Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity?

Lawrence Douglas Frank, Brian Saelens, Ken E. Powell, James E. Chapman

Abstract

Evidence documents associations between neighborhood design and active and sedentary forms of travel. Most studies compare travel patterns for people located in different types of neighborhoods at one point in time adjusting for demographics. Most fail to account for either underlying neighborhood selection factors (reasons for choosing a neighborhood) or preferences (neighborhoods that are preferred) that impact neighborhood selection and behavior. Known as self-selection, this issue makes it difficult to evaluate causation among built form, behavior, and associated outcomes and to know how much more walking and less driving could occur through creating environments conducive to active transport. The current study controls for neighborhood selection and preference and isolates the effect of the built environment on walking, car use, and obesity. Separate analyses were conducted among 2056 persons in the Atlanta, USA based Strategies for Metropolitan Atlanta’s Regional Transportation and Air Quality (SMARTRAQ) travel survey on selection factors and 1466 persons in the SMARTRAQ community preference sub-survey. A significant proportion of the population are “mismatched” and do not live in their preferred neighborhood type. Factors influencing neighborhood selection and individual preferences, and current neighborhood walkability explained vehicle travel distance after controlling for demographic variables. Individuals who preferred and lived in a walkable neighborhood walked most (33.9% walked) and drove 25.8 miles per day on average. Individuals that preferred and lived in car dependent neighborhoods drove the most (43 miles per day) and walked the least (3.3%). Individuals that do not prefer a walkable environment walked little and show no change in obesity prevalence regardless of where they live. About half as many participants were obese (11.7%) who prefer and live in walkable environments than participants who prefer car dependent environments (21.6%). Findings suggest that creating walkable environments may result in higher levels of physical activity and less driving and in slightly lower obesity prevalence for those preferring walkability.

Keywords: Built environment; Preference; Obesity; Travel; Physical activity; USA; Neighborhood

© 2007 Elsevier Ltd. All rights reserved.
Introduction

Land use patterns, transportation systems, and site design are aspects of the built environment that relate with travel behavior. Urban transportation planning (Cervero & Kockelman, 1997; Frank & Pivo, 1995; Handy, 1996) and public health (Craig, Brownson, Cragg, & Dunn, 2002; Saelens, Sallis, & Frank, 2003) research documents associations between built environment, transportation modes and distances, and overall physical activity. With few exceptions, however, current research has been cross-sectional, making the directionality of influence between place and behavior uncertain (TRB & IOM, 2005). Uncertainty also exists regarding whether observed differences in behavior across urban forms are due to individuals’ shared underlying preferences for travel activity and residential choices or an independent function of environment.

The current study examines the relative effect of neighborhood self-selection, preference for neighborhood type, and the built environment on walking, vehicle miles traveled (VMT), and obesity. Neighborhood selection is defined as the self-reported level of importance of ease of walking for moving to one’s neighborhood. Neighborhood preference is defined as an individual’s self-reported preference for neighborhoods that are typically suburban or urban, regardless of current neighborhood.

Literature review

The built environment, physical activity, and vehicle travel

Several recent and comprehensive reviews (Frank, Engelke, & Schmid, 2003; Frumkin, Frank, & Jackson, 2004; Heath et al., 2006; Saelens, Sallis, Black, & Chen, 2003; Sallis, Frank, Saelens, & Kraft, 2004; Transportation Research Board and Institute of Medicine, 2005) report the presence of a consistent positive association between built environment walkability and engagement in non-motorized travel. Residents of neighborhoods with higher population density, proximity to commercial destinations, and good public transportation are more physically active than residents of less walkable neighborhoods often deemed “suburban” (Frank, Schmid, Sallis, Chapman, & Saelens, 2005). Walk trips are more frequent and per capita distances traveled in cars is less for those living in more compact urban settings (Ewing & Cervero, 2001; Frank et al., 2006; Holtzclaw, Clear, Dittmar, Goldstein, & Haas, 2002). Several recent studies also show an inverse association between neighborhood walkability and obesity (Ewing, Schmid, Killingsworth, Zlot, & Raudenbush, 2003; Frank, Andresen, & Schmid, 2004; Lopez, 2004; Saelens et al., 2003), although these associations are less consistent in the literature.

Self-selection

Observed neighborhood-associated behavioral differences may be spurious and merely reflect shared underlying lifestyle preferences that impact both location and travel choice. Recent studies suggest that both self-selection and built environment are responsible for the travel and activity differences between residents of urban and suburban neighborhoods (Bagley & Mokhtarian, 2002; Handy, Cao, & Mokhtarian, 2006; Khattak & Rodriguez, 2005; Kitamura, Mokhtarian, & Laidet, 1997; Schwanen & Mokhtarian, 2004, 2005a, 2005b). Their conclusions derive from the ability to stratify individuals according to their neighborhood preferences as well as the physical features of their neighborhoods. Using a quasi-longitudinal design, Handy et al. (2006) concluded that the built environment influences walking behavior after taking neighborhood preferences and attitudes into account.

Establishing causation will require disentangling the effect of attitudes and preferences from the physical environment on transportation, health, and environmental outcomes. The current study contributes to the evidence base by exploring how factors influencing neighborhood selection and neighborhood preferences influence the observed association between built environment and walking for non-discretionary (transport) or discretionary (leisure) purposes, vehicle miles traveled, and obesity.

