March, 2012

Adolescent Physical Activity and the Built Environment: A Latent Class Analysis Approach

Kelsey McDonald
Mary Hearst
Kian Farbakhsh
Carrie Patnode
Ann Forsyth, et al.

Available at: http://works.bepress.com/john_sirard/7/
Adolescent Physical Activity and the Built Environment: A Latent Class Analysis Approach

Kelsey McDonalda, Mary Hearsta, Kian Farbakhsha, Carrie Patnodeb, Ann Forsythc, John Sirardb and Leslie Lytlea

Carrie Patnode: Carrie.D.Patnode@kpchr.org; Ann Forsyth: forsyth@cornell.edu; John Sirard: jrs2wg@virginia.edu
aDivision of Epidemiology and Community Health, School of Public Health, University of Minnesota, 1300 South Second Street, Suite 300, Minneapolis, MN, 55454, USA
bKaiser Permanente Center for Health Research, Portland, OR, USA
cCity and Regional Planning, Cornell University, 106 West Sibley Hall, Ithaca, NY 14853, USA
dKinesiology Program, Curry School of Education, University of Virginia, Memorial Gymnasium, Room 223A, Charlottesville, VA 22904, USA

Abstract

This study used latent class analysis to classify adolescent home neighborhoods (n=344) according to built environment characteristics, and tested how adolescent physical activity, sedentary behavior, and screen time differ by neighborhood type/class. Four distinct neighborhood classes emerged: 1) low-density retail/transit, low walkability index (WI), further from recreation; 2) high-density retail/transit, high WI, closer to recreation; 3) moderate-high-density retail/transit, moderate WI, further from recreation; and 4) moderate-low-density retail/transit, low WI, closer to recreation. We found no difference in adolescent activity by neighborhood class. These results highlight the difficulty of disentangling the potential effects of the built environment on adolescent physical activity.

Keywords

Latent class analysis; neighborhoods; physical activity; adolescents

INTRODUCTION

Low levels of physical activity (PA) and high levels of sedentary time (particularly during screen time) in adolescents are of critical concern in public health. Regular PA has long been associated with a decreased risk of many diseases, such as coronary heart disease, colon cancer, and type II diabetes (U. S. Department of Health Human Services, 1996). Yet, in a recent study, only 8% of adolescents aged 12–15 years met the recommendation of 60 minutes or more of moderate to vigorous PA on 5 of 7 days (Troiano et al., 2008). As adolescents move into adulthood their PA levels continue to decline (Dowda et al., 2003;
Gordon-Larsen et al., 2004; Trost et al., 2002). Recent findings suggest that adolescent sedentary behavior is independent of PA (Biddle et al., 2004; Taveras et al., 2007). Sedentary behaviors have been associated with risk factors for disease such as obesity (Epstein et al., 1995). A substantial portion of adolescent sedentary behavior is spent using screen media (i.e. TV/video, computer, or video game). Youth aged 8–18 years average over 5.7 hours (342 minutes) of screen time per day (Henry J. Kaiser Family Foundation, 2005). Identifying points for PA and sedentary behavior interventions in adolescents may have a significant effect on PA into adulthood, and thereby decrease risk of disease.

The built environment has been associated with certain types of PA (e.g. travel and recreational-related activity) in adults. The built environment is broadly defined as the physical structure of neighborhoods, or the physical environment. Features of the built environment thought to influence adult PA include greater residential population density and street connectivity, a diverse land-use mix, and availability and/or access to recreational facilities, trails, and parks (Humpel et al., 2002; Saelens et al., 2003; Wendel-Vos et al., 2007). Hypothesized reasons for these associations include increased access to PA facilities and/or destinations that people may travel to using active transportation (such as walking or cycling).

Researchers have only recently begun to examine the associations between built environment features and adolescent PA patterns. This line of research is important given that different factors may influence PA in children as compared to adults (Krizek et al., 2004). Evidence from research on adolescents is somewhat sparse and often contradictory (Davison and Lawson, 2006; Ferreira et al., 2007; Giles-Corti et al., 2009). One recent review failed to find convincing evidence of an association between neighborhood environment factors and PA in adolescents (Ferreira et al., 2007). Two other review articles (Davison and Lawson, 2006; Giles-Corti et al., 2009) identified several physical environment characteristics as positively associated with forms of children’s PA: access and/or proximity to recreational facilities and schools (although these associations were more commonly reported in girls than boys); presence of sidewalks and controlled intersections; access/proximity to destinations and/or public transportation (Davison and Lawson, 2006; Giles-Corti et al., 2009) and higher density and/or mixed-use neighborhoods (Giles-Corti et al., 2009). The number of roads to cross and traffic density/speed, as well as crime and area deprivation, were negatively associated with children’s PA (Davison and Lawson, 2006). A number of studies have shown that proximity to parks and recreational facilities seem to be associated with adolescent PA (Cohen et al., 2006; Gordon-Larsen et al., 2006; Norman et al., 2006; Patnode, 2010).

