



Practice of Epidemiology

Construct Validation of 4 Food-Environment Assessment Methods: Adapting a Multitrait-Multimethod Matrix Approach for Environmental Measures

Leia M. Minaker*, Kim D. Raine, T. Cameron Wild, Candace I. J. Nykiforuk, Mary E. Thompson, and Lawrence D. Frank

* Correspondence to Dr. Leia M. Minaker, School of Public Health and Health Systems, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada (e-mail: lminaker@uwaterloo.ca).

Initially submitted May 15, 2013; accepted for publication October 15, 2013.

Few studies have assessed the construct validity of measures of neighborhood food environment, which remains a major challenge in accurately assessing food access. In this study, we adapted a psychometric tool to examine the construct validity of 4 such measures for 3 constructs. We used 4 food-environment measures to collect objective data from 422 Ontario, Canada, food stores in 2010. Residents' perceptions of their neighborhood food environment were collected from 2,397 households between 2009 and 2010. Objective and perceptual data were aggregated within buffer zones around respondents' homes (at 250 m, 500 m, 1,000 m, and 1,500 m). We constructed multitrait-multimethod matrices for each scale to examine construct validity for the constructs of food availability, food quality, and food affordability. Convergent validity between objective measures decreased with increasing geographic scale. Convergent validity between objective and subjective measures increased with increasing geographic scale. High discriminant validity coefficients existed between food availability and food quality, indicating that these two constructs may not be distinct in this setting. We conclude that the construct validity of food environment measures varies over geographic scales, which has implications for research, policy, and practice.

food availability; food environment; food supply; nutrition; nutrition policy; psychometrics; residence characteristics

Abbreviations: AUC, area under the curve; ICC, intraclass correlation coefficient; MTMM, multitrait-multimethod; NEM-S, Nutrition Environment Measures Survey in Stores; NEWPATH, Neighborhood Environments in Waterloo Region: Patterns of Transportation and Health; RFEI, Retail Food Environment Index; ROC, receiver operating characteristic.

Recognition of environmental determinants of food-related behavior has gained momentum over the past 20 years (1). Relative to individual-level measures like diet records, environmental assessments such as neighborhood availability of fruits and vegetables might have potential for less bias and higher cost-effectiveness (2) and may also have a sustained impact on health outcomes in comparison with programmatic strategies.

Neighborhood food environments consist of several constructs: food access, food availability, food quality, and food affordability (3). The extant research has examined associations between at least 1 construct and at least 1 health outcome, or associations with area-level disadvantage (e.g., food desert research) (4, 5). The concept of geographic food access refers to local availability of food outlets and is

often used as a proxy for the underlying construct of food availability—the kinds of foods that are available where people shop. Food quality (the quality of available foods) and affordability (aggregated food costs in a given area) have also been studied (6). With geographic measures of food access, researchers categorize food outlets as “healthy” (e.g., grocery stores) or “unhealthy” (e.g., convenience stores) based on the differential distribution of nutritious foods at different outlets (7). The mere presence of a given type of food outlet (the food access measure) does not fully capture people's access to food, since other individual and store-level factors (e.g., income, mobility, aesthetics, safety, and service) also contribute to people's use of food outlets (8). Despite this, food environments are an important public-health setting given their role in driving obesity rates (9).

Over 500 food-environment assessment methods exist (10), and few studies have compared different measures (11–15). Researchers should explicitly employ hypothesized causal models to link environmental features with diet-related diseases, and measures should accurately reflect the environmental constructs (16). The field of food-environment assessment is characterized by inconsistent operational definitions (17–19) and inconsistent geographic scales on which environmental exposures are hypothesized to operate (17). Equivocal findings may hamper attempts to implement programs and policies (e.g., zoning restrictions).

The measurement imperative

In 2009, only 13.1% of existing food-environment measures had been tested for any psychometric properties (20), and recent reviews recommended that assessment tools be psychometrically evaluated (4, 18, 21–24). Psychometric evaluation addresses overarching theories and techniques of measurement and is generally conducted at an individual level (16). Construct validity, an important consideration in psychometrics, is comprised of 2 main components: *operational* definitions (how are constructs measured?) and *syntactical* definitions (how are constructs related to one another within a theoretical system?) (25). In terms of *operational* definitions, specific questions relate to: the reliability of measures (e.g., test-retest reliability, interrater reliability, and internal consistency); convergent validity (measures assessing the same construct should be highly correlated); and discriminant validity (measures assessing theoretically unrelated constructs should not be highly correlated) (26). Ecological assessment methods have lagged behind psychometric approaches to measurement (27), although several studies have examined construct validity by comparing food-environment measures within the same population (11–15, 28). Results generally indicate that there are modest associations between objective and perceptual measures of food environment and that there is an inconsistent ability to detect associations with outcomes of interest (11, 13, 15, 28).

Only 2 studies used more than 1 objectively defined measure in their analyses (11, 28). Thus, the question of whether different objective measures of food-environment variables are truly measuring the constructs they purport to measure remains. Because measures differ in terms of resources needed to implement them, there is a need among both researchers and practitioners to know how closely a food environment characterized by means of one method corresponds to its characterization by another method. Examining several methods simultaneously will clarify how closely different methods are related to each other, particularly given that measures purportedly assessing underlying constructs of “healthy food availability” or “healthy food quality” are so diverse.