Conceptual framework

The conceptual framework for our analyses (Fig. 1) considers the built environment to be an enabler or disabler for active transportation. The model proposes that demographic characteristics and life experiences create (arrow a) a wide ranging set of lifestyle preferences, including preferences for residential location, character, and travel options.
To the extent these preferences can be met (arrow b.1), the transportation system, land use patterns, and design of the built environment facilitate or at least allow the desired travel behaviors, including frequency, distance, and mode choice (e.g., automobile, bicycle, on foot) of trips made for occupational, daily chore, and leisure purposes. These travel behaviors contribute importantly to a wide range of economic, environmental, and health outcomes such as motor vehicle-related expenditures, air pollution, activity, obesity, and associated morbidity and mortality rates.

However, research has shown that in many instances lifestyle preferences for residential location, character, and travel options are not met. The availability and cost of housing, real and perceived quality of schools, job location, and other preferences require compromise. People, therefore, often live in a built environment that is not to their choosing (arrow b.2) or are "mismatched." The transportation system, land use patterns, and design of the un-preferred built environment may impede or in some cases completely prevent use (arrow c.2) of preferred travel behaviors. Some may use the automobile less, others more, and some may alter their levels of walking and transit use. Comparing the travel behavior of persons with similar preferences but located within different types of environments enables the isolation of the physical environment on behavior from attitudinal predisposition.

**Methods**

**Sample**

Two sub-samples were drawn from the 2001/2002 Strategies for Metropolitan Atlanta’s Regional Transportation and Air Quality (SMARTRAQ) study of 8069 households (Frank et al., 2004). These sub-samples are called neighborhood selection ($n = 2088$; $n = 2056$ with complete data) and neighborhood preference ($n = 1455$). Both sub-samples have one person from each household. Eligibility in the Neighborhood Selection sub-sample was based on tenure in residence of less than 3 years. The Neighborhood Preference Sample was derived from a representative sample of the larger SMARTRAQ survey across income and net residential density. Two-hundred and sixty-two persons were in both samples, but samples were evaluated separately in the present study.

The larger SMARTRAQ travel survey comes from the 13-county Atlanta region and those results conform with national travel data, showing similar distributions of trips across purpose (e.g. work and non-work) and across modes of travel (walk, drive, bike, transit, etc.). Data collection was conducted in.
the spring and fall of 2001 and 2002 to avoid seasonal influences of summer and Christmas break when children are not in school and when extreme weather patterns impact travel patterns as well.

Detailed recruitment procedures are described elsewhere (see Frank et al., 2004, 2005). Briefly, a computer-aided telephone interview (CATI) script was employed for recruitment that screened and selected participants based on household income, household size, and net residential density. This stratification was designed to create a statistically representative sample of the households across income and household size ranges that are located within differing urban environments. The overall response rate for the SMARTRAQ survey (30.4%) was calculated by multiplying the response rate for recruitment (44.8%) and for retrieval of data (67.8%). Willing travel survey participants tend to be higher in income, more frequently white, and older than the regional average.

The Neighborhood Selection Sample provided responses to a Neighborhood Selection Questionnaire administered via CATI. Eligibility for participation in this sub-sample included being an adult head of household, a renter or owner, and to have moved within the past 3 years. The Neighborhood Preference Sample is based on a more in-depth stand alone sub-sample on community preferences. Survey completion rates for the selection and preference samples were above 97.5% for eligible households, since households had already been recruited into the larger SMARTRAQ survey. Representativeness varies for both surveys across age, gender, income, and ethnicity. The sample was not weighted to be representative of regional demographic or urban form characteristics.

The land area within the buffer was categorized by use, e.g., commercial, office, multi-family residential (Fig. 2). The buffer areas, land use categories, and census data were used to develop four measures of urban form: the ratio of retail floor area to retail ground area, land use mix, net number of licensed drivers per vehicle was also derived.

Travel behavior

Travel data were collected through a 2-day travel diary which captures the location of all destinations visited, modes of travel, time of day, day of week, duration, distance, and how many people accompanied the participant. Primary and secondary activities in which participants engaged at each destination were recorded. About 70% of all participants in both surveys were assigned two consecutive weekdays. Thirty percent of the Neighborhood Selection and approximately 29% of the Neighborhood Preference participants had both a weekday and weekend day in their survey period. About 85% of all the travel days in each analysis were weekdays.

Walk trips were characterized as non-discretionary if the reported destination activities included eating, work, school, shop/errand, personal, and/or transport. Discretionary walk trips had a destination activity that included entertainment, visiting friends/relatives, community meetings, political/civic events, public hearing, worship/religious meeting, fitness/exercising, and/or outdoor recreation. Walk trips for recreation such as around the block were included as discretionary.

Average daily number of vehicle miles traveled was estimated using GIS software to determine the shortest-time path between origin and destination. Local road travel is estimated by assigning trips to street segments that form the shortest path from the destination visited to the modeled road network.

Built environment measures

County-level tax assessor’s data, regional land use data from digital aerial photography, street network data, and census data were integrated within a Geographic Information System to measure urban form characteristics of the immediate neighborhood (buffer) around each household. The buffers used were road-network based polygons created by traveling 1-km from each participant’s house. The road network used only contained roads on which pedestrian travel was not restricted.

The land area within the buffer was categorized by use, e.g., commercial, office, multi-family residential (Fig. 2). The buffer areas, land use categories, and census data were used to develop four measures of urban form: the ratio of retail floor area to retail ground area, land use mix, net...
residential density, and street connectivity. These four measures were then used to calculate a measure of walkability around each participant’s place of residence.

