Neighborhood Environment Walkability Scale: Validity and Development of a Short Form

ESTER CERIN[1], BRIAN E. SAELENS[2], JAMES F. SALLIS[3], and LAWRENCE D. FRANK[4]

[1]Institute of Human Performance, The University of Hong Kong, Hong Kong, CHINA; [2]Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, OH; [3]Department of Psychology, San Diego State University, San Diego, CA; and [4]School of Community and Regional Planning, The University of British Columbia, Vancouver, CANADA

ABSTRACT

CERIN, E., B. E. SAELENS, J. F. SALLIS, and L. D. FRANK. Neighborhood Environment Walkability Scale: Validity and Development of a Short Form. *Med. Sci. Sports Exerc.*, Vol. 38, No. 9, pp. 1682–1691, 2006. **Purpose:** The aim of this study was to examine the factorial and criterion validity of the Neighborhood Environment Walkability Scale (NEWS) and to develop an abbreviated version (NEWS-A). **Methods:** A stratified two-stage cluster sample design was used to recruit 1286 adults. The sample was drawn from residential addresses within eight high- and eight low-walkable neighborhoods matched for socioeconomic status. Subjects completed the NEWS and reported weekly minutes of walking for transport and recreation using items from the International Physical Activity Questionnaire. **Results:** Multilevel confirmatory factor analysis was used to develop measurement models of the NEWS and NEWS-A. Six individual-level and five blockgroup-level factors were identified. Factors/scales gauging presence of diversity of destinations, residential density, walking infrastructure, aesthetics, traffic safety, and crime were positively related to walking for transport. Aesthetics, mixed destinations, and residential density were associated with walking for recreation. **Conclusions:** The NEWS and NEWS-A possess adequate levels of factorial and criterion validity. Alternative methods of scoring for different purposes are presented. **Key Words:** MULTILEVEL CONFIRMATORY FACTOR ANALYSIS, WALKING FOR TRANSPORT, WALKING FOR RECREATION, PHYSICAL ACTIVITY, BUILT ENVIRONMENT

To identify correlates of physical activity that can guide improved interventions, researchers have been shifting from individually based theories to multilevel ecological approaches (21,26,27). Ecological models posit that built and natural environmental factors play important roles in shaping physical activity.

Empirical support for the significant impact of environmental attributes on physical activity has been accumulating in multiple disciplines. Reviews of the transportation and urban planning literature found that residents from neighborhoods with higher levels of residential density, street connectivity, and land use mix reported more walking and cycling than their counterparts (14,25,30). A review of the health and behavioral science literature found relatively consistent positive associations between physical activity and accessibility to recreational facilities and aesthetic attributes, but not safety from traffic and crime (17).

Despite these encouraging findings, there are still many questions to be answered about environment–physical activity relationships. An accurate analysis of these relationships requires the use of valid measures of attributes of the built environment as well as of physical activity. Attributes of the built environment can be measured objectively (e.g., using geographic information systems data) and subjectively (e.g., using questionnaires).

The Neighborhood Environment Walkability Scale (NEWS) (24) is one of several recently developed questionnaires designed to measure residents' perceptions of the environmental attributes of their local area (5). The NEWS was designed to obtain residents' perceptions of how neighborhood characteristics found in the transportation and urban planning literature were related to a higher frequency of walking and cycling trips (25). Additional NEWS items were created based on input from local planning and transportation experts. Collectively, these neighborhood characteristics were theorized to operationalize the larger construct of neighborhood walkability, which we hypothesized would be related to the level of walking among residents. Items were *a priori* grouped into subscales to assess the underlying constructs of residential density, proximity to stores and facilities, perceived access to these destinations, street connectivity, facilities for walking and cycling, aesthetics, and safety from traffic and crime. Initial evidence for validity was based on mean differences found between NEWS subscale scores of residents living in neighborhoods known to differ on neighborhood walkability characteristics (e.g., the neighborhood differed in objectively measured residential density, and residents perceived this difference) (20,24). However, prior studies using the NEWS did not assess the factorial validity of the NEWS subscales, that is, whether items on subscales formed coherent factors and

Address for correspondence: Ester Cerin, Institute of Human Performance, The University of Hong Kong, 111-113 Pokfulam Road, Pokfulam, Hong Kong, China; E-mail: ecerin@hku.hk.
Submitted for publication September 2005.
Accepted for publication April 2006.

0915-9131/06/3809-1682/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/01.mss.0000227639.83607.4d

whether individual factors derived from subscales were related to walking (5,20,24).

Three studies have supported test–retest reliability of the NEWS (5,20,24), and other studies have provided partial support for its construct validity by reporting significant differences on some NEWS subscales between neighborhoods selected to differ on walkability (20,24) and modest correlations between NEWS subscales and accelerometer (1) and self-reported (9) estimates of physical activity. Several of these studies were limited in terms of neighborhood variability (e.g., only two neighborhoods were examined) and participant sample sizes. The *a priori* scales used may not be optimal, which could partly explain why only modest correlations were observed between the original *a priori* NEWS subscales and physical activity outcomes. Hence, the aim of this study was to provide a more comprehensive evaluation of the construct validity (factorial and criterion) of the NEWS. This was achieved by analyzing data collected on a large sample of residents from neighborhoods in an urban area of the United States with wide variability in neighborhood types. Finally, because the NEWS is a relatively long questionnaire (68 items) and response rates may be negatively affected by survey length (3), an additional purpose was to develop an abbreviated version of the NEWS (NEWS-A).

METHODS

Subjects

A stratified two-stage cluster sample design was used to recruit 1286 adults aged 20–65, all of whom were residents of private dwellings in King County, WA. The study sample was drawn from residential addresses within 16 selected neighborhoods (103 census blockgroups) and classified, based on their walkability characteristics and median household income, into four strata: high walkable/high socioeconomic status (SES); low walkable/high SES; high walkable/low SES; and low walkable/low SES. Neighborhood was defined as a cluster of adjacent blockgroups. Neighborhood walkability was determined using geographic information systems (GIS) data on four neighborhood attributes: residential density (number of residential units per acre), street connectivity (number of intersections per square kilometer), land use mix (evenness of distribution of building floor area of residential, retail, entertainment, office, and institutional development), and retail floor area ratio (ratio of retail building floor area to land area), with higher values of these characteristics indicating more walkable neighborhoods (12,13).

Census blockgroups within the county were ranked and divided into deciles based on a walkability index. The top four and bottom four deciles represented high-walkability and low-walkability areas. Census blockgroups in the county were also deciled by median household income, using U.S. Census data, and categorized into high income and low income. Areas with median household incomes less than $15,000 and greater than $150,000 were not included in the sampling frame to avoid extreme income values. The second, third, and fourth deciles constituted the low-income category, and the seventh, eighth, and ninth deciles made up the high-income category. Simple random sampling, without replacement, was used to select households from each neighborhood. Only one respondent per household was asked to participate. Individuals who were in group living establishments (e.g., nursing homes, military barracks), unable to walk without assistance, or unable to take part in surveys in English were excluded. Subjects' sociodemographic characteristics are shown in Table 1. Written informed consent was obtained from the subjects. This study was approved by the ethics committee of the local university.