In addition to studies examining individual neighborhood characteristics, other studies have combined built environment variables into a walkability index (WI) including variables such as land-use mix, retail floor area ratio or retail density, intersection density, and residential density. The hypothesis is that certain built environment characteristics make a neighborhood more walkable and that a more walkable neighborhood will encourage PA. Several studies with adolescents found a positive association between a WI and minutes of moderate to vigorous PA (MVPA) (Kligerman et al., 2007; Patnode, 2010).

Very few studies have examined the association between the built environment and inactivity (i.e. sedentary behavior). Several of the studies have focused on multiple features of the built environment (including crime/safety, parks, gyms, walking and cycling ease or facilities, tidiness, and street access and condition), but they only found associations of sidewalk characteristics and hills with sedentary behavior (Jago et al., 2006a; Jago et al., 2005; Norman et al., 2005). Even less research exists on the potential effect of the built environment on screen time, defined as the time spent watching TV/videos, using the

Health Place. Author manuscript; available in PMC 2013 March 01.
internet/computer, or playing computer games. The limited research on screen time does suggest that it is neither highly nor consistently correlated with PA (Biddle et al., 2004; Eisenmann et al., 2002; Taveras et al., 2007); therefore, screen time may be worthy of consideration as a separate outcome.

A newer approach to studying the effect of the built environment on PA seeks to first classify neighborhoods according to different environmental characteristics and then examine any potential neighborhood effect on PA patterns. These methods may be better at capturing the complexity of neighborhoods than risk factor analysis, where singular characteristics or indices characterize the built environment (Nelson et al., 2006). Several studies have used cluster analysis and/or factor analysis to classify neighborhoods and evaluate the effect of neighborhood class on health and/or PA (Jago et al., 2006a; Jago et al., 2005; Li and Chuang, 2008; Nelson et al., 2006; Zhang et al., 2006). Results from these studies suggest possible associations between neighborhood type and PA; however, these results are also mixed. One study found an association between some neighborhood types and adolescent MVPA (Nelson et al., 2006), while another set of studies found no association between neighborhood type and PA (except for light PA) (Jago et al., 2006a; Jago et al., 2005).

An alternative and novel method of classifying neighborhoods according to a set of environmental characteristics is latent class analysis (LCA), which is related to a specific type of cluster analysis, called multivariate mixture estimation. LCA has typically been used to classify people according to different characteristics, and we know of only one other study using LCA to classify neighborhoods (Weden et al., 2010). That study characterized 6 types of neighborhoods across the entire U.S. to examine neighborhood change from 1990 to 2000.

Our study is novel in that it uses LCA to classify neighborhoods to examine the association between neighborhood class and health-related outcomes. The purpose of this study was two-fold. First, we used LCA to classify neighborhoods based on geographical information system (GIS) built environment data available for each participant within a 1600-meter network buffer or measured as the distance to nearest feature. Second, we used these neighborhood classes to determine how PA, sedentary behavior, and screen time differ by neighborhood class. We hypothesized that we would identify classes of neighborhoods with distinct characteristics and that PA, sedentary time, and screen time patterns of adolescents would differ across the neighborhood ‘types,’ in particular, those with more features that support PA being associated with higher mean levels of moderate-to-vigorous PA and less sedentary and screen time behavior.

### METHODS

#### Participants

The participants (n=344) were from the Transdisciplinary Research on Energetics and Cancer – Identifying Determinants of Eating and Activity (TREC-IDEA) study, a 3-year longitudinal study examining the etiology of childhood obesity by focusing on societal and environmental factors (Lytle, 2009). Adolescents aged 10–16 years (at baseline) were recruited from the Minneapolis-St. Paul metropolitan area using three sources: the existing Minnesota Adolescent Community Cohort Tobacco Study cohort (Widome et al., 2007), the Minnesota Department of Motor Vehicle list, and a convenience sample. Each adolescent was recruited along with one parent or guardian. Exclusion criteria included plans to move from the area within 3 years, being a non-English speaker, having a medical condition affecting growth, or having any other condition affecting diet or activity levels or constraining ability to obtain measurements.
This study used data from the baseline measurement (2006). From the total TREC-IDEA sample of 349 adolescents, we dropped 5 due to incomplete data. Additional details of the recruitment and study design have been detailed previously (Lytle, 2009). The University of Minnesota Institutional Review Board approved this study.

**Measures**

Individual-level survey data collection for adolescents and a parent occurred at the Epidemiology Clinical Research Center (ECRC) of the University of Minnesota and was conducted by trained staff from September 2006-May 2007. At that appointment, trained research staff weighed and measured each adolescent and parent and asked them to complete surveys. The youth participants were fit with an accelerometer and were provided with written and verbal instructions to wear the monitor for the subsequent 7 days and mail it back in an addressed, postage-paid envelope.