A multitrait-multimethod (MTMM) matrix is a psychometric approach assessing the *operational* component of construct validity. MTMM matrices assess “the adequacy of tests as measures of a construct, rather than the adequacy of a construct as determined by the confirmation of theoretically predicted associations with measures of other constructs” (26, p. 100). In this procedure, a set of c constructs is measured by means of m methods (26). The resulting MTMM cor-

relation matrix represents cm measures. Construct validity is demonstrated through high reliability, high convergent validity between measures purportedly assessing the same constructs, and low discriminant validity between measures assessing different constructs.

Objectives

The objectives of the present study were to 1) use MTMM matrices to systematically examine the operational component of the construct validity of 4 different food-environment measures purportedly assessing 3 different constructs: food availability, food affordability, and food quality, and 2) examine how geographic scale influences construct validity. Each of the tools described in this article has been explained in the academic literature and attempts to capture the concept of “food availability,” “food affordability,” and/or “food quality”; we selected 4 popular measures to reflect variation in implementation costs. Therefore, we examined convergent and discriminant validity among published tools that purportedly measure different aspects of these underlying constructs.

METHODS

Sampling

This study was undertaken in collaboration with the Neighborhood Environments in Waterloo Region: Patterns of Transportation and Health (NEWPATH) Study in Ontario, Canada, which examined the relationship between urban built environments and various health outcomes. The sample was stratified by neighborhood walkability, household income, and household size, with allocation to achieve high statistical power to detect hypothesized effects of walkability. Within each level of walkability, we used proportional allocation to recruit a stratified random sample of households that were representative of the study area in terms of income and household size according to 2006 census data. Response rates conditional on initial agreement varied between 56% and 64% over 6 phases of data collection. The units of analysis were households; respondents were representatives of those households (the male or female household head). Household- and individual-level survey weights were calibrated to ensure that households were geographically representative of their city and walkability level. Participating households completed a paper questionnaire that included perceptions of food environment (see Table 1).

The home addresses of NEWPATH participants with complete survey data in 3 urban centers (Kitchener, Cambridge, and Waterloo) in the Region of Waterloo ($n = 2,397$) were geocoded and represent the centroids from which buffer zones were established. Because of the lack of consensus around which geographic scale is most appropriate for examining food environments, ArcGIS software (Esri, Redlands, California) was used to establish 250-m, 500-m, 1,000-m, and 1,500-m Euclidean distance buffer zones around respondents' homes. Preliminary analyses revealed that 82% of the sample lived within 1,500 m of the nearest supermarket (mean distance = 1,022 m); buffer zones for objective food-environment assessment data were created to be consistent

Table 1. Characteristics of Food-Environment Measures Included in an Evaluation of the Construct Validity of Food-Environment Measures, Region of Waterloo, Ontario, Canada, 2010

Instrument	Type of Food Outlet Assessed	Food Construct(s) Addressed	Methodology	Psychometric Tests Conducted Previously	First Author, Year (Reference No.)
Shelf-space measures	Stores	Availability	Cumulative linear shelf space of fruits and vegetables (fresh, frozen, and canned varieties).	Interrater reliability; face and construct validity	Spence, 2009 (29)
NEMS-S checklist	Stores	Availability Affordability Quality	Objective audits of food stores, including fruits and vegetables (fresh, frozen, and canned varieties), grain products (breads and cereals), milk products, meat products (ground beef, hot dogs), frozen dinners, and other foods (including soda, chips, and baked goods).	Interrater and test-retest reliability; face and construct validity	Glanz, 2007 (6)
Retail Food Environment Index	Stores and restaurants	Availability	Geographic analysis of ratio of the number of fast-food outlets and convenience stores to the number of grocery stores and specialty stores.	None	Rose, 2009 (30)
Food-environment perceptions	Stores and restaurants	Availability Affordability Quality	<p>Neighborhoods were defined as the area "within a 10- to 15-minute walk of home (1–1.5 km)." Perceptions of the food environment were assessed using agreement with the following statements on a 4-point Likert scale.</p> <p><i>Availability:</i> "There is a large selection of fresh fruits and vegetables available in my neighborhood"; "There is a large selection of low-fat products available in my neighborhood"; "It is easy to eat healthily at the restaurants in my neighborhood."</p> <p><i>Affordability:</i> "I shop elsewhere because the prices in my neighborhood are too high"; "The produce in my neighborhood is more expensive than that in other neighborhoods"; "The low-fat products in my neighborhood are more expensive than those in other areas."</p> <p><i>Quality:</i> "The fresh produce in my neighborhood is of high quality"; "The low-fat products in my neighborhood are of high quality."</p>	Inter-item and test-retest reliability	Moore, 2008 (13); Mujahid, 2007 (27)

Abbreviation: NEMS-S, Nutrition Environment Measures Survey in Stores.

with perceptual data (Table 1). Previous research using buffer zones has tested different geographic scales, and significant findings have been reported at some scales but not others; therefore, all 4 buffer zone sizes were examined (29, 30). Moreover, we employed concentric, ego-centered buffers to maximize convergence with the perceived measures (Table 1). Data presented here represent correlations between constructs within each buffer zone, rather than between

buffer zones. Given that each buffer zone contained data from the smaller buffer zones nested within it, the point of the analysis was to determine the scale at which geographic measures showed the highest convergent and discriminant validity and reliability.