Measures

Neighborhood environment walkability scale (NEWS). This 68-item instrument measured perceived attributes of the local environment hypothesized to be related to physical activity and, particularly, to walking for transport and walking for recreation. Concepts and subscales were based on variables believed to relate to walking and other physical activities that are discussed in the urban planning literature (11). The questionnaire assessed the following environmental characteristics: a) residential density; b) proximity to nonresidential land uses, such as restaurants and retail stores (land use mix–diversity); c) ease of access to nonresidential uses (land use mix–access); d) street

TABLE 1. Sociodemographic characteristics of the sample (N = 1286).

Characteristic	Estimate
Gender (%)	
Male	54.7
Missing values	0.2
Ethnicity (%)	
Caucasian	82.3
African American	3.3
Asian American	6.5
Pacific Islander	1.4
Native American	1.2
Other	4.7
Missing values	0.5
Marital status (%)	
Married	57.0
Widowed/divorced/separated	15.9
Single and never married	20.8
Living with partner	6.1
Missing values	0.2
Age (yr)	
Mean (SD)	44.0 (11.0)
Missing values (%)	0.2
Educational attainment (%)	
Primary or less	1.3
Secondary	35.4
Tertiary	60.1
Missing values	0.2
Annual household income (%)	
< $19,500	23.8
$19,500–$39,500	50.2
$39,500–$59,500	22.2
> $59,500	3.5
Missing values	0.2
Children in household (%)	
Yes	38.5
Missing values	0.2

connectivity; e) walking/cycling facilities, such as sidewalks and pedestrian/bike trails; f) aesthetics; g) pedestrian traffic safety; and h) crime safety. With the exception of the residential density and land use mix–diversity subscales, items were rated on a 4-point Likert scale from 1 (strongly disagree) to 4 (strongly agree). Residential density items asked about the frequency of various types of residences, from single-family detached homes to 13-story or higher apartments/condominiums, with a response range of 1 (none) to 5 (all). Residential density items were weighted relative to the average density of single-family detached residences (e.g., 7- to 12-story apartments and condominiums were considered to be 50 times more person-dense than single-family residences), and weighted values were summed to create a residential density subscale score. Land use mix–diversity was assessed by the walking proximity from home to various types of stores and facilities, with responses ranging from 1- to 5-min walking distance (coded as 5) to > 30-min walking distance (coded as 1). Higher scores on land use mix–diversity indicated closer average proximity. The NEWS and its abbreviated version are available at www.drjamessallis.sdsu.edu.

Self-reported walking for transport and recreation. Because the variables assessed in the NEWS are hypothesized to be differentially associated with walking for different purposes, self-reported walking was used as the validity criterion. Weekly minutes of walking for transport and walking for recreation were assessed using the long version of the International Physical Activity Questionnaire (IPAQ). Subjects were instructed to report the frequency and duration of walking for transport and recreation during the past 7 d. Weekly minutes of walking for transport and walking for recreation were computed, and outlier values were truncated to the 99th percentile. A recent report on data collected in 12 countries showed that the IPAQ had comparable reliability and validity to other self-report measures of physical activity (7). There is evidence for an acceptable degree of reliability and validity of the transportation and leisure-time physical activity items of the IPAQ, which include the walking items used in this study. For these items, intraclass correlations ranging from 0.60 and 0.82 have been reported (32). Also, moderate correlations (0.50–0.63) were found between diary measures of transport-related and leisure-time physical activity and the corresponding IPAQ items (32). In the present study, the blockgroup-level correlation between self-reported total weekly minutes of walking, and accelerometry-estimated weekly minutes of moderate-intensity activity was 0.54, whereas that at the individual-level was 0.19, indicating moderate criterion validity.

Sociodemographic characteristics. Subjects were asked to provide information on their age, sex, educational attainment, marital status, ethnicity, annual household income, and number of children (< 18 yr old) in the household.

Procedure

Households within the selected blockgroups were identified by a marketing company and were sent an introductory letter. Households were called within 2 wk of the expected receipt of this letter, with an adult in the household queried about interest and study eligibility. One interested and eligible adult per household was sent the consent form and, upon its return, was sent questionnaires with instructions and a postage-paid return envelope. The recruitment rate (subjects/eligible people contacted) was 28%.

Data Analytic Plan

Phase I: Confirmatory factor analysis of the long version of the NEWS. Confirmatory factor analysis (CFA) was conducted on the items from six of the eight subscales of the NEWS. These were the subscales using the same 4-point rating scales: land use mix–access, street/connectivity, infrastructure for walking/cycling, aesthetics, traffic safety, and crime safety. Given their formats, it was not appropriate to factor analyze the subscales of residential density and land use mix–diversity. For the items included in the CFA, univariate normality of the individual deviations from the blockgroup mean scores and the disaggregated blockgroup mean scores was examined. To establish whether a multilevel rather than a single-level CFA of the NEWS was needed, intraclass correlation coefficients (ICC) were computed for each of the NEWS items. It is generally maintained that when ICC exceed values of 0.10 and group sizes exceed 15, the multilevel structure of the data should be modeled (22).

Multilevel CFA of the six subscales of the NEWS was carried out using Bentler and Liang's maximum likelihood estimation (MLE) method (2). This method uses an expectation maximization–type gradient algorithm for computing the MLE for two-level structural equation models, of which confirmatory factor analysis is a special case. This algorithm is applicable to any study with balanced or unbalanced design and is preferable to other methods when the sample sizes vary substantially among clusters (blockgroups).

The construction of a two-level measurement model of the long version of the NEWS included two main steps. First, an *a priori* two-level measurement model with six oblique factors, as originally defined by its developers (24), was examined. For this model, the factor structures at the individual and blockgroup levels were defined to be equal. In the second step of the analyses, respecification of the original model was conducted according to Jöreskog and Sörbom's iterative model-generating approach (18). This approach consists of testing the viability of initial hypothetical models in terms of whether they satisfactorily fit the observed data. If the results indicate a lack of fit based on empirical or substantive evidence, the models are respecified. The ultimate goal of model respecification is to identify models that can provide a statistically acceptable fit and a theoretically meaningful interpretation of the data. Model respecification was based on the analysis of standardized factor loadings, the analysis of three empirical indices of poor model fit (standardized residual covariances, univariate Lagrange multiplier tests, and Wald tests), and substantive considerations (i.e., salience) (18). Factor loadings greater than |0.30| were considered to be significant (6).