**Outcome Variables**—This study examined several PA outcomes, a measure of sedentary time, and screen time. PA and sedentary time were measured using the ActiGraph activity monitor, model 7164 (ActiGraph, LLC, Pensacola, FL). Activity was monitored over seven days using 30-second epochs (data collection intervals). The monitor is an objective measure of PA and has been previously validated for use with children in laboratory and field settings (Eston et al., 1998; Louie, 1999; Trost et al., 1998). ActiGraph data were reduced using a custom software program (Catellier et al., 2005). Daily inclusion criteria were set to identify days and times with acceptable accelerometer data. Missing data within an adolescent’s 7-day record were replaced via imputation based on the Expectation Maximization (EM) algorithm (Catellier et al., 2005). On average, approximately 22 hours of data (about 17.5%) per adolescent were imputed over the 7 days of data collection. Similar physical activity studies imputed on average 12 hours of accelerometer data across 6 days per adolescent girl (Pate et al., 2006), and 37% of data across 7 days per adolescent girl and boy (Kitzman-Ulrich et al., 2010). Summary PA variables were calculated using the Freedson age-specific count cutoffs (Freedson et al., 1997) distinguishing sedentary, light, moderate, and vigorous intensity based on age-adjusted MET values (Harrell et al., 2005; Sirard and Welk, 2009). Mean daily minutes of moderate-to-vigorous PA (MVPA), vigorous PA (VPA), total PA (light, moderate, and vigorous PA), and sedentary time were calculated in addition to total movement, which was a 7-day average count per minute per day.

Screen time was reported by adolescents using a self-administered survey, which assessed typical number of hours of screen time on weekdays and weekends. The question asked, “On a typical weekday, how many hours do you spend doing the following: watching TV, DVDs or videos, playing computer games, or using the internet/computer?” The same question was asked for a typical weekend day. The overall estimate of mean minutes of screen time per day was a weighted average of weekday and weekend screen time reported in a range from “None” to “6+ hours per day” (French et al., 2001; Gortmaker et al., 1999; Utter et al., 2003; Wolf et al., 1994).

**Exposure Variable**—The primary exposure is neighborhood class. This variable was generated with a LCA that used 10 neighborhood characteristics that, based on the literature, are potentially related to PA. In identifying our final set of neighborhood characteristics for the LCA, we also considered the number and distribution of variables compared to the total number of participants to identify a final set of variables that would allow our LCA model to converge. The characteristics were dichotomized using a median split for ease of interpretation. All neighborhood characteristics were defined using Geographic Information Systems (GIS) software (ESRI ArcGIS ArcMap 9, Redlands, CA, 2005). Details on the data sources for each characteristic are available in custom GIS Protocols (Forsyth, 2007).
This analysis included network distances (the distance along the street network) from each adolescent’s home to defined built environment features and the density (calculated by dividing the total number of the specific feature [e.g. transit stops] by the land area, excluding water) of a specific feature within a 1600-m network buffer. Given the age range of the participants, a 1600-m buffer anticipated the greater independent walking and bicycling mobility of adolescents as well as ease of access for parents; adolescents can typically walk 1600 meters within 20 minutes (Dengel et al., 2009). Specifically, features included distance to the nearest trail, recreation center, park and gym; distance to school attended; 1600-m density of transit stops, retail food outlets and busy streets. The recreation center variable was calculated as the distance to nearest recreation center, community center, or school. A walkability index (WI) was calculated by summing the z-scores of residential population density, intersection density, and retail employment density, approximating earlier measures given available data (Frank et al., 2005).

The distance measures assume that the closer these PA amenities and destinations, the more likely an adolescent is to be physically active (Giles-Corti et al., 2009). Likewise, the density measures, including the WI, assume that more amenities in the vicinity increase an adolescent’s use of active transport to travel to these amenities (Forsyth et al., 2008; Forsyth et al., 2007).

Potential Confounders—Additional variables were included as potential confounders: age, gender, race/ethnicity, pubertal status, household education, number of people in house, and free/reduced lunch. Demographic information and pubertal status were assessed using survey-based instruments during the measurement visit. Adolescents were asked to define their race/ethnicity according to seven categories. Because the large majority of adolescents were white (93.6%), race/ethnicity was classified as white or non-white. Pubertal status was assessed by the self-report Pubertal Development Scale (PDS) (Petersen and Crossett, 1988). The PDS is a five question summed score with good internal consistency (alpha = 0.77) and high correlation between the PDS and physician rating (0.61–0.67) (Petersen and Crossett, 1988). Parents completed a self-report survey that elicited information on socioeconomic status including the highest level of education completed by the adults in the household, number of people in the house, and if the family received free or reduced cost lunch at the child’s school. Household education was dichotomized into college degree or higher versus less than a college degree. The adolescent’s school was also recorded, so that analyses could account for clustering of students within schools.

Analysis

Data management and all analyses except the LCA were conducted using v.9.2 of the SAS System for Windows (Cary, NC: SAS Institute Inc.), while v.9.1 was used for the LCA.

LCA was chosen to classify neighborhoods by built environment characteristics in part because we hypothesized that we would identify discrete classes or neighborhoods (vs. classes on a continuum) (Hagenaars and Halman, 1989). Compared to cluster analysis LCA is also computationally superior and allows a test of model fit (Hagenaars and Halman, 1989; Rapkin and Luke, 1993; Weden et al., 2010).