*Commercial floor area ratio:* Commercial floor area ratio is the ratio of total commercial building floor area ratio (FAR) to the total land area devoted to commercial uses within the buffer. FAR provides an indicator of the degree to which a piece of land is covered in surface parking or where storefronts are set up close to the sidewalk allowing easy access from pedestrian areas. A ratio greater than one indicates the building contains more area than the land it occupies. A low ratio, e.g. 0.1, indicates the building covers very little of the parcel it is on (10%), and given the commercial use it is likely surrounded by parking and expected to encourage automobile and restrict pedestrian access.

*Land-use mix:* The measure of mixed use captures the evenness of distribution of square footage of development across residential, commercial and office land uses within each participant’s buffer around the household as follows:

\[
LUM = - \sum_{i=1}^{n} (p_i \ln p_i) \ln n,
\]

where \(p_i\) is the proportion of estimated building square footage attributed to land use \(i\), and \(n\) is the number of land uses.

The proportion of estimated square footage obtained from the county tax assessment database (proportional to the size of the network buffer) controls for land uses within the network buffers that were not considered to encompass walkable destinations such as industrial areas. This method of measurement ensures that an area that is evenly distributed with respect to the three land uses, but has a relatively small amount of these uses (residential, commercial, and office) will not have the same value as an area that is also evenly distributed with a much greater critical mass of uses evenly balanced. The measure is an entropy index and was developed by Cervero (1988) and used by others (Cervero & Kockelman, 1997; Frank & Pivo, 1995). Land-use mix ranges from zero to one, with zero representing a single land-use environment and one representing a highly mixed use environment. Only residential, commercial, and office uses are employed in the mixed use measure for the current study.

*Net residential density:* Net residential density equals the number of residential units per residential acre within the household’s buffer. The count of residential units is based on census block group. The amount of residential land is parcel data based.
Connectivity: Connectivity was defined as the number of intersections with more than three legs per square kilometer within the household buffer. Street networks capture the degree to which destinations can be reached in a direct, rather than an indirect pathway, and predict the relative ease of walking, which is highly sensitive to distance (Frank et al., 2003).

A composite measure of walkability

To avoid problems with multicollinearity, a walkability index was established that integrates the four walkability variables. This index was created for the entire SMARTRAQ survey of 8069 households and normalized distributions (z scores) were taken for each of four urban form variables measured using a 1-km network buffer. The z scores were then combined using the formula below:

\[
\text{Walkability index} = \frac{z}{C} \text{-score of commercial building floor area to land area ratio} + \frac{z}{C} \text{-score land use mix} + \frac{z}{C} \text{-score net residential density} + \frac{z}{C} \text{-score intersection density.}
\]

The walkability index was then quartiled for the set of households in the full SMARTRAQ (8069 Household) survey, thus establishing a consistent definition of walkability across both the neighborhood selection and neighborhood preference samples (although resulting in somewhat unequal sample sizes within quartiles). Quartiling walkability allowed us to examine ‘non-linear’ associations with travel and obesity outcomes across significant degrees of change in neighborhood urban form.

Table 1

<table>
<thead>
<tr>
<th>Self-selection item descriptives and factor loadings (n = 2088)</th>
<th>Mean (SD)</th>
<th>Non-motorized selection factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low crime</td>
<td>3.83 (1.33)</td>
<td>-0.139</td>
</tr>
<tr>
<td>Affordability</td>
<td>3.51 (1.33)</td>
<td>0.109</td>
</tr>
<tr>
<td>Closeness to job</td>
<td>3.51 (1.54)</td>
<td>0.394</td>
</tr>
<tr>
<td>Near shops and services</td>
<td>3.51 (1.30)</td>
<td>0.275</td>
</tr>
<tr>
<td>Near major roads and interstates</td>
<td>3.48 (1.38)</td>
<td>0.230</td>
</tr>
<tr>
<td>Ease of walking</td>
<td>3.21 (1.52)</td>
<td>0.549^a</td>
</tr>
<tr>
<td>Low transportation costs</td>
<td>3.15 (1.51)</td>
<td>0.596^a</td>
</tr>
<tr>
<td>Near outdoor recreation</td>
<td>2.94 (1.41)</td>
<td>0.195</td>
</tr>
<tr>
<td>Quality of schools</td>
<td>2.76 (1.76)</td>
<td>-0.072</td>
</tr>
<tr>
<td>Near to public transit</td>
<td>2.57 (1.64)</td>
<td>0.712^a</td>
</tr>
</tbody>
</table>

^aItems selected to constitute the non-motorized factor.

Neighborhood selection

The Neighborhood Selection Questionnaire included 10 items assessing reasons for moving to one’s neighborhood (see Table 1 for items), including a 5-point Likert-type response format (1 = not at all important to 5 = very important). Principal components analysis with an oblique Promax rotation identified a single factor explaining the most variance. The items “ease of walking”, “low transportation costs”, and “near to public transit” (see Table 1) loaded most highly on this factor labeled “non-motorized selection”. Individuals were assigned values for this factor by averaging responses to these three items. Quartiling of these values resulted in the lowest quartile having values ≤2, the second quartile with values >2 and <3, the third quartile with values ≥3 and <4, and the fourth quartile having values ≥4.

Neighborhood preference

The Stated Preference survey for SMARTRAQ was a choice-based conjoint method used by marketing professionals both to gauge the extent to which demand for a specific product exists, and to understand the attributes of specific products consumers value when deciding between competing products or goods (Boyle, Holmes, Teisl, & Roe, 2001; Green & Srinivasan, 1990; Wittink & Cattin, 1989). Thus, it assumes that, for a static cost, the respondent can meaningfully choose between two competing development types, each providing particular benefits and disadvantages.” (Levine & Frank, 2006).

