 15.1/5.3 (47) | 15.4/5.6 (48) | .91 | .94 | 0.88 (0.79–0.83) |
| Safety from crime | 3 | 3–15 | 9.2/2.7 (48) | 9.1/2.9 (48) | .77 | .82 | 0.80 (0.67–0.88) |
| Violence | 4 | 4–16 | 8.8/2.9 (46) | 8.5/3.0 (46) | .85 | .83 | 0.78 (0.64–0.87) |
| Social cohesion | 5 | 5–25 | 14.2/3.6 (48) | 14.4/4.1 (48) | .82 | .86 | 0.90 (0.84–0.94) |

ICC, intraclass correlation coefficient.

*Increasing score (mean) represents increasing disadvantage.

†For the fast food item alone (which was dropped from the scale), mean and standard deviations were 4.2 (SD = 1.0) at time 1 and 4.4 (SD = 0.8) at time 2. The κ was 0.58 (0.39–0.77), and weighted κ was 0.58 (0.39–0.78).

**TABLE 5. Mean levels at both administration times and reliability estimates for recreational facilities indices, neighborhood participation index, and neighborhood problems index**

| Index | Number of items per scale | Score range* | Mean/SD (n), time 1 | Mean/SD (n), time 2 | Test–retest ICC (95% confidence interval) |
|---|---|---|---|---|---|
| Presence of recreational facilities index | 8 | 0–8 | 6.0/1.9 (38) | 6.5/1.8 (42) | 0.85 (0.75–0.92) |
| Quality of recreational facilities index | † | 0–4 per item | 13.0/5.4 (37) | 14.3/5.4 (42) | 0.81 (0.68–0.89) |
| Neighborhood participation index | 12 | 1–12 | 1.9/2.2 (47) | 1.5/2.2 (48) | 0.73 (0.56–0.84) |
| Neighborhood problems index | 18 | 18–54 | 31.8/9.1 (45) | 30.4/8.2 (47) | 0.91 (0.84–0.95) |

ICC, intraclass correlation coefficient.

*Increasing score (mean) represents increasing disadvantage.

†Total quality score includes items for which quality was reported and was restricted to persons with complete data on presence of facilities.

**TABLE 6. Distribution of responses (%) and agreement statistics for general neighborhood quality questions**

| | Time 1 (n = 48) | Time 2 (n = 48) |
|---|---|---|
| How safe from crime do you consider your neighborhood to be? | | |
| Extremely safe | 6 | 8 |
| Quite safe | 31 | 31 |
| Slightly safe | 52 | 48 |
| Not at all safe | 10 | 13 |
| κ/weighted κ | 0.64 (0.44–0.83)/0.72 (0.55–0.88) | |
| Overall how would you rate your neighborhood as a place to live?* | | |
| Excellent | 13 | 15 |
| Good | 52 | 46 |
| Only fair | 27 | 31 |
| Poor | 8 | 8 |
| κ/weighted κ | 0.71 (0.54–0.88)/0.78 (0.64–0.93) | |
| And how do you think your neighborhood compares to other neighborhoods in the city? | | |
| Much better | 19 | 23 |
| Better | 19 | 10 |
| Same | 44 | 42 |
| Worse | 17 | 21 |
| Much worse | 2 | 4 |
| κ/weighted κ | 0.65 (0.49–0.82)/0.76 (0.63–0.88) | |

*General neighborhood quality.

when compared to other neighborhoods in the city where they lived. The reliability of these three questions was relatively high (weighted κ's ranging from 0.72 to 0.78).

Table 7 shows Spearman correlations among scales, indices, and responses to the individual question on general neighborhood quality. All scales were correlated with the general neighborhood quality question in the expected direction: Persons who reported poorer neighborhood quality also reported a less-pleasant aesthetic environment, an environment less conducive to walking/exercise, less access to healthy foods, less safety from crime, more violence, and less social cohesion. Persons who reported poorer neighborhood quality also reported more problems and fewer recreational facilities. The other two indices (quality of recreational facilities and neighborhood participation index) were not correlated with general neighborhood quality. High correlations (0.7 or more) were observed for aesthetic quality and walking/exercise environment, presence and quality of recreational facilities, aesthetic quality and neighborhood problems, aesthetic quality and general neighborhood quality, and neighborhood problems and general neighborhood quality.