Several measures of absolute and incremental fit were used to evaluate the goodness-of-fit of the measurement models (4,19). Absolute-fit indices describe the ability of the model to reproduce the original covariance matrix. The absolute-fit indices reported in this paper are the χ^2 test (specifically, the Bentler–Liang likelihood ratio statistic), the goodness-of-fit index (GFI), and the root mean square error of approximation (RMSEA) (4,19). Two incremental fit indices, assessing the degree to which a specified model is better than a baseline model that specifies no covariances, were used. These were the nonnormed fit index (NNFI) and the comparative fit index (CFI). In addition to the above indices of fit, the standardized root mean squared residual (SRMR) was computed. The SRMR is a standardized summary of the average covariance residuals (19). We also used the Aikake information criterion (AIC) of the standard goodness-of-fit χ^2 statistics that includes a penalty for complexity. This index can be used for the comparison of nonnested measurement models (models that are not subsets of one another) (19). Hu and Bentler's recommended cutoff values for specific fit indices were adopted (16). According to these authors, GFI, CFI, and NNFI with values exceeding 0.95 are generally indicative of a good model fit. A favorable value for the SRMR is less than 0.08; for RMSEA, a favorable value is less than 0.06. All CFA-related analyses were conducted using EQS 6.1 (Multivariate Software Inc., 2004).

Phase II: Selection of items for the NEWS-A. To create an abbreviated version of the NEWS, pairs of items overlapping in content were identified, and those with the better psychometric properties were selected. The following psychometric properties were taken into account:

1. Criterion validity: direction and magnitude of associations (correlations) between ratings on a specific item and reported weekly minutes of walking for transport and walking for recreation at the individual level (i.e., within blockgroups) and the blockgroup level (i.e., between blockgroups) after controlling for sociodemographic factors (age, gender, educational attainment, children in household, and annual household income). These correlations were computed as specified by Snijder and Bosker (29).
2. Contribution of a specific item to the criterion validity of its factor in relation to walking for transport and walking for recreation. This was established by examining the change in shared variance (R^2) between a factor and measures of walking following the exclusion of the item ($dR^2 = R^2$ of factor without item $- R^2$ of factor with item).
3. Magnitude of ICC (i.e., how much of the total item variance is due to differences between blockgroups).
4. Test–retest reliability (intraclass correlation observed in an earlier study) (24).
5. Magnitude of the standardized loading of a specific item on its corresponding underlying dimension (factor).

Higher values of the above criteria were considered to be desirable. In an overlapping pair of items, the item with a greater tally of higher values was retained. Any other items with low criterion validity, either as a single item (see #1 above) and as an element of a factor (see #2 above) or of potentially low theoretical salience, were excluded from the abbreviated scale (NEWS-A).

Phase III: Confirmatory factor analysis of the NEWS-A. Confirmatory factor analysis of the NEWS-A was conducted using the same procedure outlined in phase I (CFA of the long version of the NEWS) by 1) testing the fit of a truncated measurement model with a structure equivalent to that identified for the long version of the NEWS and 2) defining a respecified model providing a statistically acceptable fit and a theoretically meaningful interpretation of the data.

Phase IV: Criterion validity of the NEWS and NEWS-A. Criterion validity of the multilevel CFA-derived factors of the NEWS and NEWS-A and the residential density and land use mix–diversity scales were determined by examining the individual- and blockgroup-level associations between scores on the identified factors/scales and walking for transport and recreation. Individual- and blockgroup-level correlations were estimated following the procedure specified by Snijder and Bosker (29). The associations between self-reported walking and the NEWS and NEWS-A factors/scales were estimated after controlling for sociodemographic confounders (partial correlations). Huber–White sandwich estimators of standard errors were used to compute the associations with the nonnormal variables of walking for transport and walking for recreation (33). The same analyses were carried out using normalized variables. Normalization involved assigning expected values from the standard normal distribution according to the ranks of the original values so that normal equivalent deviates were computed for (i-0.5)/n where i is the rank of the original values and n is the total number of values (23). In case of no substantial difference between the two sets of analyses, the results with the original outcome values were reported. These analyses

TABLE 2. Results of the multilevel CFA of the NEWS and NEWS-A.

Model	χ^2	*df*	GFI	RMSEA (90% CI)	SRMR	NNFI	CFI	AIC
NEWS								
Model 1: *A priori*	5701.5	1374	0.92	0.050 (0.048–0.051)	0.074	0.81	0.82	2954
Model 2: Respecified	3400.2	1135	0.98	0.040 (0.038–0.041)	0.063	0.92	0.92	1130
NEWS-A								
Model 1a: NEWS based*	1052.9	442	1.00	0.033 (0.030–0.035)	0.052	0.97	0.97	169
Model 2a: Respecified	1020.7	445	1.00	0.032 (0.029–0.034)	0.067	0.97	0.97	131

* Based on the respecified measurement model of the NEWS (model 2).

TABLE 3. Standardized factor loadings and uniquenesses for final respecified individual-level and neighborhood-level measurement models of the NEWS and NEWS-A (in parentheses).