Respondents were asked to consider specific tradeoffs between aspects of travel convenience and neighborhood design. Respondents provided...
their degree of preference (using an 11-point Likert-type response format). The following seven trade-offs were posed to respondents:

- walkability versus commercial–residential land use separation;
- commute distance versus residential density (e.g., lot size);
- urban vitality versus low-density and single-use neighborhoods;
- commute distance versus living on a quieter cul-de-sac street (e.g., streets permitting through traffic versus streets arranged in cul-de-sacs);
- availability of alternatives to the car versus home size;
- accommodation of the automobile versus accommodation of pedestrians and cyclists; and
- availability of alternatives to the car versus neighborhood privacy (defined as cul-de-sac layout and little pedestrian or auto traffic from other neighborhoods).

An example of the layout of each of the questions is provided in Fig. 3. Respondents responded for preferences alone, as price was constant, in order to facilitate this comparison.

A single measure of neighborhood preference ranging from auto to pedestrian/transit oriented was operationalized from the set of seven paired trade-offs using principle components analysis (see Levine & Frank, 2006). Table 2 shows the factor loadings.

The neutral position (between pedestrian-oriented and auto-oriented neighborhoods) was set at zero by normalizing the neighborhood preference factor. Positive values indicate a preference for transit- and pedestrian-oriented neighborhoods, and negative values indicate preferences for auto-oriented neighborhoods. Quartiling of these values resulted in the lowest quartile having values $\leq (-0.43)$, the second quartile with values $>-0.43$ and $<0.53$, the third quartile with values $\geq 0.53$ and $<1.3$, and the fourth quartile having values $\geq 1.3$.

Table 2
Neighborhood preference factor loadings ($n = 1455$)

<table>
<thead>
<tr>
<th>Loadings on the neighborhood preference factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to walk to nearby shops and services</td>
</tr>
<tr>
<td>Lot size and travel distance</td>
</tr>
<tr>
<td>Level of activity and mix of housing</td>
</tr>
<tr>
<td>Street pattern and travel distance</td>
</tr>
<tr>
<td>House size and travel options</td>
</tr>
<tr>
<td>Street design</td>
</tr>
<tr>
<td>Street types and travel options</td>
</tr>
</tbody>
</table>

Data analytic plan

Separate analyses were conducted on the two study samples using the same analytic approach. The primary purpose of the investigation was to determine the relative contribution of urban form (walkability), neighborhood selection or preference, and demographic covariates in explaining physical activity (discretionary or non-discretionary walking), distances traveled (VMT), and obesity.

We first employed chi-square analysis to assess if there was a significant difference in the likelihood people walked for all purposes, non-discretionary (work and food related), and discretionary (leisure, recreation, social, and entertainment) and obesity across quartiled levels of walkability. Similar assessments were conducted to assess if walking, vehicle travel, and obesity levels varied across quartiled levels of neighborhood selection or preference. For individuals that took at least one walk trip, ANOVA analyses were conducted to assess how each quartile increase in walkability, non-motorized selection, and neighborhood preference translated into more trips walking trips.

Separate logistical regression analyses were then performed with the following dependent variables: the likelihood participants reported walking for any purpose, non-discretionary walking, discretionary walking, and obesity. The models aimed to identify the independent relationship with demographic factors, area walkability, neighborhood selection and neighborhood preference. For instances in which the prevalence exceeded 10% and the logistic odds ratios associated with preference/selection or walkability were significantly different than the reference group, corrected odds ratios are reported to provide a more accurate estimate of relative risk (Zhang & Yu, 1998). While there is some debate about the preferred estimate for relative risk (e.g., prevalence rate ratios versus odds ratios; Zocchetti, Consonni, & Bertazzi, 1997), the odds ratio, particularly when adjusted in the case of high prevalence, appears an acceptable metric (Pearce, 2004).

Linear regression analyses were performed to assess the relative contribution of walkability versus neighborhood preference and in a separate assessment, neighborhood selection, in explaining the variation in vehicle miles traveled when adjusting for demographic factors. Partial r-squared results are provided ($r^2$) to denote the amount of explained variance attributable to each variable.

Participants in the residential preference sample were then placed into groups based on the match between their residential preference for a more or less walkable environment and conditions where they actually live (the classification is explained below). Descriptive analyses were conducted to convey the mean percent that walked, were obese, and miles driven for each group.

Results

Descriptive statistics are presented in Table 3 characterizing the Neighborhood Selection and the Neighborhood Preference samples. Participants in the neighborhood selection sample are considerably younger, of higher minority concentration, and have an overall income that is lower than those in the neighborhood preference sample. Participants on the neighborhood selection sample took fewer walk trips and had a slightly higher proportion that are obese than those in the neighborhood preference sample. Both samples were representative of the regional distribution across gender and household size. The selection sample was a bit younger and the preference sample older than the regional average for age. The selection sample was closer to the regional distribution in terms of ethnicity and income. Forty-nine percent of the Neighborhood Selection and 85% of the Neighborhood Preference participants owned or were buying their place of residence. The current study does not address how selection and preferences may differ for renters and owners.