### In-Person Questionnaire

Results from the in-person questionnaire showed that approximately 80% of persons reported that the place where they did most of their food shopping was about 1 mile or less from their home (time 1=79%, time 2=81%), and 71% reported that at least three quarters of their food shopping was done in their neighborhood (Table 8). Supermarkets were the most common source for obtaining groceries among the study participants. Among persons who reported exercising regularly, the most commonly used place for exercise was the street or sidewalks (time 1=35%, time 2=33%). The majority of study participants reported that the place they used most often to get exercise was about 1 mile or less away from their home (time 1=71%, time 2=75%).

In general, the items on the amount of time spent in the neighborhood yielded more inconsistent distribution of responses across the study periods. In this sample, 44% reported at time 1 that they spent all or most of Saturday and Sundays in their neighborhoods, in comparison to 48% at time 2. There were 46% who reported spending all or most of their time in the neighborhood Monday to Friday from 7 AM to 5 PM at time 1, and 42% reported this at time 2. For the amount of time spent in the neighborhood from 5 to 9 PM, 48% of the sample reported spending all or most of their time in the neighborhood at time 1, and 67% did so at time 2.

Weighted κ's indicated that reliability was generally acceptable (weighted $\kappa > 0.65$) for questions related to place where household shopped, amount of shopping done in neighborhood, types of stores used most often, frequency of eating fast food in general, place used most often to get exercise and its distance from home, and time spent in the neighborhood during the day Monday–Friday. However, reliability was lower for items related to frequency of eating fast food in the neighborhood (weighted $\kappa = 0.42$), frequency of exercising in the neighborhood (weighted $\kappa = 0.43$), and time spent in the neighborhood Monday–Friday evenings, nights, and Saturdays and Sundays (weighted κ ranged from .32 to .53).

## DISCUSSION

This investigation was one of the first studies to evaluate systematically the reliability of self-reported measures of distinct domains of the neighborhood environment theoretically related to cardiovascular disease. The few studies that have explicitly examined the reliability of self-reported measures of neighborhood attributes have

**TABLE 7. Spearman correlations among scales, indices, and responses to general neighborhood quality question**

| | Walking/exercise environment | Access healthy foods | Safety from crime | Violence | Social cohesion | Presence of recreational facilities | Quality of recreational facilities index | Neighborhood Problems Index | Neighborhood participation index | General neighborhood quality |
|---|---|---|---|---|---|---|---|---|---|---|
| Aesthetic quality | 0.82* | 0.59* | 0.54* | 0.36† | 0.52* | –0.34† | 0.11 | 0.77* | –0.14 | 0.70* |
| Walking/exercise environment | | 0.57* | 0.56* | 0.32† | 0.47† | –0.41† | –0.02 | 0.66* | –0.05 | 0.59*† |
| Access to healthy food | | | 0.33† | 0.43† | 0.46† | –0.11 | 0.19 | 0.59* | –0.09 | 0.50† |
| Safety from crime | | | | 0.51† | 0.61* | –0.21 | 0.09 | 0.60* | 0.05 | 0.62* |
| Violence | | | | | 0.44† | 0.04 | 0.21 | 0.57* | 0.02 | 0.40† |
| Social cohesion | | | | | | –0.15 | 0.06 | 0.52* | –0.29† | 0.59* |
| Presence of recreational facilities | | | | | | | 0.78* | –0.15 | 0.11 | –0.38 |
| Quality of recreational facilities index | | | | | | | | 0.23 | 0.10 | 0.01 |
| Neighborhood participation index | | | | | | | | | –0.05 | –0.15 |
| Neighborhood problems index | | | | | | | | | | 0.70* |

Correlations refer to first administration.
*$P \leq .0001$.
†$P < .05$, but $> .0001$.