The NEWPATH Study received ethics clearance from the University of Waterloo's Office of Research Ethics and the University of British Columbia's Behavioral Research Ethics Board.

Assessment of food environment

We applied 4 food-environment assessment methods in the Waterloo region: 3 objective measures (a Canadian adaptation of the Nutrition Environment Measures Survey in Stores (NEMS-S) (6; Susan Buhler, Alberta Health Services (Edmonton, Alberta, Canada), personal communication, 2010), a measure of shelf space (31), and the Retail Food Environment Index (RFEI) (29)) and 1 subjective measure (NEUPATH participants' perceptions) (Table 1). Two objective methods (shelf space and NEMS-S) were used to assess 1 of each chain convenience store, pharmacy, and superstore in the 3 cities (since chains strive to be identical in terms of advertising, promotions, price, and available foods) and each grocery store and independently owned convenience store, pharmacy, and specialty store in the 3 cities. Because grocery stores in the Waterloo region differ in size (and may offer different products), all grocery stores were assessed. Names and locations of food outlets were identified using the Region of Waterloo's public health inspection database, which includes geocoded data for each food outlet in the region inspected by public health inspectors, and by systematic direct observation to ensure maximum coverage (32). Outlets identified in the public health inspection database were coded using North American Industry Classification System codes (33) for the RFEI calculation (Table 1). Objective data were collected between May and August of 2010; NEUPATH data were collected from 2009 to 2010.

Training of raters

Six observers with at least 2 years of university education collected food-environment data. The first author (L.M.M.) trained raters to use the NEMS-S (34) and shelf-space measures (30, 31). Training included classroom sessions and fieldwork, with feedback on results. It took approximately 1 week for raters to consistently achieve correct answers on all of the food-environment measures. During debriefing sessions, consensus on appropriate data was reached. Decision rules were added to protocol handbooks, which were taken to each outlet assessment. Additionally, throughout data collection, different raters periodically assessed the same outlet to provide immediate feedback and limit the occurrence of drift in observations. The first author participated in all data collection.

Statistical analyses

Two sets of MTMM matrices were constructed for each of the 4 buffer zone sizes (250 m, 500 m, 1,000 m, and 1,500 m). One set of matrices included all of the data needed to assess construct validity in buffer zones where food outlets may or may not have been present. The second set included only buffer zones where food outlets existed, to answer the question, "In buffer zones containing food outlets, what is the convergent and discriminant validity between and reliability within measures?" Buffer zone databases contained all method-construct variables (e.g., NEMS-S food availability, shelf-space food availability).

As mentioned above, the operational component of construct validity is concerned with the measures' reliability in addition to convergent and discriminant validity. In the current study, reliability tests differed for each tool employed. For perceptions, internal consistency was calculated using Cronbach's α , since residents completed only 1 survey. Because of resource constraints, only 1 NEMS-S survey was completed in each store; therefore, Cronbach's α was used to calculate the internal consistency of the NEMS-S survey. Two raters collected shelf-space data in each store; thus, interrater reliability was calculated. Although the RFEI was originally developed to assess geographic food *access*, this study included RFEI scores as an *availability* variable, because availability is thought to be the underlying mechanism by which access affects outcomes. Moreover, the RFEI represents the least resource-intensive method employed here, and, importantly, a modified version of the RFEI is being used by the US Centers for Disease Control and Prevention to produce census-tract-level state maps of food environments (35). The RFEI measure did not lend itself to the calculation of reliability coefficients because it is a 1-item measure, does not change over a short period of time, and is rated by only 1 rater. Spearman correlations were calculated rather than Pearson correlations because metrics were different for each assessment tool. Correlations were calculated using SPSS Statistics 19 (IBM, Armonk, New York).

The NEMS-S has been classified as a "proxy gold standard" (11) and has been used in many published food-environment research papers (28, 36–40). To further assess the validity of the other food-environment measures as compared with the NEMS-S, we plotted a receiver operating characteristic (ROC) curve (11). The ROC curve (using 1,000-m buffer-zone data) was created by estimating sensitivity and specificity for the measures using threshold values based on distributional deciles.

RESULTS

Databases contained data from 422 food stores, representing all food stores that were open for business in the 3 Waterloo cities at the time of data collection. Assessments were carried out on 47 grocery stores, 47 specialty stores (including ethnic stores, representing 51 unique locations), 9 pharmacies (representing 22 unique locations), 3 superstores (representing 13 unique locations), and 169 convenience stores (representing 289 unique locations). Using the Waterloo public health inspection database to identify food outlets, we found that 11% of grocery stores and 16% of convenience stores either were not relevant to the NEMS-S or shelf-space measures (i.e., they did not sell items rated by either tool) or were not located at the address given (see Table 2). In total, we carried out assessments on 68% of stores that were at the location listed and were open during business hours; the highest refusal rate was among convenience stores (39% of convenience stores refused to allow raters to complete assessments). A probability-based technique was used to randomly assign NEMS-S and shelf-space scores from the observed stores to missing data from similar types of stores ($n = 131$; 31% of store locations), consistent with previous food-environment research (30). Preliminary results revealed that

Table 2. Distribution of Local Food Outlets According to Measure of Neighborhood Food Environment, Region of Waterloo, Ontario, Canada, 2009–2010

Outlet Type	No. of Outlets Assessed	No. of Unique Locations Included in Buffer-Zone Analyses (Including All Chains)	No. of Outlets Found in PHID That Were Not at the Given Address or Were Closed	No. of Outlets Found in PHID That Were Irrelevant ^a	No. Found Through Direct Observation	Refusal Rate, ^b %
Grocery store	47	47	0	5	1	0
Convenience store	169	289	10	17	4	39
Pharmacy	9	22	0	0	0	0
Superstore	3	13	0	0	0	0
Specialty store	47	51	1	3	2	8

Abbreviations: NEMS-S, Nutrition Environment Measures Survey in Stores; PHID, public health inspection database.