For the LCA, we estimated four latent class models, beginning at two latent class models, with increasing numbers of classes based on maximum likelihood methods using SAS PROC LCA (Lanza et al., 2007). We repeated estimation of each of the 2-5-class models 100 times using different starting values to detect any model identification problems. Model fit was assessed according to the likelihood ratio value ($G^2$) and associated degrees of freedom, AIC and BIC. Specifically, if the $G^2$ estimates were less than the model’s degrees of freedom, the model was identified as having reasonably good fit, with lower values of
AIC and BIC preferred (Lanza et al., 2003). Finally, we also considered conceptual interpretation of the resultant classes when determining the best model. For detailed discussions of LCA methodology, see Lanza et al. (Lanza et al., 2003).

Following the estimation of the appropriate number of latent classes, adolescents were assigned to the class in which they had the highest probability of membership (i.e., maximum-probability assignment rule). We examined differences in the socio-demographic variables (i.e., age, gender, race/ethnicity, pubertal status, household education, number of people in house, and free/reduced lunch) between the classes using chi-square tests.

Lastly, the effect of neighborhood class on PA, sedentary behavior, or screen time was explored using PROC GENMOD. Generalized estimated equation (GEE) models were used to account for potential clustering of observations by school (a violation of the assumption of independence of observations), and thereby provide robust standard errors. Adolescent PA, sedentary behavior, and screen time were the dependent variables in multivariate linear regression models. Independent variables included the class to which adolescents were assigned, and the demographic and socio-economic characteristics listed above.

RESULTS

Descriptive Characteristics of the Sample

There were 344 participants from the TREC-IDEA study included in the analyses with complete data on all variables. The sample was comprised of a well-educated, largely white, middle class sample, with even gender distribution of the adolescents. The mean age of the analytic sample was 15.4 years. On average 2.7 (SD 3.2) students attended each school (n=124 schools; not including 11 students who were home schooled). The adolescents in this sample engaged, on average, in 30 minutes of MVPA and 10 hours of sedentary time per day, which included 5 hours of screen time per day. See Table 1A for additional sample description of key analytic variables. Table 1B presents the distributions of the built environment variables included in the LCA. The distributions show that for each built environment variable the study contained a typically diverse range of values for an urban plus suburban area – see e.g. (Frank et al., 2005; Jago et al., 2006b; McDonald et al., 2011).

Estimation of the Number of Latent Classes

The appropriate number of classes can be determined by comparing the goodness of fit. BIC is typically used mostly since it also rewards more parsimonious models. The model with lowest value of the information criterion, BIC, is the best fitting model (Nylund et al., 2007). To avoid converging on a local solution, the estimation algorithms were run several times with different parameter start values. Based on the model fit indices (see Table 2), a 4-class model seemed most appropriate for the data. The selected model also had a high entropy value of 0.85. Entropy is an index for assessing the precision of assigning latent class membership; higher values indicate greater precision of classification.

Approximately 80% of adolescents had a 90% or higher probability of falling within only one class (91% had a 70% or higher probability). Those adolescents with a probability of less than 0.50 in any one group (n=2) were dropped from the analysis. The average probability of membership was calculated in each class to evaluate the assignment accuracy (see Table 3). All classes had a mean probability of 0.89 or higher, indicating good assignment accuracy and distinct classes.
Latent Class Profiles

Table 4 presents the mean item-response probabilities for each built environment characteristic used in the LCA by latent class. These indicate the probability of reporting a value above the median value within a class. Figure 1 represents the same information, but in a graphical format.

Given distinctive item-response probabilities the four neighborhood classes identified in the latent class analysis are characterized as follows: **Class 1** low-density retail/transit, low WI, further from recreation; **Class 2** high-density retail/transit, high WI, closer to recreation; **Class 3** moderate-high-density retail/transit, moderate WI, further from recreation; and **Class 4** moderate-low-density retail/transit, low WI, closer to recreation.

Class 1 (low-density retail/transit, low WI, further from recreation) was the largest class and accounted for 32.8% of the sample. Neighborhoods in this class were distinguished by the lowest probabilities of a high density of transit stops (item-response probability = 0.00) and retail establishments (0.03) within the 1600-meter buffer. They also have a very low probability of a high WI (0.02). In this class, neighborhoods also had a higher probability of greater distances to recreational opportunities, such as recreation centers (0.90) and gyms (0.87).

Class 2 (high-density retail/transit, high WI, closer to recreation) accounted for 29.2% of the sample. These neighborhoods were distinguished by the highest probabilities of a high density of transit stops (0.90) and retail establishments (0.93), and a high WI score (0.97). Neighborhoods were also more likely to be closer to many recreational opportunities, except trail access (0.50) and parks (0.74).

Class 3 (moderate-high-density retail/transit, moderate WI, further from recreation) contained 26.6% of the sample. Neighborhoods in this class have a moderately high probability of having a higher density of transit stops (0.74) and retail (0.63), and a higher WI score (0.73). Although trail access (0.27) and school attended (0.42) were likely to be closer in this class, other recreational opportunities were more likely to be further away, such as distance to a park (0.84), gym (0.74), or recreation center (0.74).