**TABLE 8. Distribution of responses (%) and reliability estimates for in-person interview questions**

| Question | Time 1 | Time 2 |
|---|---|---|
| Place where household shops | N = 48 | N = 48 |
| 1 mile or less | 79 | 81 |
| 2–5 miles | 17 | 13 |
| 6–10 miles | 2 | 5 |
| More than 10 miles | 2 | 2 |
| κ/weighted κ | 0.75 (0.53–0.97)/0.89 (0.76–1.0); N = 48 | |
| Amount of shopping done in neighborhood | N = 48 | N = 48 |
| All or almost all | 50 | 52 |
| About three quarters | 21 | 17 |
| About a half | 17 | 17 |
| About one quarter | 4 | 8 |
| None or almost none | 8 | 6 |
| κ/weighted κ | 0.47 (0.28–0.66)/0.67 (0.43–0.91); N = 48 | |
| Type of food stores most commonly used | N = 48 | N = 48 |
| Supermarkets/superstores | 83 | 81 |
| Grocery stores/bodegas/delis | 10 | 13 |
| Convenience stores | 0 | 0 |
| Specialty stores | 6 | 6 |
| κ/weighted κ | 0.80 (0.57–1.0)/0.95 (0.87–1.0); N = 48 | |
| Frequency of eating at a fast food place | N = 48 | N = 48 |
| Almost never or never | 33 | 33 |
| Less than once a week | 21 | 13 |
| 1–2 times a week | 27 | 38 |
| 3–4 times a week | 10 | 15 |
| Five or more times a week | 8 | 2 |
| κ/weighted κ (n) | 0.44 (0.27–0.62)/0.72 (0.55–0.89); N = 48 | |
| How often fast food is eaten in neighborhood | N = 43 | N = 42 |
| All or almost all the time | 37 | 52 |
| About three quarters of the time | 21 | 24 |
| About a half of the time | 16 | 14 |
| About one quarter of the time | 9 | 2 |
| None or almost none of the time | 16 | 7 |
| κ/weighted κ | 0.40 (0.20–0.60)/0.42 (0.13–0.72); N = 42 | |
| Place used most commonly to exercise | N = 42 | N = 41 |
| Public parks or other public facilities | 19 | 18 |
| Streets or sidewalks | 35 | 33 |
| Schools | 5 | 0 |
| Religious institutions | 0 | 0 |
| Private gyms, clubs, dance studios/ YMCA/YWCA** | 29 | 38 |
| Own home | 12 | 13 |
| κ/weighted κ | 0.64 (0.46–0.83)/0.76 (0.59–0.92); N = 39 | |

**TABLE 8.** ***Continued***

| Question | Time 1 | Time 2 |
|---|---|---|
| Distance traveled from home to exercise | N = 41 | N = 41 |
| 1 mile or less | 71 | 75 |
| 2–5 miles | 17 | 18 |
| 6–10 miles | 7 | 3 |
| More than 10 miles | 5 | 5 |
| κ/weighted κ (n) | 0.58 (0.30–0.86)/0.79 (0.56–1.0); N = 48 | |
| Frequency of exercising in neighborhood | N = 42 | N = 41 |
| All or almost all the time | 57 | 63 |
| About three quarters of the time | 10 | 10 |
| About a half of the time | 17 | 13 |
| About one quarter of the time | 7 | 5 |
| None or almost none of the time | 10 | 10 |
| κ/weighted κ | 0.20 (0.0–0.44)/0.43 (0.08–0.77); N = 39 | |
| Time spent in neighborhood, per week Saturday and Sunday* | N = 48 | N = 48 |
| All or most of the time | 44 | 48 |
| About half of the time | 42 | 38 |
| Less than half of the time | 15 | 15 |
| κ/weighted κ | 0.32 (0.11–0.53)/0.34 (0.07–0.62); N = 48 | |
| Monday–Fri days* | N = 48 | N = 47 |
| All or most of the time | 46 | 42 |
| About half of the time | 10 | 17 |
| Less than half of the time | 44 | 40 |
| κ/weighted κ | 0.75 (0.60–0.92)/0.84 (0.70–0.98); N = 47 | |
| Mon–Friday evenings* | N = 48 | N = 48 |
| All or most of the time | 48 | 67 |
| About half of the time | 33 | 25 |
| Less than half of the time | 19 | 8 |
| κ/weighted κ | 0.36 (0.16–0.55)/0.42 (0.20–0.64); N = 48 | |
| Monday–Friday nights* | N = 48 | N = 47 |
| All or most of the time | 90 | 85 |
| About half of the time | 10 | 13 |
| Less than half of the time | 0 | 2 |
| κ/weighted κ | 0.53 (0.20–0.86)/0.57 (0.27–0.86); N = 47 | |