^a Outlets selling no food or beverages or selling only a limited number of canned goods that were not rated by any tool.

^b Store employees instructed raters not to conduct shelf-space or NEMS-S checklist measurements.

refusals were not geographically clustered; therefore, while the imputation of missing data may have decreased the accuracy of the data, it is not expected to have created a substantial bias given that measures were aggregated geographically. Reasons for missing data included the food outlet's being closed for renovations or owner or manager refusal.

Each tool had generally good reliability, with the exception of internal consistency for NEMS-S items measuring food affordability ($\alpha = -0.173$) (see Table 3). Internal consistency for NEMS-S food availability and food quality was good and excellent ($\alpha = 0.861$ and $\alpha = 0.925$), respectively. Interrater reliability for the shelf-space method ranged from good (intraclass correlation coefficient (ICC) = 0.858 for canned fruit) to excellent (ICC = 0.996 for fresh vegetables), with excellent mean reliability (ICC = 0.940) across the 11 items assessed. Internal consistency for perceptual items was acceptable to good ($\alpha = 0.746$ for availability, $\alpha = 0.833$ for quality, and $\alpha = 0.772$ for affordability). Table 3 shows an MTMM matrix using 1,000-m buffer-zone data. In this matrix, which restricts data to buffer zones that contain at least 1 food store, convergent validity coefficients were lower, indicating that measures were better at discriminating whether food stores existed in that zone than at exhibiting good convergent validity in zones where food stores existed. Results from MTMM matrices restricted to buffer zones in which food stores existed were similar using 250-m, 500-m, and 1,500-m data.

Convergent and discriminant validity between objective measures generally decreased with increasing geographic size, whereas correlations between objective and perceived measures increased or remained stable with increasing geographic size (see Table 4).

For both sets of matrices, there were high discriminant validity coefficients between food availability and food quality for the NEMS-S (at 250 m, $\rho = 0.785$ for the first set and $\rho = 0.577$ for the second set) and perceived measures ($\rho = 0.830$). Conversely, there were low discriminant validity coefficients between food affordability and the other constructs. Convergent validity coefficients between objective and perceived measures were very low in both sets of matrices.

In the ROC curve (see Figure 1), the area under the curve (AUC) represents the probability that a higher score for shelf

space, RFEI, or perceived availability for a randomly chosen high NEMS-S availability score will exceed the result for a randomly chosen negative case. The AUC was highest for shelf space (AUC = 0.666, $P < 0.001$), indicating that shelf space was the most accurate predictor of "good food availability" as characterized by the NEMS-S, despite its being "fair" by conventional interpretations.

DISCUSSION

This study focused on the *operational definition* component of construct validation of food-environment measures (25). Each measure was reliable, which is consistent with previous reports for the NEMS-S (6) (excepting affordability), shelf space (31), and residents' perceptions (13, 27). However, the second set of matrices showed low convergent validity coefficients at all geographic scales for the multi-method assessment of the availability construct. These results suggest a method effect: The construct measured by these instruments varies substantially by method.

Measures used here were selected to represent the state of the evidence and to vary in terms of resource-intensiveness. All of the measures attempt to define the potentially nebulous concepts associated with a "healthy food environment." Therefore, MTMM matrices were appropriate evaluation tools. Low convergent validity between the RFEI and NEMS-S food availability or shelf space was not surprising, given that the RFEI includes both fast-food outlets and stores, while the NEMS-S and shelf-space measures address only stores. Relatively low convergent validity between shelf space and the NEMS-S might be due to the fact that shelf space did not measure all of the foods assessed by the NEMS-S, which additionally examines low-fat and whole-grain products. Every method used in the current study has shown some correlation with weight status (29, 30, 37, 41) or diet quality (11, 13, 42), despite low convergent validity. Low convergent validity might indicate different pathways through which the actual constructs measured are associated with outcomes. Measures may be tapping into constructs that exhibit different (perhaps even contradictory) effects on

Table 3. Multitrait-Multimethod Matrix for Measures of Neighborhood Food Environment Using 1,000-m Buffer-Zone Data, Region of Waterloo, Ontario, Canada, 2009–2010^a

	NEMS-S			Shelf Space (Food Availability)	RFEI (Food Availability)	Residents' Perceptions		
	Food Availability	Food Quality	Food Affordability			Food Availability	Food Quality	Food Affordability
NEMS-S								
Food availability	0.861 ^b							
Food quality	0.427 ^{c,d}	0.925 ^b						
Food affordability	-0.265 ^c	-0.234 ^c	-0.173 ^b					
Shelf space								
Food availability	0.310 ^{e,f}	-0.008 ^c	0.015 ^c	0.858–0.996 ^b				
RFEI								
Food availability	0.028 ^e	-0.084 ^c	-0.247 ^c	0.192 ^{e,f}				
Residents' perceptions								
Food availability	0.169 ^{e,f}	0.035 ^c	-0.031 ^c	0.262 ^{e,f}	0.035 ^e	0.746 ^b		
Food quality	0.166 ^{e,f}	0.035 ^e	-0.067 ^c	0.210 ^{c,f}	0.034 ^c	0.830 ^{c,f}	0.833 ^b	
Food affordability	-0.014 ^c	-0.110 ^c	0.029 ^e	-0.001 ^c	0.055 ^c	-0.171 ^c	-0.110 ^c	0.772 ^b

Abbreviations: NEMS-S, Nutrition Environment Measures Survey in Stores; RFEI, Retail Food Environment Index.