Neighborhood selection, walking, and obesity

As seen in Table 4, the reported likelihood of any walk trip, or a non-discretionary or discretionary walk trip was significantly higher in the third and fourth non-motorized selection and walkability quartiles. Differences by non-motorized selection or walkability quartile were less common when examining number of walk trips among the subsample of individuals taking at least one walk trip. Vehicle miles traveled were significantly higher for respondents whom the non-motorized factors were least important (41.0) versus those who rated them as most important (28.2) in choosing their residence. Respondents in the least walkable environments generated the most (45.5) miles of travel versus those in the most walkable environments (28.3). Unexpectedly, obesity prevalence was higher in the...
second versus 1st non-motorized selection quartile. As expected, prevalence was lower in the fourth (most walkable) versus the first (least walkable) walkability quartile.

Results from logistic regressions on the neighborhood selection sample predicting the likelihood to walk for any purpose, for non-discretionary or discretionary purposes, and on the likelihood of being obese are provided in Table 5. Lower age, fewer motorized vehicles, lower proportion of licensed drivers, increased importance of non-motorized selection, and increased walkability were all significant predictors of increased likelihood of any walk trips (pseudo \( R^2 = 0.15 \)).

Similar demographic variables along with higher household income, and greater importance of non-motorized selection and greater walkability, were also related to the likelihood of any non-discretionary walk trips (pseudo \( R^2 = 0.14 \)). Individuals in both the third and fourth quartiles of the non-motorized selection factor and walkability had significantly higher odds of any walk trips and non-discretionary walk trips than first quartile individuals for the selection and walkability factors (see Table 5).

Those who were non-black, had fewer licensed drivers in the household, reported greater non-motorized selection and greater walkability, had an increased likelihood of any discretionary walk trips (pseudo \( R^2 = 0.14 \)). The odds of walking for discretionary purposes were higher in the third and fourth quartile of non-motorized selection than the lowest quartile, although only the fourth quartile group on walkability showed significantly greater odds of a discretionary walk trip.

Being older, black, poorer, and in a larger household was associated with increased odds of obesity. Greater walkability was associated with lower odds of obesity, especially when comparing the fourth and first walkability quartile (see Table 5). Non-motorized selection was not a significant predictor of being obese, although the overall model fit for obesity was lower than for the walking outcomes (pseudo \( R^2 = 0.08 \)). Gender was tested but was not found to be significant in any of the models presented in Table 5.

Please cite this article as: Frank, L. D., et al. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and.... *Social Science & Medicine* (2007), doi:10.1016/j.socscimed.2007.05.053
Neighborhood preference, walking, and obesity

The likelihood of taking any walk trips, or walk trips for non-discretionary, or discretionary purposes was significantly higher with each quartile increase in the preference for a walkable environment (see Table 6). The same pattern was observed with each quartile increase in walkability. Only those that prefer the most walkable environments reported significantly more walk trips for any purpose or for non-discretionary walk trips. There were no differences across quartiles of neighborhood preference in the number of total, non-discretionary, or discretionary walk trips among respondents taking at least one walk trip. A similar pattern emerged across quartiles of neighborhood walkability, however, those in the most walkable environments (fourth quartile of walkability) were much more likely to take more walk trips for non-discretionary purposes.

Vehicle miles traveled were significantly higher for respondents who do not prefer (41.1) versus those that most prefer (29.9) to live in a walkable environment. Respondents in the least walkable environments generated the most (46.0) miles of travel versus those in the most walkable environment drove the fewest miles per day (26.7). Obesity prevalence was much higher (23.8%) for those that prefer the least versus the most walkable environments (9.0%). Obesity prevalence was also higher for those in the least (22.2%) versus those in the most walkable environments (15.3%).

Table 7 summarizes the logistic models on the neighborhood preference sample for any walk trips, non-discretionary walk trips, discretionary walk trips, and obesity.

---

Table 4
Walk trips, vehicle miles traveled, and obesity by neighborhood selection and walkability quartile

<table>
<thead>
<tr>
<th>Walkability quartile (% of sample)</th>
<th>% Taking a walk trip</th>
<th>% Taking a non-discretionary walk trip</th>
<th>Number of walk trips</th>
<th>Number of non-discretionary walk trips</th>
<th>Number of discretionary walk trips</th>
<th>Vehicle miles traveled</th>
<th>% Obese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest (17%)</td>
<td>5.9</td>
<td>5.4</td>
<td>1.4</td>
<td>1.3 (0.8)</td>
<td>1.1 (0.8)</td>
<td>0.6 (0.2)</td>
<td>45.5</td>
</tr>
<tr>
<td>2nd quartile (21%)</td>
<td>6.2</td>
<td>5.5</td>
<td>1.4</td>
<td>1.5 (1.2)</td>
<td>1.2 (0.8)</td>
<td>0.7 (0.3)</td>
<td>37.5</td>
</tr>
<tr>
<td>3rd quartile (28%)</td>
<td>15.9*</td>
<td>14.4*</td>
<td>5.0</td>
<td>1.6 (1.1)</td>
<td>1.3 (0.9)</td>
<td>0.6 (0.2)</td>
<td>32.8</td>
</tr>
<tr>
<td>Highest quartile (34%)</td>
<td>23.8*</td>
<td>21.4*</td>
<td>9.4</td>
<td>1.7 (1.2)</td>
<td>1.4 (0.9)</td>
<td>0.7 (0.5)</td>
<td>28.3</td>
</tr>
</tbody>
</table>

*Significantly different (p<0.05) than the lowest quartile reference group.