*Responses for four categories (all or most; about half; about a quarter; none or almost none) were available for 41 participants: κ/weighted κ 0.38 (0.14–0.60)/0.36 (0.0–0.72) for Saturday/Sunday; κ 0.63 (0.45–0.81)/0.83 (0.68–0.98) for Monday–Friday days, 0.37 (0.16–0.57)/0.36(0.16–0.57) for Monday–Friday evenings, and 0.59(0.25–0.92)/0.62 (0.33–0.91) for Monday–Friday nights.

**For 32 participants, responses were available splitting private gyms, clubs, and dance studios from YMCAs/YWCAs: κ 0.62 (0.42–0.83), weighted κ 0.73 (0.54–0.91).

largely focused on features relevant to physical activity levels[18,32,33] or have limited the assessment of these features to measures related to esthetic quality and general safety concerns.[11,19] Our study builds on this body of work and examined the reliability of a broader set of neighborhood attributes, ranging from features conducive to walking/exercise and social cohesion to access to healthy foods. Results indicated that the majority of the scales examined had high internal consistency, and that estimates

of internal consistency were similar at both times of administration. For four of the six scales, test–retest was above 0.88. The safety from crime and violence scale had lower but acceptable test–retest results of 0.80 and 0.78, respectively.

In general, Spearman correlation results showed evidence of convergent validity between the scales and the general neighborhood quality measure (range of 0.38 to 0.70). In addition, the neighborhood problems index was strongly correlated with the safety from crime scale (0.60) and the violence scale (0.57). These results indicated that the correlations were in the expected direction and suggested that the measures were behaving in a predictable fashion. Although encouraging, correlations in the expected direction should be regarded as preliminary evidence of the validity of self-reported measures of neighborhood attributes. More elaborate assessments of construct validity await future research.[10]

As in other measurement domains, the assessment of neighborhood attributes is likely to be best captured through scales or indices comprised of multiple items. In this study, several of the domains assessed in the in-person questionnaire included only single items, possibly explaining the lower reliability of some of these measures. Alternatively, although the neighborhood demarcation used in this investigation has been used in previous studies,[17] it is nonetheless an arbitrary definition that may not be suitable across people or places or across the range of activities measured in the present study. This may have made it difficult for persons to respond to these questions, resulting in lower test–retest reliability. In a recent article by Macintyre et al.,[3] the authors suggested that different activities may require distinct spatial scales. An important methodological challenge in the study of the effect of places on health is to define what constitutes an appropriate spatial scale and range of resources for specific human needs. In addition, several of the questions with low reliability, such as those related to frequency of eating fast foods or exercise, could have been influenced by perceptions of social desirability, possibly also affecting the reliability of the question.