^a Data were restricted to those buffer zones that contained food stores within 1,000 m of respondents' homes.

^b Reliability coefficient.

^c Discriminant validity coefficient.

^d $P < 0.001$.

^e Convergent validity coefficient.

^f $P < 0.01$.

outcomes or may be tapping into different mechanisms by which outcomes occur.

High discriminant validity between food availability and food quality indicates that these may not have been distinct constructs in this study setting: In this setting, available foods were of high quality. In other contexts, where low-income and/or minority residents are concentrated and poor food quality has been identified (43), assessment of food quality may be an important feature of overall food-environment assessment (6, 44, 45). Methodologically speaking, including variables that are not highly correlated would more meaningfully contribute to predictive models.

Conversely, low discriminant validity coefficients between food affordability and other measures would argue for the inclusion of affordability measures in food-environment research, since the low discriminant validity coefficients indicate that current affordability measures assess a different construct than food availability or food quality. That said, although perceived affordability showed adequate reliability, the NEMS-S affordability component showed poor internal consistency. Improving affordability measures may be an important direction for future research. One potential avenue of exploration is the cost of nutritious food baskets, which have been identified as a meaningful way to capture affordability. Nutritious food baskets represent the quantity and purchase units of a set of foods that constitute a nutritious diet for persons in various age and gender groups. Comparing the cost of a healthy diet (using nutritious food baskets) with governmental social assistance payments contextualizes the cost of a healthy diet in a given geographic area (46, 47).

The second objective of the current study was to examine how geographic scale influences validity coefficients. In general, the convergent validity of objective measures decreased with increasing geographic scale. If two variables are approximately bivariate-normal and are positively associated because they measure the same construct, the typical pattern would be a decrease in correlations as buffer zone size increases, because with larger buffer zones, observations increasingly violate independence assumptions (i.e., more overlap between buffer zones). Here, correlations between perceived and objective measures seemed to increase with increasing geographic scale. This finding may reflect the wording of the survey question: "Please think of your neighborhood as only the area within about a 10- to 15-minute walk (1–1.5 km) from your home"; this potentially indicates that residents may be able to conceptualize a 1-km buffer zone around their homes.

Regarding the ROC analyses, the statistically significant AUCs for shelf space and perceptions indicate that these two measures predict availability, as per the NEMS-S, significantly better than chance. The AUC is a useful 1-statistic summary of the accuracy of the test as a predictor of the gold-standard test, which is helpful, given the 3 different methods that were compared with the NEMS-S. In another ROC curve analysis of perceived availability versus NEMS-S availability, Moore et al. (11) reported that the AUC for perceived availability was 0.658, slightly higher than that in the current study (AUC = 0.580).

Assessment of the measurement properties of environmental characteristics is less developed than is psychometrics

Table 4. Convergent and Discriminant Validity Coefficients for an Evaluation of the Construct Validity of Food-Environment Measures, Region of Waterloo, Ontario, Canada, 2009–2010^a

Method-Trait Variables	Correlation Coefficient			
	250-m Buffer	500-m Buffer	1,000-m Buffer	1,500-m Buffer
Food availability				
NEMS-S and shelf space	0.911	0.849	0.455	0.174
NEMS-S and RFEI	0.851	0.724	0.213	0.069
NEMS-S and perceptions	-0.010	0.036	0.213	0.069
Shelf space and RFEI	0.789	0.798	0.343	-0.019
Shelf space and perceptions	0.014	0.081	0.262	0.238
RFEI and perceptions	0.005	0.020	0.035	0.007
Food quality				
NEMS-S and perceptions	-0.045	-0.032	0.035	0.030
Food affordability				
NEMS-S and perceptions	0.022	0.013	0.029	0.036
NEMS-S (quality) and NEMS-S (availability)	0.577	0.499	0.427	0.548
NEMS-S (quality) and NEMS-S (affordability)	-0.054	-0.105	-0.234	-0.208
NEMS-S (availability) and NEMS-S (affordability)	-0.323	-0.229	-0.265	-0.234
Shelf space and NEMS-S (quality)	0.406	0.290	-0.008	-0.018
Shelf space and NEMS-S (affordability)	-0.015	-0.073	0.015	0.070
RFEI and NEMS-S (quality)	0.120	0.121	-0.084	-0.115
RFEI and NEMS-S (affordability)	-0.132	-0.078	-0.247	-0.291
Perceptions (availability) and NEMS-S (quality)	-0.041	-0.007	0.035	-0.009
Perceptions (availability) and NEMS-S (affordability)	0.029	0.041	-0.031	-0.013
Perceptions (availability) and perceptions (quality)	0.830	0.830	0.830	0.830
Perceptions (availability) and perceptions (affordability)	-0.171	-0.171	-0.171	-0.171
Perceptions (quality) and NEMS-S (availability)	-0.027	0.006	0.166	0.167
Perceptions (quality) and NEMS-S (affordability)	0.033	0.037	-0.067	-0.059
Perceptions (quality) and shelf space	-0.006	0.041	0.210	0.214
Perceptions (quality) and RFEI	-0.022	-0.007	0.034	-0.004
Perceptions (quality) and perceptions (affordability)	-0.110	-0.110	-0.110	-0.110
Perceptions (affordability) and NEMS-S (quality)	0.032	0.046	-0.110	0.000
Perceptions (affordability) and NEMS-S (availability)	0.042	0.053	-0.014	0.011
Perceptions (affordability) and shelf space	0.034	0.032	-0.001	0.015
Perceptions (affordability) and RFEI	0.042	0.094	0.055	0.047