Please cite this article as: Frank, L. D., et al. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and... Social Science & Medicine (2007), doi:10.1016/j.socscimed.2007.05.053
Being younger, having access to fewer vehicles, greater preference for pedestrian-oriented neighborhoods, and greater walkability (significant for the fourth quartile at the \( p = 0.07 \) level) where one lives were associated with an increased likelihood of any walk trips (pseudo \( R^2 = 0.20 \)). Similar demographic variables, and greater preference for pedestrian-oriented neighborhoods and greater walkability (significant for the fourth quartile at the \( p = 0.06 \) level), were also related to the likelihood of any non-discretionary walk trips (pseudo \( R^2 = 0.21 \)) (pseudo \( R^2 = 0.13 \)). Being white, female, higher income, and one who moderately or strongly prefers a more pedestrian-oriented neighborhood is associated with a lower likelihood of obesity (pseudo \( R^2 = 0.07 \)).

**Neighborhood selection, preference, and vehicle miles traveled**

For the Neighborhood Selection Sample, higher household income, more vehicles in the household, more children < 18 years in the household, greater proportion of licensed drivers per vehicle related to higher vehicle miles traveled (VMT). Increased importance of non-motorized selection and living in a more walkable environment was related with lower VMT (see Table 8; \( R^2 = 0.08 \)). Being older, white, male, having higher household income, and a greater proportion of licensed drivers to vehicles were related to higher VMT. Both non-motorized neighborhood selection and greater walkability were related to lower VMT (see Table 8; \( R^2 = 13.9 \)).

**Matched versus mismatched neighborhood walkability and preference**

Walking, vehicle miles traveled, and obesity is compared amongst participants grouped based on their neighborhood’s walkability and their preferences for a pedestrian or auto-oriented environment. Therefore, four groups were formed as follows: Those that:
Table 6 Walking, driving, and obesity by neighborhood preference and walkability

<table>
<thead>
<tr>
<th>Neighborhood preference</th>
<th>% Taking any walk trips</th>
<th>% Taking any non-discretionary walk trips</th>
<th>% Taking any discretionary walk trips</th>
<th>Number of walk trips</th>
<th>Number of non-discretionary walk trips</th>
<th>Number of discretionary walk trips</th>
<th>Number of vehicle miles traveled</th>
<th>% Obese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest quartile</td>
<td>4.7</td>
<td>4.1</td>
<td>1.7</td>
<td>1.2 (1.6)</td>
<td>0.9 (0.8)</td>
<td>0.6 (0.2)</td>
<td>41.1 (29.7)</td>
<td>23.8</td>
</tr>
<tr>
<td>2nd quartile</td>
<td>9.1**</td>
<td>7.2</td>
<td>3.9</td>
<td>1.4 (1.1)</td>
<td>1.2 (1.1)</td>
<td>0.6 (0.2)</td>
<td>33.3 (25.7)*</td>
<td>23.6</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>20.1*</td>
<td>19.6*</td>
<td>4.7**</td>
<td>1.6 (1.2)</td>
<td>1.2 (0.9)</td>
<td>0.7 (0.4)</td>
<td>31.2 (23.7)*</td>
<td>15.9**</td>
</tr>
<tr>
<td>Highest quartile</td>
<td>34.3*</td>
<td>32.6*</td>
<td>13.0*</td>
<td>2.1 (1.7)**</td>
<td>1.6 (1.3)**</td>
<td>0.8 (0.5)</td>
<td>29.9 (23.2)*</td>
<td>9.0*</td>
</tr>
<tr>
<td>Walkability quartile (% of sample)</td>
<td>8.0</td>
<td>6.7</td>
<td>2.7</td>
<td>1.1 (0.6)</td>
<td>1.1 (0.6)</td>
<td>0.8 (0.4)</td>
<td>46.0 (32.7)</td>
<td>22.2</td>
</tr>
<tr>
<td>Lowest quartile (21%)</td>
<td>8.3</td>
<td>7.7</td>
<td>1.9</td>
<td>1.8 (1.6)</td>
<td>1.3 (1.0)</td>
<td>0.6 (0.2)</td>
<td>36.0 (26.3)*</td>
<td>17.9</td>
</tr>
<tr>
<td>2nd quartile (23%)</td>
<td>17.1*</td>
<td>16.0*</td>
<td>5.2</td>
<td>1.6 (1.3)</td>
<td>1.2 (1.0)</td>
<td>0.7 (0.3)</td>
<td>30.0 (20.6)*</td>
<td>17.6</td>
</tr>
<tr>
<td>3rd quartile (25%)</td>
<td>29.8*</td>
<td>28.3*</td>
<td>11.4*</td>
<td>2.0 (1.7)**</td>
<td>1.5 (1.3)**</td>
<td>0.8 (0.5)</td>
<td>26.7 (20.8)*</td>
<td>15.3**</td>
</tr>
</tbody>
</table>

*Significantly different (p<0.001) than the reference group of lowest quartile.
**Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Significantly different (p<0.05) than the reference group of lowest quartile.

Sample sizes across quartiles are not equal because the cut points are based on the distribution of participants in the full travel survey across walkability.

1. live in low and prefer high walkability;
2. live in high and prefer high walkability;
3. live in low and prefer low walkability;
4. live in high and prefer low walkability.

These four groups correspond with the four rows in Table 9. Inclusion in each of the four groups was based on participants having greater than 0.5 or less than –0.5 on the walkability and preference scales. This criterion results in greater variation between groups and eliminates observations in a boundary condition. Table 9 presents the proportion of participants that walked and are obese and the mean vehicle miles driven within each of the four groups.

For those that live in low and prefer high walkability, 16.0% walked for any purpose, 14.9% walked for non-discretionary, and 2.7% for discretionary purposes. In contrast, 33.9% of those that both prefer and live in a high walkability walked for any purpose, and 32.7% for non-discretionary, and 12.1% for discretionary purposes.