Item #	Item	Individual Level					Blockgroup Level				
		Standardized Loading		Standardized Uniqueness		Latent Factor	Standardized Loading		Standardized Uniqueness		Latent Factor
A1	I can do most of my shopping at local stores	0.68	(—)	0.54	(—)	IL1 (—)	0.80	(—)	0.36	(—)	BL1 (—)
A2	Stores are within easy walking distance at my home	0.89	(0.76)	0.21	(0.43)	IL1 (IL1$_A$)	0.88	(0.80)	0.23	(0.36)	BL1 (BL1$_A$)
A3	Parking is difficult in local shopping areas	Max std load = −0.11 on factor IL1 (−0.10 on factor IL1$_A$)					0.50	(0.54)	0.75	(0.71)	BL1 (BL1$_A$)
A4	There are many places to go within walking distance at my home	0.74	(0.85)	0.46	(0.28)	IL1 (IL1$_A$)	0.86	(0.80)	0.26	(0.36)	BL1 (BL1$_A$)
A5	It is easy to walk to a transit stop (bus, train) from my home	0.45	(0.50)	0.80	(0.75)	IL1 (IL1$_A$)	0.80	(0.81)	0.36	(0.35)	BL1 (BL1$_A$)
A6	The streets in my neighborhood are hilly, making my neighborhood difficult to walk in	Max std load = −0.17 on factor IL1 (−0.15 on factor IL1$_A$)					0.72	(0.78)	0.48	(0.40)	BL2 (BL2$_A$)
A7	There are many canyons/hillsides in my neighborhood that limit the number of routes for getting from place to place	Max std load = 0.18 on factor IL1 (0.21 on factor IL1$_A$)					0.79	(0.86)	0.37	(0.26)	BL2 (BL2$_A$)
B1	The streets in my neighborhood do not have many cul-de-sacs	Max std load = 0.16 on factor IL2 (0.16 on factor IL2$_A$)					0.78	(0.74)	0.40	(0.45)	BL1 (BL1$_A$)
B2	There are walkways in my neighborhood that connect cul-de-sacs to streets, trails, or other cul-de-sacs	Max std load = 0.15 on factor IL2 (—)					Max std load = 0.17 on factor BL1 (—)				
B3	The distance between intersections in my neighborhood is usually short	0.43	(0.37)	0.81	(0.86)	IL2 (IL2$_A$)	0.89	(0.89)	0.20	(0.20)	BL1 (BL1$_A$)
B4	There are many four-way intersections in my neighborhood.	0.46	(—)	0.79	(—)	IL2 (—)	0.95	(—)	0.10	(—)	BL1 (—)
B5	There are many alternative routes for getting from place to place in my neighborhood	0.61	(0.62)	0.63	(0.61)	IL2 (IL2$_A$)	0.98	(0.96)	0.04	(0.08)	BL1 (BL1$_A$)
C1	There are sidewalks on most of the streets in my neighborhood	0.61	(0.49)	0.62	(0.76)	IL3 (IL3$_A$)	0.70	(0.74)	0.51	(0.45)	BL1 (BL1$_A$)
C2	The sidewalks in my neighborhood are well maintained	0.63	(—)	0.61	(—)	IL3 (—)	−0.74	(—)	0.45	(—)	BL2 (—)
C3	There are bicycle or pedestrian trails in or near my neighborhood that are easy to get to	0.45	(—)	0.80	(—)	IL3 (—)	−0.50	(—)	0.75	(—)	BL2 (—)
C4	Sidewalks are separated from the road/traffic in my neighborhood by parked cars	0.38	(0.36)	0.85	(0.86)	IL3 (IL3$_A$)	0.97	(0.98)	0.06	(0.05)	BL1 (BL1$_A$)
C5	There is a grass/dirt strip that separates the streets from the sidewalks in my neighborhood	0.30	(0.33)	0.90	(0.89)	IL3 (IL3$_A$)	0.81	(0.83)	0.35	(0.32)	BL1 (BL1$_A$)
C6	It is safe to ride a bike in or near my neighborhood	0.61	(—)	0.62	(—)	IL3 (—)	−0.82	(—)	0.32	(—)	BL2 (—)
D1	There are trees along the streets in my neighborhood	0.32	(0.34)	0.90	(0.88)	IL4 (IL4$_A$)	0.89	(0.85)	0.21	(0.29)	BL3 (BL3$_A$)
D2	Trees give shade for the sidewalks in my neighborhood	0.34	(—)	0.88	(—)	IL4 (—)	0.82	(—)	0.34	(—)	BL3 (—)
D3	There are many interesting things to look at while walking in my neighborhood	0.69	(0.72)	0.53	(0.48)	IL4 (IL4$_A$)	0.94	(0.96)	0.12	(0.08)	BL3 (BL3$_A$)
D4	My neighborhood is generally free from litter	0.49	(—)	0.76	(—)	IL4 (—)	0.79	(—)	0.37	(—)	BL3 (—)
D5	There are many attractive natural sights in my neighborhood	0.76	(0.76)	0.43	(0.42)	IL4 (IL4$_A$)	0.95	(0.95)	0.10	(0.10)	BL3 (BL3$_A$)
D6	There are attractive buildings/homes in my neighborhood	0.74	(0.72)	0.46	(0.48)	IL4 (IL4$_A$)	0.97	(0.96)	0.06	(0.07)	BL3 (BL3$_A$)
E1	There is so much traffic along the street I live on that it makes it difficult or unpleasant to walk in my neighborhood	0.72	(—)	0.48	(—)	IL5 (—)	0.84	(—)	0.29	(—)	BL4 (—)
E2	There is so much traffic along nearby streets that it makes it difficult or unpleasant to walk in my neighborhood	0.70	(0.71)	0.51	(0.49)	IL5 (IL5$_A$)	0.96	(0.95)	0.08	(0.11)	BL4 (BL4$_A$)
E3	The speed of traffic on the street I live on is usually slow	−0.61	(—)	0.63	(—)	IL5 (—)	−0.76	(—)	0.42	(—)	BL4 (—)
E4	The speed of traffic on most nearby streets is usually slow	−0.59	(−0.64)	0.65	(0.59)	IL5 (IL5$_A$)	−0.92	(−0.90)	0.15	(0.18)	BL4 (BL4$_A$)
E5	Most drivers exceed the posted limits while driving in my neighborhood	0.46	(0.45)	0.79	(0.80)	IL5 (IL5$_A$)	0.93	(0.90)	0.13	(0.18)	BL4 (BL4$_A$)
E6	My neighborhood is well lit at night	0.43	(0.50)	0.81	(0.75)	IL3 (IL3$_A$)	−0.76	(−0.77)	0.42	(0.41)	BL4 (BL4$_A$)
E7	Walkers and bikers on the streets in my neighborhood can be easily seen by people in their homes	0.39	(0.46)	0.85	(0.79)	IL3 (IL3$_A$)	−0.83	(−0.83)	0.32	(0.31)	BL4 (BL4$_A$)
E8	There are crosswalks and pedestrian signals to help walkers cross busy streets in my neighborhood	0.35	(0.34)	0.88	(0.88)	IL3 (IL3$_A$)	0.89	(0.74)	0.20	(0.45)	BL1 (BL1$_A$)
E9	The crosswalks in my neighborhood help walkers feel safe crossing busy streets	0.43	(—)	0.81	(—)	IL3 (—)	0.93	(—)	0.13	(—)	BL1 (—)
E10	When walking in my neighborhood there are a lot of exhaust fumes	0.50	(—)	0.75	(—)	IL5 (—)	0.81	(—)	0.35	(—)	BL5 (—)
F1	I see and speak to other people when I am walking in my neighborhood	Max std load = −0.19 on factor IL6 (—)					0.58	(—)	0.67	(—)	BL3 (—)
F2	There is a high crime rate in my neighborhood	0.68	(0.67)	0.54	(0.55)	IL6 (IL6$_A$)	0.99	(0.98)	0.02	(0.04)	BL5 (BL5$_A$)
F3	The crime rate in my neighborhood makes it unsafe to go on walks during the day	0.55	(0.54)	0.70	(0.71)	IL6 (IL6$_A$)	0.92	(0.92)	0.17	(0.16)	BL5 (BL5$_A$)
F4	The crime rate in my neighborhood makes it unsafe to go on walks at night	0.82	(0.83)	0.34	(0.31)	IL6 (IL6$_A$)	0.97	(0.97)	0.06	(0.05)	BL5 (BL5$_A$)
F5	My neighborhood is safe enough so that I would let a 10-yr-old boy walk around my block alone in the daytime	−0.39	(—)	0.85	(—)	IL6 (—)	0.91	(—)	0.17	(—)	BL3 (—)

Max std load, maximal standardized loading; (—), not applicable.

A priori factors: A, land use mix–access; B, street connectivity; C, infrastructure for walking/cycling; D, aesthetics; E, traffic safety; F, crime safety.

Latent individual-level factors: IL1 and IL1$_A$, land use mix–access; IL2 and IL2$_A$, street connectivity; IL3, infrastructure and safety for walking/cycling; IL3$_A$, infrastructure and safety for walking; IL4 and IL4$_A$, aesthetics; IL5 and IL5$_A$, traffic hazards; IL6 and IL6$_A$, crime.

Latent blockgroup-level factors: BL1 and BL1$_A$, land use mix–access and infrastructure for walking; BL2, physical obstacles to walking/cycling; BL2$_A$, physical obstacles to walking; BL3, aesthetics and friendliness; BL3$_A$, aesthetics; BL4 and BL4$_A$, traffic hazards; BL5 and BL5$_A$, crime.