Abbreviations: NEMS-S, Neighborhood Environment Measures Survey in Stores; RFEI, Retail Food Environment Index.

^a Data were restricted to buffer zones where at least 1 food outlet existed (Spearman correlation coefficients).

(16, 27). Lytle defined construct validity as “the extent to which the measure ‘behaves’ in a way consistent with theoretical hypotheses” (16, p. S136), which reflects a *syntactic* definition. In focusing on *operationally* defining food-environment assessment methods, this paper grounds future work in *syntactically* defining constructs by first describing the convergent and discriminant validity of different method-construct combinations.

This study faced several limitations. First, our use of buffer zones around households as the geographic unit meant that independence assumptions were most likely violated, since neighboring, ego-centered buffers may have overlapped.

Second, the study could have been strengthened by including an absolute measure of food affordability (e.g., nutritious food baskets). Third, low convergent validity between the NEMS-S and perceptions of affordability may have been due to the low reliability of NEMS-S affordability. The NEMS-S has been applied in many settings (28, 36–39, 48, 49); therefore, this study informs the future use of this important tool. Third, the use of Euclidean rather than network distance buffer zones may have been a limitation; however, when respondents report on food-environment features within 1–1.5 km of their homes, it is unlikely that they reflect on precise network distances. Additionally, data were

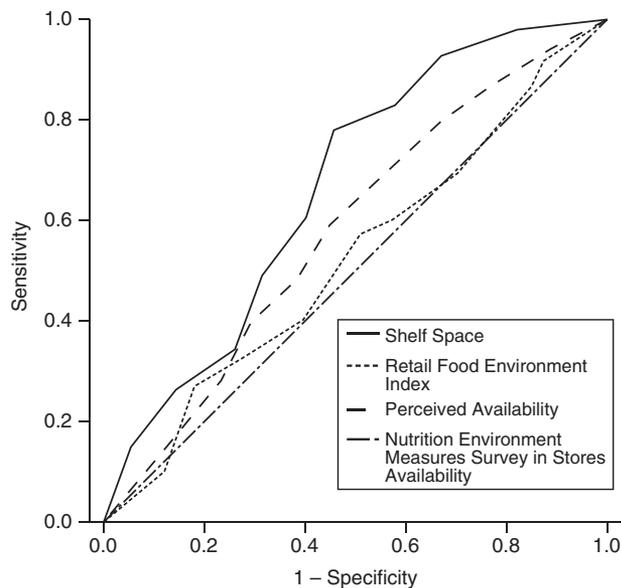


Figure 1. Sensitivity and specificity of food-environment measures versus directly measured food availability (receiver operating characteristic curve), Region of Waterloo, Ontario, Canada, 2009–2010. Directly measured availability (Nutrition Environment Measures Survey in Stores) was dichotomized at the median; curves reflect data discretized at the distributional deciles. Data assessed at the 1,000-m buffer are displayed here.

collected from 3 midsized urban municipalities in southern Ontario, Canada, with a median family income that was Can\$7,000 higher than the national average in 2010 (50); results may not be generalizable to other contexts. Similarly, although households were recruited to be representative in terms of household income and household size to the extent possible, as with most complex surveys, respondents tended to be better educated than nonrespondents. In this analysis, however, the same education bias applied across all geographic areas and thus would not have influenced the correlations described here. Finally, we sampled only 1 chain convenience store, 1 chain pharmacy, and 1 chain superstore because of resource constraints; in future research, investigators could examine the extent to which food-environment constructs vary within chain stores in different areas. Future researchers should carefully consider which food-environment assessment methods and geographic scales are most appropriate.

The current study does not support the argument that environmental assessments include less bias than individual data (2); instead, the evidence indicates that the selection of food-environment method may indeed influence how food environments are characterized. While popular methods may adequately distinguish between the presence and absence of food outlets, they show less agreement in areas where food stores exist. Finally, in terms of relevance for policy-makers, this study shows that food-environment assessment methods are not interchangeable. In the context of restricted resources in local public health departments, if measures were highly correlated, a recommendation could

be made to use the least expensive method, since the measures assess the same construct. On the basis of these findings, this recommendation does not hold. Future work examining *syntactic* definitions will help to elucidate whether less expensive measures predict outcomes as well as (or better than) more comprehensive measures (51).