Class 4 (moderate-low-density retail/transit, low WI, closer to recreation) only accounted for 11.4% of the sample, making this the smallest class. These neighborhoods were distinguished by the very low probability of being walkable according to the WI (0.04), as well as the high probability of having a greater density of busy streets (0.89). Most recreational opportunities were more likely to be closer in these neighborhoods, except for school attended (0.72). These neighborhoods were moderately less likely to have a high density of transit stops (0.30) and retail (0.39).

Comparison of Outcomes by Class

Results from the adjusted multivariate linear regression model using adolescent PA, sedentary time, and screen time as the outcome variables and the four neighborhood classes as the exposure variables are presented in Table 5, both by class and overall. We failed to find evidence of an effect of neighborhood class on adolescent PA, sedentary time, or screen time in either the unadjusted or adjusted analyses.

Independent of the outcomes, household education, gender, and race/ethnicity did not vary across the four neighborhood classes; however, free/reduced lunch did vary ($\chi^2=14.09$, p-value 0.003).
DISCUSSION

Few studies of adolescent PA and the built environment have attempted to compare the effect of distinct classes or types of neighborhoods on PA. To our knowledge none of these studies has used a latent class analysis (LCA) to classify neighborhoods. The purpose of this study was to classify neighborhoods based on built environment characteristics using a novel approach – LCA – and then to determine whether the resulting neighborhood classes were associated with adolescent PA, sedentary time, or screen time. In our sample we identified four classes in the best fitting model, which we labeled as: 1) low-density retail/transit, low walkability index (WI), further from recreation; 2) high-density retail/transit, high WI, closer to recreation; 3) moderate-high-density retail/transit, moderate WI, further from recreation; and 4) moderate-low-density retail/transit, low WI, closer to recreation. We failed to find evidence of an effect of neighborhood class (i.e. the built environment) on adolescent PA and sedentary behaviors.

Our results add to the literature on the association of the built environment with adolescent PA, which to date has documented conflicting results (Davison and Lawson, 2006; Ferreira et al., 2007; Giles-Corti et al., 2009). The use of varied measures, operational definitions of measures, and methods may help explain these conflicting patterns. First, the built environment has been measured using many different characteristics, and these may be single risk factors, indices, or a combination of both. Furthermore, the operational definitions of these characteristics are highly variable (Brownson et al., 2009). For example, walkability has been measured using many different features or indices, ranging from street connectivity to population density to indices with 3+ features. Furthermore, a meaningful WI for adolescents might require a modified set of characteristics compared to that used for adults. However, similar walkability indices have been used with adolescents previously (Kligerman et al., 2007; Patnode, 2010). When classes of neighborhoods are defined, different methods have been used, including principal component analysis and cluster analysis, perhaps also producing differing results.

Another complication is that built environment research lacks an agreed-upon definition of a neighborhood area. Although buffers have become a common means of defining a neighborhood with advances in GIS, the size and type (straight-line vs. network) of the buffer varies and may affect study results. Jago et al (Jago et al., 2006b) found an association between access to parks and light PA levels in adolescent males with a 400-meter straight-line buffer, but not with a 1-mile buffer. Furthermore, the buffer is usually created around the home, but an adolescent may spend only a small proportion of their time in this particular built environment. This study used a 1600-meter (approximately 1-mile) buffer around the adolescents’ homes. It is possible that this buffer was too large to find an association, as is suggested by several recent studies on proximity to parks and PA (Cohen et al., 2006; Jago et al., 2006b). However, as was found by Laska et al. among food purchasing patterns in young adults (Laska et al., 2010), meaningful physical activity space for adolescents is likely outside of traditional GIS buffers. Middle-class adolescents, including our population, may have access to economic resources that promote mobility (e.g. cars), so that those who want to be active can also travel outside of their residential environment to do so, meaning that economic resources and attitudes may overshadow the effects of any particular environment. This is an important issue, in that many interventions are focused on the residential environment which may not be the best scale.

Associations between the built environment and PA may also differ depending upon the type of PA targeted: recreational, travel, or overall (Brownson et al., 2009). Results from the Twin Cities Walking Study suggest a different association between built environment characteristics (block size and density) when comparing type of adult PA (travel walking,
leisure walking and total physical activity) (Oakes et al., 2007). Research on adolescents rarely explicitly addresses the effect of a wide range of built environment characteristics on PA by type of PA; however, some research suggests the potential for different associations (Jago et al., 2006a; Jago et al., 2005).

Studies vary in their use of objective and/or perceived measures of the built environment. Even if an individual has destinations to which they can walk in their neighborhood, they may not accurately perceive those destinations as walkable. It may be that an individual’s perception of their neighborhood is more important in predicting their behavior than the actual amenities and state of the neighborhood. Furthermore, measuring PA objectively versus self-report may produce differing results. One study failed to find an association between neighborhood environment variables and an objective measure of overall PA in girls, but found an association with self-reported frequency of walking (Hume et al., 2007).