Autocorrelated within-factor error terms were modeled for the following items: D1 and D2 ($r = 0.60$; $t = 19.5$; $P < 0.001$), and E8 and E9 ($r = 0.65$; $t = 20.4$; $P < 0.001$).

TABLE 4. Correlations between individual-level latent factors of the NEWS (above the diagonal) and NEWS-A (below the diagonal).

NEWS-A Factors		(IL2)	(IL3)	(IL4)	(IL5)	(IL6)	NEWS Factors
Land use mix–access ($IL1_A$)		0.33	0.23	0.20	< 0.10*	< 0.10*	Land use mix–access (IL1)
Street connectivity ($IL2_A$)	0.35		0.41	0.23	−0.21	< 0.10*	Street connectivity (IL2)
Infrastructure and safety for walking ($IL3_A$)	0.28	0.54		0.48	−0.54	−0.35	Infrastructure and safety for walking/cycling (IL3)
Aesthetics ($IL4_A$)	0.23	0.30	0.44		−0.41	−0.38	Aesthetics (IL4)
Traffic hazards ($IL5_A$)	< 0.10*	−0.27	−0.49	−0.36		0.56	Traffic hazards (IL5)
Crime ($IL6_A$)	< 0.10*	< 0.10*	−0.28	−0.30	0.48		Crime (IL6)
	($IL1_A$)	($IL2_A$)	($IL3_A$)	($IL4_A$)	($IL5_A$)		

* Constrained to zero in the final model as correlation coefficients smaller than |0.10|. The subscript A stands for NEWS-A.

were conducted using MLwiN version 2 (Multilevel Models Project, Institute of Education, 2004).

The scores on the residential density and land use mix–diversity scales were calculated as explained by Saelens et al. (24). Participants' scores on each of the individual-level CFA-derived factors were defined in two different ways: 1a) as their average rating on the items loading on the specific factor, and 1b) as the deviation of their score from the mean score of their blockgroup for the specific factor (equation 1). Specifically, the score on an individual-level factor for the ith participant ($Y_{ILF.i}$) residing in the jth blockgroup was computed as

$$Y_{ILF}.i = \sum[x_i j.k - \text{mean}(x_j.k)]/l \qquad [1]$$

where mean ($x_{j.k}$) is the average score on the kth item for the jth blockgroup, $x_{ij.k}$ is the ith resident's (from blockgroup j) score on the kth item of a specific individual-level factor (ILF), and l is the total number of items loading on the specific ILF. Participants' scores on the blockgroup-level factors were defined as 2a) their average rating on the items loading on the specific factor, and 2b) the mean score of their blockgroup on the specific factor (equation 2). The score on a blockgroup-level factor (BLF) for the ith participant ($Y_{BLF.i}$) residing in the jth blockgroup was computed as

$$Y_{BLF}.i = \sum[\text{mean}(x_j.k)]/l \qquad [2]$$

representing the mean score on a blockgroup-level factor for the jth blockgroup (i.e., all residents from the same blockgroup are assigned the same score on a specific blockgroup-level factor).

In practice, scores based on the individual-level factors operationalized as in 1a provide measures that best differentiate individual perceptions of the same environment. In contrast, scores based on the blockgroup-level factors operationalized as in 2a provide measures that best differentiate between perceptions of residents from different blockgroups and are to be used when the focus is more on the actual environment rather than individual perceptions of the environment (see Discussion section for more details). The computation of scores on the individual-level factors as individual deviations from the blockgroup mean (as per 1b) and of scores on the blockgroup-level factors as the mean score for a specific blockgroup (as per 2b) is recommended when a researcher wishes to separately estimate interindividual differences in the perception of a same blockgroup (represented by 1b) and actual differences between blockgroups (represented by 2b). This approach facilitates the analysis of the independent effects of idiosyncratic perceptions of the environment and the actual environment on outcomes of interest (e.g., walking for transport).

RESULTS

Phase I: Confirmatory factor analysis of the long version of the NEWS. In the present study, the number of respondents per census blockgroup ranged from 1 to 46. The ICC of the NEWS items were small to moderate in size and ranged from 0.02 to 0.49. The mean ICC was 0.23 (SD = 0.12). These results confirmed the need for a multilevel CFA of the NEWS using Bentler–Liang's MLE method. Items' univariate skewness and kurtosis values were within acceptable limits for the use of maximum likelihood estimation (8).

The *a priori* measurement model of the NEWS exhibited an unacceptable level of fit, with two indices (RMSEA and SRMR) meeting, and two indices being substantially lower than the adopted cutoff values (Table 2, part a). In contrast, the final two-level respecified model showed a relatively good fit to the data, with three fit indices meeting, and the remaining two approaching the adopted cutoff values. Inspection of the standardized factor loadings and modification indices of the *a priori* model suggested that, at the individual level, items A3, A6, A7, B1, B2, and F1 did not substantively load on any of the factors (Table 3).

Items E6–E9 showed a stronger association with the infrastructure for walking/cycling than, as originally

TABLE 5. Correlations between blockgroup-level latent factors of the NEWS (above the diagonal) and NEWS-A (below the diagonal).

		(BL2)	(BL3)	(BL4)	(BL5)	NEWS Factors
Land use mix–access and infrastructure for walking ($BL1_A$)		−0.41	0.38	−0.72	0.25	Land use mix–access and infrastructure for walking (BL1)
Physical obstacles to walking ($BL2_A$)	−0.41		−0.13	0.43	0.20	Physical obstacles to walking/cycling (BL2)
Aesthetics ($BL3_A$)	0.62	< 0.10*		−0.75	−0.70	Aesthetics and friendliness (BL3)
Traffic hazards ($BL4_A$)	−0.62	0.26	−0.82		0.30	Traffic hazards (BL4)
Crime ($BL5_A$)	< 0.10*	< 0.10*	−0.69	0.40		Crime (BL5)
	$BL1_A$	$BL2_A$	$BL3_A$	$BL4_A$		

* Constrained to zero in the final model as correlation coefficients smaller than |0.10|. The subscript A stands for NEWS-A.

TABLE 6. Partial correlations (r_p) between walking for transport, walking for recreation, and the factors/scales of the NEWS and NEWS-A (in parentheses).