ACKNOWLEDGMENTS

Author affiliations: School of Public Health and Health Systems, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, Ontario, Canada (Leia M. Minaker); Centre for Health Promotion Studies, Edmonton Clinic Health Academy, Edmonton, Alberta, Canada (Kim D. Raine, Candace I. J. Nykiforuk); Addiction and Mental Health Research Laboratory, Edmonton Clinic Health Academy, Edmonton, Alberta, Canada (T. Cameron Wild); Department of Statistics and Actuarial Science, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada (Mary E. Thompson); and School of Population and Public Health, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada (Lawrence D. Frank).

Funding for the NEWPATH projects was obtained from the Heart and Stroke Foundation, the Canadian Institutes of Health Research (CIHR), and the Region of Waterloo, Ontario, Canada. L.M.M. received a Canada Graduate Scholarship from the CIHR, as well as a training fellowship from Alberta Innovates–Health Solutions (Edmonton, Alberta, Canada) during the writing of this paper. L.M.M. received support through the CIHR Population Interventions for Chronic Disease Prevention training program. K.D.R. received salary support as a CIHR/Heart and Stroke Foundation Applied Public Health Chair. T.C.W. received salary support as an Alberta Innovates–Health Solutions Health Scholar.

Conflict of interest: none declared.

REFERENCES

1. Richard L, Gauvin L, Raine K. Ecological models revisited: their uses and evolution in health promotion over two decades. *Annu Rev Public Health*. 2011;32:307–326.
2. Lytle LA, Fulkerson JA. Assessing the dietary environment: examples from school-based nutrition interventions. *Public Health Nutr*. 2002;5(6A):893–899.
3. Zenk SN, Schulz AJ, Odoms-Young AM. How neighborhood environments contribute to obesity. *Am J Nurs*. 2009;109(7):61–64.
4. Fleischhacker SE, Evenson KR, Rodriguez A, et al. A systematic review of fast food access studies. *Obes Rev*. 2011;12(5):e460–e471.
5. Giskes K, van Lenthe F, Avendano-Pabon M, et al. A systematic review of environmental factors and obesogenic dietary intakes among adults: are we getting closer to understanding obesogenic environments? *Obes Rev*. 2011;12(5):e95–e106.
6. Glanz K, Sallis JF, Saelens BE, et al. Nutrition Environment Measures Survey in Stores (NEMS-S): development and evaluation. *Am J Prev Med*. 2007;32(4):282–289.
7. Kerr J, Frank L, Sallis JF, et al. Predictors of trips to food destinations. *Int J Behav Nutr Phys Act*. 2012;9:58.