Although the current study provides a robust method of classifying neighborhoods according to built environment characteristics, it has several limitations. First, the results may only be generalizable to populations that are largely white, of higher SES, and in the non-rural Midwest. Furthermore, the use of LCA makes it difficult to compare results across studies, because the classes are defined anew in each study.

Other potential limitations related to the use of a LCA include the dichotomous categorization of the built environment characteristics using the median split. If we had used different cutpoints than the median, different classes might have emerged from the LCA. Furthermore, it is possible that the median split failed to create meaningfully distinct categories within each characteristic. We explored using the neighborhood characteristics as continuous variables; however, the results were not easily interpretable. We also found high correlation between the walkability index and two of the ten individual features, which may also have altered the class composition, but was unlikely to have altered the relationship with the outcomes.

Another potential limitation is that adolescents were assigned to a class based on the highest probability of belonging to a class; those with <100% probability did not fit exactly within a class. This may have resulted in some misclassification error; however, almost 80% of the sample had at least a 90% probability of falling within only one class. In addition, our outcome measures do not explicitly exclude time during the school day, or other times when students are away from their neighborhood. Finally, the imputation algorithm for the missing accelerometer data produced an average 18-hour awake time, which may overestimate the true awake times; however, we would expect only a slight, if any, increase in PA as a result.

Like most built environment and PA studies, this study is observational, which makes it harder to disentangle causation than with a natural experiment. With neighborhood studies the particular problem with the observational study design is structural confounding – i.e. lack of exchangeability (Messer, 2007; Messer et al., 2010; Oakes, 2004; Oakes et al., 2010; Westreich and Cole, 2010). This is the idea that specific types of people “select” into specific types of neighborhoods, whether by choice or driven by socioeconomic processes, and as a result some neighborhoods may lack specific types of people. This could limit the ability to infer from regression analyses the effect of neighborhood class on an outcome. (Although here youth don’t “select” their neighborhood, their parents do.) In our study the relative homogeneity of demographic and SES variables across the classes is desirable in that it provides participants who are relatively similar (i.e. exchangeable) except for their neighborhood class (i.e. the exposure of interest), limiting structural confounding and improving the estimation of the effect of neighborhood class on the PA outcomes.
A study (Nelson et al., 2006) similar to the current study, but using cluster analysis in a nationally representative dataset of adolescents, identified six distinct neighborhood patterns based on neighborhood characteristics thought to be associated with PA including: income, race/ethnicity, SES, crime, road type, street connectivity, and recreation facilities. The study identified these neighborhoods as: 1) rural working class; 2) exurban; 3) newer suburban; 4) upper-middle class, older suburban; 5) mixed-race urban; and 6) low-socioeconomic-status inner-city. When they examined the association between neighborhood type and self-reported adolescent PA (and weight), they found differences in PA (and overweight) by some neighborhood patterns. We failed to find similar results possibly because our neighborhoods and our population, all from the Twin Cities in Minnesota, were more homogeneous than the national sample studied by Nelson et al. Still, we were able to identify distinct neighborhood classes based on the built environment, populated by a rather homogeneous sample, and were not able to see that the built environment made a difference in adolescent levels of PA, sedentary activity, or screen time. The differing results may also be due to our classification method, sample size, study area, or other factors discussed above. However, similar to Nelson et al. (2006) we agree that the typical urban-suburban-rural classification might hide important differences among neighborhood and PA patterns. We examined the degree to which the neighborhood classes in this study identified through LCA conformed to the conventional definitions of outer ring suburb, inner ring suburb, and urban area. Using kappa statistics we evaluated the concordance by comparing outer ring suburb, inner ring suburb, and urban area, as established through zip codes, with the LCA-defined class typologies. We found very low levels of agreement (kappa = 0.20) suggesting that there is something unique about the characteristics of neighborhoods that goes beyond the conventional definitions and that the conventional classification may mask important differences.

A key strength of this study was the novel extension of an LCA to classify neighborhoods and then to examine the association of neighborhood class with adolescent PA. Using LCA to identify classes of exposures explicitly acknowledges the complexity of built environments, i.e. the web of factors that make up a neighborhood. Another strength of the study was the use of objective measures of PA, which limits the measurement error of the PA outcome variables, in comparison to the more often used self-reported PA. Finally, the relative homogeneity of the participants across the neighborhood classes provided exchangeability, which allowed us to better identify the effect of neighborhood class on adolescent PA, sedentary behavior, and screen time.

Although we failed to find compelling evidence of an association between neighborhood class and adolescent PA, sedentary time, or screen time, our results add to the literature in this area. Together these results suggest that the association between the built environment and PA in general, and adolescent PA more specifically, is complicated to untangle given the diversity of challenges in this area of research. If we hope to draw conclusions from observational studies, at a minimum we need to establish clear definitions of neighborhoods and consistent measures of the built environment. We may also need to expand the area we consider an individual’s neighborhood to include the area surrounding school and other non-home locations where the adolescent spends a significant portion of their time (e.g. friends’ homes or sports clubs). In depth examination of qualitative data related to perceptions and preferences will be an important advance to this literature. Further research is also warranted using neighborhood classification methods, such as LCA, because they explicitly address the complex web of characteristics that individuals are exposed to in a neighborhood, which may affect their PA. Perhaps with additional research using LCA in other locations consistent patterns will emerge.
Acknowledgments

Funding: The project described was supported by grant number R01HL085978 from the National Heart, Lung, and Blood Institute and grant 1U54CA116849-01 from National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung and Blood Institute or the National Cancer Institute.