	Method of Factor/Scale/Item Scoring				
	Disaggregated Blockgroup Mean Score on Factor (Equation #2)	Individual Deviations from Blockgroup Average Score on Factor (Equation #1)	Individual Average Rating on Factor's Items (Conventional Scoring)/Single Items and Predefined Scales		
Factor/Scale (Label)	Walking for Transport r_p Blockgroup Level	Walking for Transport r_p Individual Level	Walking for Transport r_p Blockgroup Level	Walking for Transport r_p Individual Level	Walking for Recreation r_p Individual Level
Block-Level Factors					
Land use mix–access and infrastructure for walking ($BL1_{(A)}$)	0.72‡ (0.73‡)	—	0.70‡ (0.71‡)	0.05 (0.06*)	0.00 (0.03)
Physical obstacles to walking/cycling ($BL2_{(A)}$)	−0.12 (−0.06)	—	−0.12 (−0.11)	0.00 (-0.04)	0.00 (−0.03)
Aesthetics (and friendliness) ($BL3_{(A)}$)	0.41‡ (0.53‡)	—	0.43‡ (0.52‡)	0.13‡ (0.11‡)	0.04 (0.08*)
Traffic hazards ($BL4_{(A)}$)	−0.55‡ (−0.62‡)	—	−0.32‡ (−0.41‡)	0.00 (0.00)	0.10‡ (0.02)
Crime ($BL5_{(A)}$)	0.26† (0.21†)	—	0.28† (0.21†)	0.04 (0.09†)	0.09† (0.08*)
Individual-Level Factors					
Land use mix–access ($IL1_{(A)}$)	—	0.09† (0.09†)	0.48‡ (0.48‡)	0.10† (0.09†)	0.03 (0.04)
Street connectivity ($IL2_{(A)}$)	—	0.03 (0.06)	0.61‡ (0.61‡)	0.00 (0.05)	0.01 (0.03)
Infrastructure and safety for walking/cycling ($IL3_{(A)}$)	—	0.00 (0.03)	0.46‡ (0.59‡)	0.00 (0.00)	0.00 (0.00)
Aesthetics ($IL4_{(A)}$)	—	0.12‡ (0.12‡)	0.44‡ (0.52‡)	0.12‡ (0.11‡)	0.06* (0.08*)
Traffic hazards ($IL5_{(A)}$)	—	−0.04 (−0.01)	−0.16 (−0.33‡)	−0.03 (0.00)	0.11‡ (−0.02)
Crime ($IL6_{(A)}$)	—	0.04 (0.05)	0.17* (0.21†)	0.06† (0.09†)	0.07* (0.08*)
Predefined Scales					
Residential density (RD)	—	—	0.80‡ (0.80‡)	0.00 (0.00)	0.17‡ (0.17‡)
Land use mix–diversity (LUM-D)	—	—	0.53‡ (0.53‡)	0.07* (0.07*)	0.07* (0.07*)
Single Items					
Parking is difficult in local shopping areas (A3)	—	—	0.68‡ (0.68‡)	0.00 (0.00)	0.04 (0.04)
The streets in my neighborhood are hilly, making my neighborhood difficult to walk in (A6)	—	—	0.00 (0.00)	−0.03 (−0.03)	−0.04 (−0.04)
There are many canyons/hillsides in my neighborhood that limit the number of routes for getting from place to place (A7)	—	—	−0.22* (−0.22*)	−0.03 (−0.03)	−0.01 (−0.01)
The streets in my neighborhood do not have many cul-de-sacs (B1)	—	—	0.12 (0.12)	0.00 (0.00)	0.00 (0.00)
I see and speak to other people when I am walking in my neighborhood (F1)	—	—	−0.09	0.16‡	0.08†

* $P < 0.05$; † $P < 0.01$; ‡ $P < 0.001$.

hypothesized, with the traffic and safety *a priori* factor, thus forming an infrastructure and safety for walking factor. This finding is not surprising because items E6–E9 describe pedestrian-related safety features of the local area rather than presence of motorized traffic. All of the remaining items that were hypothesized to gauge traffic safety (E1–E5, E10) pertained to presence of motorized traffic. Items D1 and D2 (trees in the local area) and E8 and E9 (crossways to help cross busy streets) had correlated uniqueness.

In the final individual-level measurement model, all factor loadings and uniquenesses were significant at the 0.001 probability level (Table 3). The interrelationships between the individual-level factors of the NEWS are reported in Table 4 (above diagonal). In general, at the individual (within-blockgroup) level, support was found for the hypothesized six-factor measurement model of the NEWS, although not all items loaded on the expected factor.

At the blockgroup level, an analysis of factor loadings, modification indices, and residual covariance matrix suggested a different measurement model from that at the individual level. Five, rather than six, oblique factors were identified. These were land use mix–access and infrastructure for walking, physical obstacles to walking/cycling, aesthetics and friendliness, traffic hazards, and crime (Table 3). Weak to strong associations were observed between the latent blockgroup-level factors (Table 5). Similarly to what was observed for the individual-level measurement model, item B2 did not significantly load on any of the latent blockgroup-level factors. All other items' standardized loadings were significant at the 0.001 level.

Phase II: Selection of items for NEWS-A. Following the previously outlined procedure (see Data analytic plan: phase II), items A1, B2, B4, C2, C3, C6, D2, D4, E1, E3, E9, E10, F1, and F5 were excluded from the NEWS-A (Table 2; part b). Details on the psychometric characteristics of these items are available at www.drjamessallis.sdsu.edu.

Phase III: Confirmatory factor analysis of the NEWS-A. The two-level measurement model of the NEWS-A based on the multilevel CFA of the NEWS (see phase I) showed a good level of fit (Table 2, part b). An analysis of the interfactor correlations, standardized loadings, and modification indices suggested that no alterations to the model were needed at the individual level. At the blockgroup-level, the pattern of interfactor relationships called for some modifications (independence of the factors land use mix–access and infrastructure for walking, physical obstacles to walking, and crime). Although the factors traffic hazards and aesthetics were highly negatively correlated (r = −0.86), they were not combined into one factor because our goal was to develop an abbreviated NEWS with a measurement model comparable with its original version. The respecified model of the NEWS-A yielded a slightly better level of model fit according to most indices (Table 2, part b). All standardized loadings and interfactor correlations in the final two-level measurement model of the NEWS-A were significant at the 0.001 level (Table 3; numbers in brackets). The correlations between the scores on the NEWS and NEWS-A factors were very high and ranged between 0.82 and 0.98 at the blockgroup level and between 0.83 and 0.97 at the individual level.

Phase IV: Criterion validity of the NEWS and NEWS-A. Respondents reported, on average, 118 weekly minutes of walking for recreation (median = 60; SD = 190), and 163 weekly minutes of walking for transport (median = 60; SD = 289). All of the variance in walking for recreation was attributable to (within-blockgroup) differences between individuals. In contrast, approximately 5% of the total variance of walking for transport was due to differences between blockgroups.

Factors/scales measuring access/presence of destinations ($BL1_{(A)}$, $IL1_{(A)}$, and land use mix–diversity (LUM-D)), street connectivity ($IL2_{(A)}$), infrastructure for walking ($BL2_{(A)}$ and $IL3_{(A)}$), residential density (RD), and aesthetics ($BL3_{(A)}$ and $IL4_{(A)}$) were expected to be positively correlated, and crime ($BL5_{(A)}$ and $IL6_{(A)}$) and traffic hazards ($BL4_{(A)}$ and $IL5_{(A)}$) were negatively correlated with walking for transport. In general, this was found to be true (Table 6, walking for transport section). However, a positive relationship was found between walking for transport and the crime factors. With respect to walking for recreation, positive associations were expected with infrastructure for walking ($IL3_{(A)}$), aesthetics ($BL3_{(A)}$ and $IL4_{(A)}$), and presence of destinations (LUM-D), whereas negative associations were expected with crime ($BL5_{(A)}$ and $IL6_{(A)}$) and traffic hazards ($BL4_{(A)}$ and $IL5_{(A)}$). Partial support was found only for the hypothesized relationships with aesthetics and presence of destinations (Table 6, walking for recreation section). In contrast, weak but positive associations were observed between crime, traffic hazards, and walking for recreation.