The relatively poor (but still acceptable) reliability of the questions on the amount of time spent in the neighborhood may be related to the time intervals used in the question and to particular characteristics of the study population. Although we did not record the occupational status or working hours of the participants, it may be that this population had long or inconsistent work schedules, making it difficult for respondents to attribute how much time they spent in their neighborhood in the discrete time intervals surveyed. Moreover, because this study was conducted in a large metropolitan city, it may be that over a 2-week interval participants were engaged in varying levels of activity that would make the measures less reliable.

The particularities of a city-dwelling population may also explain why there was a restricted range of responses for some items. Further, because of limitations caused by sample size and population characteristics (75% of our sample was female, and all were urban residents), we were not able to conduct separate analyses based on gender, sociodemographic characteristics, or geographic area. Other studies with larger study populations have faced similar challenges (e.g., in a study by Brownson et al., 61–71%[33] of the study participants were female), but have been able to examine differences caused by rural/urban residence and age. Results from these studies generally suggested that certain questions appear to be more selectively reliable at younger ages and for urban or rural residents,[33,34] although most questions showed moderate to high reliability across groups. In line with these studies, we believe that the reliability of our findings is likely to apply across groups. However, this needs to be tested in larger and more varied samples.

The paradoxical finding that a large number of residents reported not feeling safe in their neighborhoods while simultaneously reporting that their neighborhoods were excellent or good places to live suggests a complex relation between people's perceptions of the physical and social environment and the value and meaning attached to a particular neighborhood. Suttles[35] proposed 30 years ago that residents of dangerous neighborhoods "know" the source(s) of danger in their communities and purposely avoid these encounters, thereby allowing them to live in relative safety. In reviewing more recent work done on neighborhoods, Kawachi and Berkman[36] also cited evidence describing this phenomenon as one in which individuals living in contexts of despair protect themselves by evaluating their lived experience as adequate and minimizing the deprivation they face. In addition, it may be that the factors driving perceptions of neighborhood safety may be very different from those driving perceptions of general neighborhood quality. Further studies are needed (possibly qualitative in nature) to elucidate better the complex relationship between specific neighborhood conditions, such as lack of safety or other neighborhood features, and perceived neighborhood quality or attachment to place. This may be particularly important to examine in different race/ethnic groups and in communities with varying levels of cohesion and attachment to place based on cultural, historical, or social reasons.

Despite the potential limitations discussed above, the measurement approach we have developed has several advantages. First, our study suggests that a broader set of neighborhood conditions can be efficiently measured using participants' self-report. The questionnaires required minimal training of the research staff, and on average, the longer of the two instruments took no more than 12 minutes to administer. Second, although our study was designed as a pilot investigation and was restricted to a small number of study participants, the results generally provide strong evidence for the reliability of self-reported measures of neighborhood attributes.

A next step in the evaluation of the utility of self-reported neighborhood characteristics is to examine the correlations in responses between participants living in the same neighborhood using techniques such as ecometrics.[10] Three-level multilevel models can be used to model variability between scale items, between persons within a neighborhood, and between neighborhoods simultaneously. This approach allows estimation of item consistency within each scale, interrater agreement (or intraclass correlation coefficients) for raters within neighborhood clusters for each scale, and an overall measurement of the reliability of the neighborhood-level measure. These models can be used to derive estimates of neighborhood-level latent variables adjusted for individual-level characteristics of respondents and for measurement error, which can then be used as predictors in analyses relating group characteristics to individual-level outcomes.

Finally, the evaluation of within-neighborhood agreement between respondents was beyond the scope of this pilot study and will require much larger data sets with multiple respondents per neighborhood. However, the demonstration that self-reported measures of neighborhood attributes are reliable at the individual level is a first step leading to the empirical examination of the specific mechanisms by which neighborhood attributes can lead to poor cardiovascular health.