References


McDonald KN, Oakes JM, Forsyth A. Effect of street connectivity and density on adult BMI: results from the Twin Cities Walking Study. J Epidemiol Community Health. 20111.0.1136/jech. 2010.122556


Health Place. Author manuscript; available in PMC 2013 March 01.


Figure 1.
Mean item-response probabilities of built environment features by class, TREC IDEA, 2006.
Table 1A
Sample Characteristics, TREC IDEA, 2006 (n=344)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>15.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Percent male</td>
<td>49.1</td>
<td></td>
</tr>
<tr>
<td>Percent white</td>
<td>93.6</td>
<td></td>
</tr>
<tr>
<td>Percent with college</td>
<td>75.0</td>
<td></td>
</tr>
<tr>
<td>Percent free/reduced</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>number of people in the</td>
<td>4.0</td>
<td>1.3</td>
</tr>
<tr>
<td>household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pubertal Status</td>
<td>3.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Activity counts per minute</td>
<td>356.5</td>
<td>161.4</td>
</tr>
<tr>
<td>Minutes of MVPA per day</td>
<td>30.7</td>
<td>16.4</td>
</tr>
<tr>
<td>Minutes of VPA per day</td>
<td>3.0</td>
<td>4.4</td>
</tr>
<tr>
<td>Minutes of activity per day</td>
<td>287.8</td>
<td>68.7</td>
</tr>
<tr>
<td>Minutes of sedentary time per day</td>
<td>601.3</td>
<td>96.9</td>
</tr>
<tr>
<td>Minutes of screen time per day</td>
<td>305.6</td>
<td>206.9</td>
</tr>
<tr>
<td>Parent walking minutes per week</td>
<td>82.3</td>
<td>86.7</td>
</tr>
<tr>
<td>Number of students per school</td>
<td>2.7</td>
<td>3.2</td>
</tr>
</tbody>
</table>
### Table 1B

Built Environment Characteristics, TREC IDEA, 2006 (n=344)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit stop density</td>
<td>0.07</td>
<td>0.07</td>
<td>0.06</td>
<td>0</td>
<td>0.30</td>
</tr>
<tr>
<td>Median block size (ha)</td>
<td>7.71</td>
<td>17.00</td>
<td>3.26</td>
<td>1.04</td>
<td>163.54</td>
</tr>
<tr>
<td>Busy street density</td>
<td>0.26</td>
<td>0.17</td>
<td>0.22</td>
<td>0.00</td>
<td>0.89</td>
</tr>
<tr>
<td>Distance to trail access</td>
<td>1427</td>
<td>3111</td>
<td>457</td>
<td>1</td>
<td>33,153</td>
</tr>
<tr>
<td>Distance to recreation center</td>
<td>1591</td>
<td>1475</td>
<td>1188</td>
<td>0</td>
<td>11,062</td>
</tr>
<tr>
<td>Distance to park</td>
<td>3468</td>
<td>2866</td>
<td>2841</td>
<td>0</td>
<td>26,486</td>
</tr>
<tr>
<td>Distance to school attended</td>
<td>6288</td>
<td>5829</td>
<td>4515</td>
<td>398</td>
<td>47,761</td>
</tr>
<tr>
<td>Distance to gym</td>
<td>2188</td>
<td>2657</td>
<td>1617</td>
<td>0</td>
<td>32,193</td>
</tr>
<tr>
<td>Retail establishment density</td>
<td>0.011</td>
<td>0.013</td>
<td>0.006</td>
<td>0.000</td>
<td>0.058</td>
</tr>
<tr>
<td>Walkability index</td>
<td>0.17</td>
<td>2.45</td>
<td>0.20</td>
<td>3.95</td>
<td>9.03</td>
</tr>
</tbody>
</table>

Distances are to the nearest feature in meters.