DISCUSSION

The present study examined the factorial and criterion validity of the original NEWS and the abbreviated NEWS-A in a large sample of adults. The type of sampling design and the presence of meaningful cluster effects called for the application of a multilevel approach to the confirmatory factor analysis (CFA) (2,22,31). The end result of a multilevel CFA is the estimation of a measurement model for each level of variation. In this study, two measurement models of the NEWS and NEWS-A were estimated: one at the individual level (based on within-blockgroup variations in the responses to the items), the other at the blockgroup level (based on between-blockgroup variations in the responses to the items).

Although similar, the two measurement models of the NEWS (and NEWS-A) were not equivalent. These findings indicated the environmental attributes measured by the items of NEWS and NEWS-A might group in different ways within and across neighborhoods/blockgroups in the geographic area examined in this study. For instance, blockgroups reported to offer a greater number of commercial and noncommercial destinations were also reported to have better street connectivity (i.e., intersections, alternative routes) and infrastructure for walking. These three attributes formed a single factor at the blockgroup level. In contrast, at the within-blockgroup (individual) level, these characteristics were not as strongly associated and formed separate low to moderately intercorrelated factors. These findings might be an expression of the way environmental attributes objectively covary within the same census blockgroup (e.g., areas with a high level of street connectivity may vary in the number and types of destinations they offer to individuals within a given blockgroup). However, the observed differences between the blockgroup- and individual-level measurement models also may be due partly to the different mechanisms determining the grouping of the items at the two levels of variations. Specifically, the blockgroup-level measurement model is likely to be more reflective of the way environmental attributes group objectively because it is based on the blockgroup average scores of the items across residents. The average resident rating is likely to be a more reliable and valid measure of the objective environment than a single resident rating.

Besides being reflective of the objective covariation of environmental attributes within a blockgroup, the individual-level measurement model may be determined by perceptual biases. For instance, the relationship between crime and traffic hazards blockgroup-level latent factors ($BL4_{(A)}$ and $BL5_{(A)}$) was consistently lower than that between the corresponding individual-level latent factors ($IL5_{(A)}$ and $IL6_{(A)}$). This could be due to respondents higher in trait anxiety giving higher ratings to all items describing potentially threatening stimuli, such as those loading on the crime and traffic hazards factors (10), even though these, in reality, were not correlated. In such case, the observed relationship between crime and traffic hazards would be stronger at the individual than at the blockgroup level. This is because the individual-level factors describe individual differences in perception of the local environment.

The above discussion calls for explicit recommendations on how to score the NEWS and NEWS-A for various study purposes. One alternative would be to use both blockgroup- and individual-level factors, with the former computed as disaggregated blockgroup mean scores and the latter computed as individual deviations from the blockgroup means. Such an approach could help differentiate between objective environmental characteristics and residents' idiosyncratic perceptions of the environment, which, in turn, could aid evaluation of their independent associations with physical activity. However, caution is warranted because of the unknown generalizability of present study measurement models to other locations. The conceptual soundness of the individual-level measurement model suggests that the individual-level scales would be applicable to other settings. However, cross-validation, particularly of the blockgroup-level factors to samples in other geographical locations, is certainly needed. The observed blockgroup-level measurement models were based on a nonrepresentative sample of census blockgroups, where only high and low, but not moderate, walkable blockgroups were selected. This likely increased the associations between certain environmental characteristics (e.g., street connectivity and land use mix). Secondly, patterns of association between environmental characteristics (e.g., high street connectivity predictive of high land-use mix) may vary across urban and other (e.g., rural) areas. If the blockgroup-level measurement models are shown to vary across locations, researchers are encouraged to carry out a multilevel CFA on their data and apply the measurement model they observe. Of course, this alternative solution applies to studies adopting a two-stage cluster sampling design with a sufficient sample size of residents and geographic units such as census blockgroups or neighborhoods (> 100) (15).

Presently, it is recommended that the NEWS and NEWS-A be scored according to the individual-level measurement model, particularly for studies that do not adopt a two-stage cluster-sampling design, are limited in size, and focus on the effect of perceived rather than objective and perceived environment on physical activity. In this case, the NEWS and NEWS-A would consist of eight subscales (residential density, land use mix–diversity, land use mix–access, street connectivity, infrastructure and safety for walking/cycling, aesthetics, traffic hazards, and crime) and five (for the NEWS) or four (for the NEWS-A) single items (access to parking, hilly streets, physical/natural obstacles, not many cul-de-sacs, and interaction with neighbors). These individual-level scales are clearly related to constructs commonly used in the urban planning and transportation fields (11). Findings based on these scales can be linked with specific policies that could improve the activity-friendliness of neighborhoods.

Overall, this study supported the construct validity of the NEWS (and NEWS-A). Six intercorrelated latent factors were identified at the individual level. In general, the observed factors matched the original hypothesized measurement model, although not all items loaded on the *a priori* factors. Importantly, the modifications made to the *a priori* measurement model of the NEWS were substantively justifiable. Convergent and divergent validity evidence was found for the observed factors in relation to walking for transport. As expected, less support was found for the criterion validity of the NEWS and NEWS-A with respect to walking for recreation. However, most of the NEWS constructs were designed to be related to active transportation, and other surveys have been designed to assess environmental factors relevant for active recreation (24). When compared with the NEWS, the measurement model of the NEWS-A showed a better fit to the data and marginally better criterion validity with respect to walking for transport. Because it is possible that these results might be specific to this location or sample, future studies need to establish whether the NEWS-A is consistently a better instrument to assess neighborhood walkability than the NEWS. At this stage, we recommend that researchers use the NEWS-A rather than the NEWS whenever participant burden is a significant concern.

It is important to mention that, although only weak associations between walking and the NEWS (and NEWS-A) factors/items were observed at the individual level, they were moderate to strong at the blockgroup level. Importantly, some individual differences in the responses to the NEWS (and the IPAQ) are due to measurement error. This indicates that the strength of the relationship between environmental attributes and physical activity behavior may be considerably greater than previously noted or suspected. However, even if these associations were truly weak, environmental factors would still need to be considered as a key component of the public health agenda. Public health effects depend on the effect size, number of people exposed, and duration of exposure. For instance, it has been estimated that about 50% of Americans live in low-density, low-walkable areas (28). Thus, the weak effect sizes are magnified by high exposure and the fact that people are exposed to neighborhood attributes every day over many years. The potential public health burden linked with low-walkable neighborhoods needs to be carefully studied to inform policy decisions. The development of high-quality measures is the first step toward this final goal.