## REFERENCES

1. Oakes M. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med*. 2004;58:1929–1952.

2. Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health*. 2001;91:1783–1789.
3. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure them? *Soc Sci Med*. 2002;55:125–139.
4. Sampson R, Raudenbush S. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol*. 1999;105:603–651.
5. Giles-Corti B, Donovan RJ. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. *Prev Med*. 2002;35:601–611.
6. 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.
7. Pikora TJ, Bull FC, Jamrozik K, Knuiman M, Giles-Corti B, Donovan RJ. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med*. 2002;23:187–194.
8. Balfour JL, Kaplan GA. Neighborhood environment and loss of physical function in older adults: evidence from the Alameda County Study. *Am J Epidemiol*. 2002;155:507–515.
9. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: a review. *Am J Prev Med*. 2002;22:188–199.
10. Raudenbush SW. Quantitative assessment of neighborhood social environments. In: Kawachi I, Berkman LF, eds. *Neighborhoods and Health*. Oxford, England: Oxford University Press; 2003: chapter 5:112–131.
11. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997;277:918–924.
12. Davey Smith G, Hart C, Watt G, Hole D, Hawthorne V. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. *J Epidemiol Community Health*. 1998;52:399–405.
13. Hart C, Ecob R, Smith GD. People, places and coronary heart disease risk factors: a multilevel analysis of the Scottish Heart Health Study archive. *Soc Sci Med*. 1997;45:893–902.
14. Diez Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med*. 2001;345:99–106.
15. Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol*. 1997;146:48–63.
16. Bureau US Census. Census 2000 Summary File 3 (SF3), Sample Data, Detailed Tables. Vol. 9 June 2004. Accessed June 11, 2004. Available at: http://www.census.gov/servlet.
17. *Local Area Survey*. Social and Public Health Sciences Unit, University of Glasgow, Scotland.
18. Ball K, Bauman A, Leslie E, Own N. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. *Prev Med*. 2001;33:434–440.
19. Sallis JF, Johnson MF, Calfas KJ, Caparosa S, Nichols JF. Assessing perceived physical environmental variables that may influence physical activity. *Res Q Exerc Sport*. 1997;68:345–351.
20. 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*. 2000;19:354–364.
21. Stahl T, Rutten A, Nutbeam D, et al. The importance of the social environment for physically active lifestyle—results from an international study. *Soc Sci Med*. 2001;52:1–10.
22. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity: views from urban planning. *Am J Prev Med*. 2002;23(2 suppl):64–73.
23. 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.
24. Neighborhood safety and the prevalence of physical inactivity–selected states, 1996. *JAMA*. 1999;281:1373.

25. Research TRCfPO. *The Social Capital Benchmark Survey (Data Codebook)*. Storrs, CT: The Roper Center for Public Opinion Research; 2002.
26. Sooman A, Macintyre S. Health and perceptions of the local environment in socially contrasting neighborhoods in Glasgow. *Health Place*. 1995;1:15–26.
27. Steptoe A, Feldman PJ. Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems, and associations with socioeconomic status and health. *Ann Behav Med*. 2001;23:177–185.
28. Perkins DD, Taylor RB. Ecological assessments of community disorder: their relationship to fear of crime and theoretical implications. *Am J Community Psychol*. 1996;24:63–107.
29. Troutt D. The thin red line: how the poor still pay more. San Francisco, CA: Consumers Union, West Coast Regional Office; 1993.
30. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;2:420–428.
31. Cicchetti DV, Allison T. A new procedure for assessing the reliability of scoring EEG sleep recordings. *Am J EEG Technol*. 1971;11:101–109.
32. Saelens BE, Sallis JF, Black JB, Chen D. Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health*. 2003;93:1552–1558.
33. Brownson RC, Chang JJ, Eyler AA, et al. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of three questionnaires. *Am J Public Health*. 2004;94:473–483.
34. 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 Public Health*. 2004;94:440–443.
35. Suttles GD. *The Social Construction of Communities*. Chicago, IL: University of Chicago Press; 1972.
36. Kawachi I, Berkman LF. Introduction. In: Kawachi I, Berkman LF, eds. *Neighborhoods and Health*. Oxford, England: Oxford Publishing Press; 2003:1–19.