Those that do not prefer a walkable environment walk very little regardless of where they live. For those that live in low and prefer low walkability, 3.3% walked for any purpose, 2.9% walked for non-discretionary, and less than 1% for discretionary purposes. Seven percent of those that live in high and prefer low walkability walked for any purpose, and 4.7% and 4.65% walked for non-
### Table 7
Neighborhood preference, walkability, walking, and obesity

<table>
<thead>
<tr>
<th></th>
<th>Any walk trips</th>
<th></th>
<th>Any non-discretionary walk trips</th>
<th>Any discretionary walk trips</th>
<th>Obesity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est. odds</td>
<td>95% Confidence interval</td>
<td>Est. odds</td>
<td>95% Confidence interval</td>
<td>Est. odds</td>
</tr>
<tr>
<td>Age</td>
<td>0.97</td>
<td>0.96–0.99</td>
<td>0.97</td>
<td>0.96–0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Race&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>0.64</td>
<td>0.47–0.87</td>
<td>0.94</td>
</tr>
<tr>
<td>Gender&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.68</td>
<td>0.56–0.83</td>
<td>0.67</td>
<td>0.55–0.82</td>
<td>0.56</td>
</tr>
<tr>
<td>Total vehicles in the household</td>
<td>1.56</td>
<td>0.89–3.12</td>
<td>1.39</td>
<td>0.71–2.74</td>
<td>1.97</td>
</tr>
<tr>
<td># Licensed drivers/vehicle</td>
<td>3.23&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.93–5.17</td>
<td>3.49&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.02–5.74</td>
<td>2.14</td>
</tr>
<tr>
<td>Neighborhood preference quartile&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4.81&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.98–7.32</td>
<td>5.05&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.02–7.95</td>
<td>4.45&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Walkability quartile</td>
<td>0.72</td>
<td>0.39–1.31</td>
<td>0.8</td>
<td>0.42–1.52</td>
<td>0.6</td>
</tr>
<tr>
<td>2nd quartile</td>
<td>1.11</td>
<td>0.64–1.91</td>
<td>1.2</td>
<td>0.67–2.16</td>
<td>1.28</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>1.62</td>
<td>0.95–2.76</td>
<td>1.72</td>
<td>0.97–3.03</td>
<td>2.14</td>
</tr>
<tr>
<td>4th quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Black; <sup>b</sup> male; <sup>c</sup> non-Black.
<sup>d</sup> Individual quartile comparisons are made relative to the lowest quartile of that variable.
<sup>e</sup> Adjusted odds ratio and 95% CI presented.

### Table 8
Neighborhood selection, preference, and VMT

<table>
<thead>
<tr>
<th></th>
<th>Vehicle miles traveled</th>
<th></th>
<th>Neighborhood preference sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est. odds</td>
<td>p-Value</td>
<td>sr²</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.67</td>
<td>0.89–3.12</td>
<td>1.39</td>
</tr>
<tr>
<td>Race&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.23&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.93–5.17</td>
<td>3.49&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gender&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.81&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.98–7.32</td>
<td>5.05&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Household income&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.72</td>
<td>0.39–1.31</td>
<td>0.8</td>
</tr>
<tr>
<td>Total vehicles in the household</td>
<td>1.11</td>
<td>0.64–1.91</td>
<td>1.2</td>
</tr>
<tr>
<td># Licensed drivers/vehicle</td>
<td>1.62</td>
<td>0.95–2.76</td>
<td>1.72</td>
</tr>
</tbody>
</table>

<sup>a</sup> Black; <sup>b</sup> male; <sup>c</sup> non-Black.
<sup>d</sup> Individual quartile comparisons are made relative to the lowest quartile of that variable.
<sup>e</sup> Adjusted odds ratio and 95% CI presented.

*Please cite this article as: Frank, L. D., et al. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and... Social Science & Medicine (2007), doi:10.1016/j.socscimed.2007.05.053*
higher likelihood of taking a walking trip for any purpose, for non-discretionary purposes and for discretionary travel, as well as lower obesity prevalence. When adjusting for neighborhood preferences, the odds ratios were less extreme. Preference-adjusted odds of walking, at the 90% confidence level, in the most walkable, compared with the least walkable parts of the Atlanta region were: 1.62 (for any purpose); 1.72 (for non-discretionary travel); and 2.14 times (for discretionary travel). Each quartile increase in walkability was associated with a 5.5 mile reduction in vehicle miles traveled when adjusting for demographics co-variates and for neighborhood preference.

Neighborhood selection was not related to the likelihood of being obese while actual area walkability, along with age, race, income, and household size were. It is noteworthy however that after controlling for neighborhood selection factors, it is only in comparing the highest to lowest level of walkability that obesity prevalence differs. In contrast, neighborhood preference along with race, gender, and household income were significant predictors of obesity. Walkability was not a significant predictor of obesity within the ‘preference sample’. Relationships between neighborhood preference, walkability, and obesity became clearer when we grouped participants based on both the walkability of their neighborhood and their preference for neighborhood type. Obesity prevalence varied for participants that prefer a walkable environment and live in one (11.7%) versus those that prefer but do not live in a walkable environment (14.9%). Obesity prevalence did not differ across walkability (about 21.5%) for participants who prefer an auto-oriented environment, but this prevalence was about twice as high as those that prefer and live in a walkable environment (11.7%).