The authors would like to thank Kerri Cain and Terry Conway for their contribution to the project and acknowledge that, at the time of the write-up of this paper, Ester Cerin was supported by NHMRC Grant #252977 and worked at the Cancer Prevention Research Centre (School of Population Health, The University of Queensland, Australia). This work was also supported by NIH grant HL67350.

REFERENCES

1. ATKINSON, J. L., J. F. SALLIS, E. SAELENS, L. CAIN, and J. B. BLACK. The association of neighborhood design and recreational environments with physical activity. *Am. J. Health Promotion* 19: 304–309, 2005.
2. BENTLER, P. M., and J. LIANG. Two-level mean and covariance structures: maximum likelihood via an EM algorithm, In: *Multilevel Modeling: Methodological Advances, Issues, and Applications*, S. P. Reise and N. Duan (Eds.). Mahwah, NJ: Lawrence Erlbaum, pp. 53–70, 2003.
3. BINER, P. M., and H. J. KIDD. The interactive effects of monetary incentive justification and questionnaire length on mail survey response rates. *Psychol. Market* 11:483–492, 1994.
4. BROWNE, M. W., and R. CUDECK. Alternative ways of assessing model fit. In: *Testing Structural Equation Models*, K. A. Bolennn and J. S. Long (Eds.). Newbury Park, CA: Sage Publications, pp. 136–162, 1993.
5. BROWNSON, R. C., J. J. CHANG, A. A. EYLER, et al. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. *Am. J. Public Health* 94:473–483, 2004.

6. BRYANT, F. B., and P. R. YARNOLD. Principal-components analysis and exploratory and confirmatory factor analysis. In: *Reading and Understanding Multivariate Statistics*, L. G. Grimm and P. R. Yarnold. Washington, DC: APA, pp. 99–136, 1994.
7. CRAIG, C. L., A. L. MARSHALL, M. SJÖSTRÖM, et al. International Physical Activity Questionnaire (IPAQ): 12-country reliability and validity. *Med. Sci. Sports Exerc.* 35:1381–1395, 2003.
8. CURRAN, P. J., S. G. WEST, and J. F. FINCH. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol. Methods* 1:16–29, 1996.
9. DE BOURDEAUDHUIJ, I., J. F. SALLIS, and B. E. SAELENS. Environmental correlates of physical activity in a sample of Belgian adults. *Am. J. Health Promotion* 18:83–92, 2003.
10. EDELMANN, R. J. *Anxiety Theory, Research and Intervention in Clinical and Health Psychology*. New York: Wiley, pp. 1–349, 1992.
11. FRANK, L. D., P. O. ENGELKE, and T. L. SCHMID. *Health and Community Design: The Impact of the Built Environment on Physical Activity*. Washington, DC: Island, pp. 99–177, 2003.
12. FRANK, L. D., J. F.SALLIS, T. L. CONWAY, et al. Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass and air quality. *J. Am. Planning Assoc.* 72:75–87, 2006.
13. FRANK, L. D., T. L. SCHMID, J. F. SALLIS, and J. E. CHAPMAN. Linking objectively measured physical activity data with objectively measured urban form: findings from SMARTRAQ. *Am. J. Prev. Med.* 28(Suppl. 2):117–125, 2005.
14. HEATH, G. W., R. C. BROWNSON, J. KRUGER, et al. The effectiveness of urban design and land use and transport policies and practices to increase physical activity: a systematic review. *J. Phys. Act. Health* 3(Suppl. 1):S55–S76, 2006.
15. HOX, J. *Multilevel Analysis: Techniques and Applications*. Mahwah, NJ: Laurence Erlbaum, pp. 173–196, 2002.
16. HU, L., and P. M. BENTLER. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equat. Model* 6:1–55, 1999.
17. HUMPEL, N., N. OWEN, and E. LESLIE. Environmental factors associated with adults' participation in physical activity: a review. *Am. J. Prev. Med.* 22:188–199, 2002.
18. JÖRESKOG, K. G., and D. SÖRBOM. *LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language*. Chicago, IL: SPSS, pp. 1–186, 1993.
19. KLINE, R. B. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press, pp. 165–208, 2005.
20. LESLIE, E., B. SAELENS, L. FRANK, et al. Residents' perceptions of walkability attributes in objectively different neighbourhoods: a pilot study. *Health Place* 11:227–236, 2005.
21. MCLEROY, K., D. BIBEAU, A. STECKLER, and K. GLANZ. An ecological perspective on health promotion programs. *Health Educ. Q.* 15:351–377, 1988.
22. MUTHÉN, B. O. Latent variable modeling of longitudinal and multilevel data. In: *Sociological Methodology*, A. E. Raftery (Eds.). Washington, DC: ASA, pp. 453–481, 1997.
23. RASBASH, J., F. STEELE, W. BROWNE, and B. PROSSER. *A User's Guide to MLwiN (version 2)*. London, UK: CMM, IE, University of London, pp. 136–153, 2004.
24. SAELENS, B. E., F. SALLIS, B. BLACK, and D. CHEN. Neighborhood-based differences in physical activity: an environmental scale evaluation. *Am. J. Public Health* 93:1552–1558, 2003.
25. SAELENS, B. E., J. F. SALLIS, and L. D. FRANK. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann. Behav. Med.* 25:80–91, 2003.
26. SALLIS, J. F., M. F. JOHNSON, K. J. CALFAS, S. CAPAROSA, and J. NICHOLS. Assessing perceived physical environment variables that may influence physical activity. *Res. Q. Exerc. Sport* 68: 345–351, 1997.
27. SALLIS, J. F., and N. OWEN. Ecological models of health behavior. In: *Health Behavior and Health Education: Theory, Research, and Practice*, K. Glanz, F. M. Lewis, and B. K. Rimer. San Francisco, CA: Jossey-Bass, pp. 403–424, 2002.
28. SCHILLING, J., and L. S. LINTON. The public health roots of zoning: in search of active living's legal genealogy. *Am. J. Prev. Med.* 28:96–104, 2005.
29. SNIJDER, T. A. B., and R. BOSKER. Modeled variance in two level models. *Sociol. Method Res.* 22:342–363, 1994.
30. TRANSPORTATION RESEARCH BOARD AND INSTITUTE OF MEDICINE. Does the Built Environment Influence Physical Activity? Examining the Evidence. Special Report 282. Washington, DC: National Academies Press, pp. 1–248, 2005.
31. VAN DE VIJVER, F. J. R., and H. POORTINGA. Structural equivalence in multilevel research. *J. Cross-Cultural Psychol.* 33:141–156, 2002.
32. VANDELANOTTE, C., I. DE BOURDEAUDHUIJ, R. PHILIPPAERTS, M. SJÖSTRÖM, and J. SALLIS. Reliability and validity of a computerized and Dutch version of the International Physical Activity Questionnaire (IPAQ). *J. Phys. Activity Health* 2:63–75, 2005.
33. WHITE, H. Maximum likelihood estimation of misspecified models. *Econometrica* 50:1–25, 1982.