8. Drewnowski A, Darmon N. Food choices and diet costs: an economic analysis. *Br J Nutr.* 2005;135(4):900–904.
9. Swinburn BA, Sacks G, Hall KD, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet.* 2011;378(9793):804–814.
10. Applied Research Program, Division of Cancer Central and Population Sciences, National Cancer Institute. *Measures of the Food Environment.* Bethesda, MD: National Cancer Institute; 2013. (<http://appliedresearch.cancer.gov/mfe/>). (Accessed October 11, 2013).
11. Moore LV, Diez-Roux A, Franco M. Measuring availability of healthy foods: agreement between directly measured and self-reported data. *Am J Epidemiol.* 2012;175(10):1037–1044.
12. Moore LV, Roux AV, Brines S. Comparing perception-based and geographic information system (GIS)-based characterizations of the local food environment. *J Urban Health.* 2008;85(2):206–216.
13. Moore LV, Roux AV, Nettleton JA, et al. Associations of the local food environment with diet quality—a comparison of assessments based on surveys and geographic information systems. *Am J Epidemiol.* 2008;167(8):917–924.
14. Freedman DA, Bell BA. Access to healthful foods among an urban food insecure population: perceptions versus reality. *J Urban Health.* 2009;86(6):825–838.
15. Giskes K, van Lenthe F, Brug J, et al. Socioeconomic inequalities in food purchasing: the contribution of respondent-perceived and actual (objectively measured) price and availability of foods. *Prev Med.* 2007;45(1):41–48.
16. Lytle LA. Measuring the food environment: state of the science. *Am J Prev Med.* 2009;36(4 suppl):S134–S144.
17. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. *Obes Rev.* 2011;12(3):217–230.
18. Feng J, Glass TA, Curriero FC, et al. The built environment and obesity: a systematic review of the epidemiologic evidence. *Health Place.* 2010;16(2):175–190.
19. Caspi CE, Sorensen G, Subramanian SV, et al. The local food environment and diet: a systematic review. *Health Place.* 2012;18(5):1172–1187.
20. McKinnon RA, Reedy J, Morrisette MA, et al. Measures of the food environment: a compilation of the literature, 1990–2007. *Am J Prev Med.* 2009;36(4 suppl):S124–S133.
21. Forsyth A, Lytle LA, Van Riper D. Finding food: issues and challenges in using geographic information systems to measure food access. *J Transp Land Use.* 2010;3(1):43–65.
22. Ohri-Vachaspati P, Leviton LC. Measuring food environments: a guide to available instruments. *Am J Health Promot.* 2010;24(6):410–426.
23. Kelly B, Flood VM, Yeatman H. Measuring local food environments: an overview of available methods and measures. *Health Place.* 2011;17(6):1284–1293.
24. Larson NI, Story MT, Nelson MC. Neighborhood environments: disparities in access to healthy foods in the U.S. *Am J Prev Med.* 2009;36(1):74–81.
25. Lord FM, Novick MR. *Statistical Theories of Mental Test Scores.* Reading, MA: Addison-Wesley Publishing Company; 1968.
26. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull.* 1959;56(2):81–105.
27. Mujahid MS, Roux AV, Morenoff JD, et al. Assessing the measurement properties of neighborhood scales: from psychometrics to econometrics. *Am J Epidemiol.* 2007;165(8):858–867.
28. Gustafson AA, Sharkey J, Samuel-Hodge CD, et al. Perceived and objective measures of the food store environment and the association with weight and diet among low-income women in North Carolina. *Public Health Nutr.* 2011;14(6):1032–1038.
29. Spence JC, Cutumisu N, Edwards J, et al. Relation between local food environments and obesity among adults. *BMC Public Health.* 2009;9:192.
30. Rose D, Hutchinson PL, Bodor JN, et al. Neighborhood food environments and body mass index: the importance of in-store contents. *Am J Prev Med.* 2009;37(3):214–219.
31. Farley TA, Rice J, Bodor JN, et al. Measuring the food environment: shelf space of fruits, vegetables, and snack foods in stores. *J Urban Health.* 2009;86(5):672–682.
32. Hosler AS, Dharssi A. Identifying retail food stores to evaluate the food environment. *Am J Prev Med.* 2010;39(1):41–44.
33. Statistics Canada. *North American Industry Classification System (NAICS)—Canada 2007.* (Catalogue no. 12-501-XIE). Ottawa, Ontario, Canada: Statistics Canada; 2010.
34. Glanz K, Clawson M, Davis E, et al. *Nutrition Environment Measures Training Manual.* Atlanta, GA: Emory University; 2008.
35. Centers for Disease Control and Prevention. *Census Tract Level State Maps of the Modified Retail Food Environment Index (mRFEI).* Atlanta, GA: Centers for Disease Control and Prevention; 2013. (ftp://ftp.cdc.gov/pub/Publications/dnpao/census-tract-level-state-maps-mrfei_TAG508.pdf). (Accessed October 13, 2013).
36. Andreyeva T, Middleton AE, Long MW, et al. Food retailer practices, attitudes and beliefs about the supply of healthy foods. *Public Health Nutr.* 2011;14(6):1024–1031.
37. Cerin E, Frank LD, Sallis JF, et al. From neighborhood design and food options to residents' weight status. *Appetite.* 2011;56(3):693–703.
38. Bertoni AG, Foy CG, Hunter JC, et al. A multilevel assessment of barriers to adoption of Dietary Approaches to Stop Hypertension (DASH) among African Americans of low socioeconomic status. *J Health Care Poor Underserved.* 2011;22(4):1205–1220.
39. Eisenmann JC, Alaimo K, Pfeiffer K, et al. Project FIT: rationale, design and baseline characteristics of a school- and community-based intervention to address physical activity and healthy eating among low-income elementary school children. *BMC Public Health.* 2011;11:607.
40. Hillier A, Cannuscio C, Karpyn A, et al. How far do low-income parents travel to shop for food? Empirical evidence from two urban neighborhoods. *Urban Geogr.* 2011;32(5):712–729.
41. Casagrande SS, Franco M, Gittelsohn J, et al. Healthy food availability and the association with BMI in Baltimore, Maryland. *Public Health Nutr.* 2011;14(6):1001–1007.
42. Franco M, Diez-Roux AV, Nettleton JA, et al. Availability of healthy foods and dietary patterns: the Multi-Ethnic Study of Atherosclerosis. *Am J Clin Nutr.* 2009;89(3):897–904.
43. Zenk SN, Odoms-Young AM, Dallas C, et al. “You have to hunt for the fruits, the vegetables”: environmental barriers and adaptive strategies to acquire food in a low-income African American neighborhood. *Health Educ Behav.* 2011;38(3):282–292.

44. Sloane DC, Diamant AL, Lewis LB, et al. Improving the nutritional resource environment for healthy living through community-based participatory research. *J Gen Intern Med.* 2003;18(7):568–575.
45. Cummins S, Smith DM, Taylor M, et al. Variations in fresh fruit and vegetable quality by store type, urban-rural setting and neighborhood deprivation in Scotland. *Public Health Nutr.* 2009;12(11):2044–2050.
46. Wong KC, Coveney J, Ward P, et al. Availability, affordability, and quality of a healthy food basket in Adelaide, South Australia. *Nutr Diet.* 2011;68(1):8–14.
47. Health Canada. *National Nutritious Food Basket.* Ottawa, Ontario, Canada: Health Canada; 2009. (<http://hc-sc.gc.ca/fn-an/surveill/basket-panier/index-eng.php>). (Accessed October 13, 2013).
48. Hubley TA. Assessing the proximity of healthy food options and food deserts in a rural area in Maine. *Appl Geogr.* 2011; 31(4):1224–1231.
49. Hillier A, Cannuscio CC, Karpyn A, et al. How far do low-income parents travel to shop for food? Empirical evidence from two urban neighborhoods. *Urban Geogr.* 2011;32(5): 712–729.
50. Statistics Canada. *Median Total Income, by Family Type, by Census Metropolitan Area.* Ottawa, Ontario, Canada: Statistics Canada; 2012. (<http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil107a-eng.htm>). (Accessed October 11, 2013).
51. Minaker LM, Raine KD, Wild TC, et al. Objective food environments and health outcomes. *Am J Prev Med.* 2013; 45(3):289–296.