Discussion

Both attitudinal predisposition for neighborhood type and the actual characteristics of the built environment in which one lives were found to impact the choice to walk and distances driven. Participants preferring and selecting neighborhoods with greater walkability walked more for both non-discretionary and discretionary travel; especially when located in environments that are conducive to walking. These findings help to further clarify and support more recent self-selection findings regarding the dual importance of attitude/preference and built form in explaining non-motorized transport.

Participants drove less when located in more walkable environments regardless of their demographic characteristics, the importance of the selection factors tested, and preferences for neighborhood type. When adjusting for neighborhood selection, participants in the most walkable versus least walkable areas of the Atlanta region had...
Considering our findings on obesity, there are many potential explanations for the discrepancy between the selection and preference samples. These samples differed in recruitment criteria, with the selection sample including only individuals who moved in the past 3 years. The selection sample was significantly younger, more ethnically and racially diverse, with slightly lower incomes, lower likelihood of walking for any purpose, and somewhat higher overall obesity prevalence than the preference sample. The preference sample resided in their neighborhoods for a longer period of time, with the possibility that preference and built environment have become more aligned, thus making it more difficult to demonstrate an independent relation with built form. Race and household income were related to obesity status in both samples, with nonwhites and lower income individuals more likely to be obese, but other significant demographic factors differed based on sample. Women were more likely to be overweight in the preference sample, but no gender effects were seen in the selection sample. Age was a positive correlate of being obese in the selection, but not the preference sample. This perhaps has implications for how walkability would affect behaviors and health after considering preference/selection factors. ‘Preference’ scores may be more highly related than ‘selection’ values to other urban form factors involved with obesity (e.g., food availability or the interaction of food availability and quality) and/or other attitudes and beliefs that prevent obesity (e.g., overall health consciousness). The forced choice method by which ‘preference’ was assessed may also more likely reflect the choices and compromises that individuals/households must make when choosing a place of residence, again better aligning ‘preference’ with actual neighborhood walkability and making it more difficult to find independent relations between obesity and actual neighborhood walkability. Clearly, more formative work is needed in best operationalizing self-selection constructs.

Amongst those preferring a walkable environment results suggest that providing more walkable environments will result in increased walking for non-discretionary and discretionary purposes, and reduced vehicle use; and in lower obesity prevalence at the upper end of walkability. However, results suggest that the absolute amount of walking is extremely low for those who do not prefer a walkable environment regardless of neighborhood walkability. Therefore, promoting physical activity by living in a walkable environment is not likely to be a sufficiently effective intervention for this segment of the population.

Neighborhood preferences were more closely related with walking whereas walkability was a better predictor of driving when adjusting for demographic factors. Strategies to promote compact environments may be more effective at reducing per capita vehicle emissions, greenhouse gases, and energy use associated with driving than increasing the amount of walking.

These results support the argument for building communities that better match reported residential preferences. Levine and Frank (2006) found an undersupply of more walkable environments in the Atlanta region. Participants preferring a more walkable environment were less likely to sort themselves into their preferred type of environment than participants who prefer low walkability environments. This suggests an undersupply of walkable environments in the Atlanta region relative to stated demand for such environments. Approximately a third of the residents of the most auto-dependent environments in that region stated they would prefer a more walkable environment. These results bolster the need to remove the regulatory and fiscal barriers to increasing the supply of more walkable environments.

These findings, combined with complementary results from other research, move the field closer to specifying the causal effect of built environment on walking behavior. However, as with most other studies, these findings are limited by being cross-sectional. A more potent design would experimentally isolate built environment effects, perhaps through examining changes in travel behavior among individuals moving from one type of built environment to another, although whether an individual would retain neighborhood attitudes and preferences across such a move is unknown. Longer periods and more complete assessments of walking (e.g., walking duration) are necessary to better specify the health impacts of built environment. In addition, the present models were incomplete, accounting for only a modest proportion of the variance explained in walking, car use, and obesity. These behaviors and outcomes, like most, are multi-determined and require consideration of many types of influences.

Lastly, there are likely many factors that influence neighborhood selection and preference that were not measured in the present study, including 

Please cite this article as: Frank, L. D., et al. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and... Social Science & Medicine (2007), doi:10.1016/j.socscimed.2007.05.053
availability, cost, and other neighborhood characteristics. Although many factors enter into the choice of residential location, cost, job location, quality of public schools, transportation choices, and opportunities for physical activity are likely to be important. Understanding this myriad of factors and their relationships will better allow for the determination of preferred built environments and independent effects of built environment on health behaviors.

Q4 Uncited references


Acknowledgements

The authors acknowledge the financial support of the Georgia Department of Transportation, in particular Mr. George Boulineau, and the Georgia Regional Transportation Authority for funding the SMARTRAQ Project, and the collaboration with the Atlanta Regional Commission to collect the data. We also acknowledge the important contributions to the project made by Dr. Jonathan Levine, Dr. Jim Sallis, Dr. Doug Hunt, Mr. Mark Bradley, Mr. Peter Engelke, Mr. Christopher Leerssen and Ms. Ann Carpenter.

References


Please cite this article as: Frank, L. D., et al. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and... *Social Science & Medicine* (2007), doi:10.1016/j.soscimed.2007.05.053

Heath, G. W., Brownson, R. C., Kruger, J., Miles, R., Powell, K. E., Ramsey, L. T., & the Task Force on Community Preventive Services (2006). The effectiveness of urban design and land use and transportation practices to increase physical activity: A systematic review. *Journal of Physical Activity and Health, 3*(S1), S55–S76.