Densities are units per total land area.
Table 2

Model fit statistics for repeated estimation of the latent class analysis, TREC IDEA, 2006

<table>
<thead>
<tr>
<th>Class</th>
<th>Log-Likelihood</th>
<th>G-squared</th>
<th>AIC</th>
<th>BIC</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Class</td>
<td>−2034.59</td>
<td>801.51</td>
<td>843.51</td>
<td>924.16</td>
<td>0.86</td>
</tr>
<tr>
<td>3-Class</td>
<td>−1989.41</td>
<td>711.16</td>
<td>775.16</td>
<td>898.06</td>
<td>0.84</td>
</tr>
<tr>
<td>4-Class</td>
<td>−1944.78</td>
<td>621.89</td>
<td>707.89</td>
<td>873.04</td>
<td>0.85</td>
</tr>
<tr>
<td>5-Class</td>
<td>−1924.34</td>
<td>581.01</td>
<td>689.01</td>
<td>896.40</td>
<td>0.84</td>
</tr>
</tbody>
</table>
### Table 3

Reliability of LCA Class Assignments (n=342), TREC IDEA 2006

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posterior Probabilities for Class 1</td>
<td>112</td>
<td>0.96</td>
<td>0.99</td>
<td>0.62</td>
<td>1.00</td>
<td>0.08</td>
</tr>
<tr>
<td>Posterior Probabilities for Class 2</td>
<td>100</td>
<td>0.93</td>
<td>0.97</td>
<td>0.54</td>
<td>1.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Posterior Probabilities for Class 3</td>
<td>91</td>
<td>0.89</td>
<td>0.97</td>
<td>0.50</td>
<td>1.00</td>
<td>0.15</td>
</tr>
<tr>
<td>Posterior Probabilities for Class 4</td>
<td>39</td>
<td>0.89</td>
<td>0.94</td>
<td>0.60</td>
<td>1.00</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Table 4

Mean item-response probabilities\(^a\) of built environment features by class, TREC IDEA, 2006.

<table>
<thead>
<tr>
<th>Latent Class</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit stop density</td>
<td>0.000</td>
<td>0.895</td>
<td>0.736</td>
<td>0.297</td>
</tr>
<tr>
<td>Median block size</td>
<td>0.788</td>
<td>0.212</td>
<td>0.452</td>
<td>0.632</td>
</tr>
<tr>
<td>Busy street density</td>
<td>0.611</td>
<td>0.356</td>
<td>0.375</td>
<td>0.892</td>
</tr>
<tr>
<td>Distance to trail access</td>
<td>0.734</td>
<td>0.501</td>
<td>0.265</td>
<td>0.388</td>
</tr>
<tr>
<td>Distance to recreation center</td>
<td>0.893</td>
<td>0.036</td>
<td>0.738</td>
<td>0.000</td>
</tr>
<tr>
<td>Distance to park</td>
<td>0.752</td>
<td>0.735</td>
<td>0.844</td>
<td>0.406</td>
</tr>
<tr>
<td>Distance to school attended</td>
<td>0.734</td>
<td>0.267</td>
<td>0.419</td>
<td>0.720</td>
</tr>
<tr>
<td>Distance to gym</td>
<td>0.872</td>
<td>0.059</td>
<td>0.738</td>
<td>0.000</td>
</tr>
<tr>
<td>Walkability index</td>
<td>0.019</td>
<td>0.968</td>
<td>0.733</td>
<td>0.038</td>
</tr>
<tr>
<td>Retail establishment density</td>
<td>0.027</td>
<td>0.931</td>
<td>0.627</td>
<td>0.394</td>
</tr>
</tbody>
</table>

\(^a\)Probability of endorsing item given latent class
Table 5

Mean values of selected PA outcomes by latent class, TREC IDEA, 2006 (n=342).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Class 1 (n=106)</th>
<th>Class 2 (n=97)</th>
<th>Class 3 (n=89)</th>
<th>Class 4 (n=39)</th>
<th>Chi-Square</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean minutes/day of MVPA</td>
<td>29.9 (1.6)</td>
<td>31.8 (1.8)</td>
<td>31.8 (2.0)</td>
<td>28.4 (2.1)</td>
<td>2.2</td>
<td>0.53</td>
</tr>
<tr>
<td>Mean minutes/day of VPA</td>
<td>3.4 (0.5)</td>
<td>2.8 (0.4)</td>
<td>3.0 (0.4)</td>
<td>2.5 (0.3)</td>
<td>2.0</td>
<td>0.57</td>
</tr>
<tr>
<td>Mean minutes/day of activity</td>
<td>292.1 (6.5)</td>
<td>288.3 (6.3)</td>
<td>287.5 (6.8)</td>
<td>280.4 (8.1)</td>
<td>1.1</td>
<td>0.78</td>
</tr>
<tr>
<td>Mean activity count/min/day</td>
<td>362.9 (16.0)</td>
<td>364.4 (14.6)</td>
<td>348.3 (16.6)</td>
<td>345.4 (16.8)</td>
<td>1.2</td>
<td>0.74</td>
</tr>
<tr>
<td>Mean minutes/day of sedentary time</td>
<td>594.4 (11.3)</td>
<td>608.6 (8.7)</td>
<td>608.2 (8.4)</td>
<td>592.4 (12.5)</td>
<td>1.6</td>
<td>0.66</td>
</tr>
<tr>
<td>Mean screen time, min/day</td>
<td>332.3 (20.7)</td>
<td>295.5 (24.0)</td>
<td>287.2 (18.9)</td>
<td>318.7 (35.7)</td>
<td>3.1</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Adjusted for age, gender, race/ethnicity, free/reduced lunch, puberty, household adult education and number of people in the house.

*Chi-square used to assess significant differences between classes.

There are no significant differences between classes.