INVESTIGATING THE ASSOCIATION BETWEEN THE BUILT ENVIRONMENT AND ACTIVE TRAVEL OF YOUNG ADULTS USING LOCATION BASED TECHNOLOGY

BY

XIAOLU ZHOU

DISSETRATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Landscape Architecture in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor William C. Sullivan, Chair
Professor Weimo Zhu
Associate Professor Brian Deal
Assistant Professor Chen-Chieh Feng
ABSTRACT

Physical inactivity is the second leading modifiable risk of chronic disease. Habitual inactivity prevents young adults from developing healthy patterns of physical activity during the important transition from adolescence to adulthood. Consistent levels of inactivity thus pose a critical, life-long health risk. One promising approach to promoting physical activity is to embed physical activity in daily life by promoting active travel as ways to commute and recreate. A growing body of evidence indicates that certain environmental characteristics promote active lifestyles in general or promote specific kinds of physical activity such as walking or biking. However, studies have also reported inconsistent and sometimes contradictory results of the influence of environment on active travel. These inconsistencies and contradictions stem partly from the challenges of collecting valid data regarding the environmental features and of the locations where people actually travel. The inconsistent results may also stem from the discrepancy between the measured environment and the perceived environment. These challenges prevent us from understanding the strength of the association between environmental features and travel behaviors, and thereby limit the potential to use research results to guide evidenced-based urban design and planning.

This dissertation explores and reduces the research gaps regarding these sets of challenges. First, to better measure an individual’s travel behavior in a convenient and cost-effective manner, I tested the applicability of using a smartphone application that I developed to simultaneously collect location, time, and accelerometer data, and developed a method to automatically classify these data into different travel modes. Second, to overcome the Uncertain Geographic Context Problem (UGCP), I measured the built environmental features at the places where active travel occurred. Then, I modeled the active travel behaviors based on the environmental features using mixed logistic regression. Third, to complement the statistical models, and to reveal people’s own perspectives about the characteristics of the environment that support active travel, I examined geo-tagged photo narrative that the participants provided. These photo narratives reveal information that grows directly from the users of the environment about the specific features of the built environment.
Results from the smartphone data classification demonstrated that smartphone devices are capable of capturing data that reveal how, where, and when people travel. The classification system used in this study achieved more than 80% accuracy in the detection of the type of travel mode people took. Results from the statistical analysis of the relationship between environment and travel behavior showed that greenness was consistently and positively associated with more recreational active travel than vehicle travel in both cities. Destinations in general showed a positive relationship with utilitarian active travel behavior. Crime did not show a significant relationship with different modes of active travel. I also found that a variety of design features such as aesthetics, functionality, destination, and safety were associated with an active lifestyle.

In this study, I employed interdisciplinary methods from geography, public health, and urban planning. I attempted to integrate realms of urban planning and public health by using innovative technologies such as GIS and smartphone sensors to examine travel behaviors. This project also reached to the scale of each individual and probed their concerns on environmental characteristics to promote active travel behavior. I hope the results of this study will help to design urban environment with more active living features. I also hope this study will contribute to combat the obesity and physical inactivity problems that plague cities and make cities more livable and people healthier.
ACKNOWLEDGMENTS

First and foremost, I would like to express my earnest gratitude to my advisor, Professor William Sullivan. This dissertation would not have been possible without his great support. He has read my numerous revisions and helped me to sharpen my ideas, challenge my understanding, clear my train of thoughts, streamline the structure, and guide my writing. I am also very thankful for Professor Sullivan for guiding me into the area of environment and human health and continued support for my study and research. I am sincerely grateful for his supervision in the past four years.

In addition, I would like to thank professors on my doctoral committee, Professor Weimo Zhu, Brian Deal, and Chen-chieh Feng. Thank you all for supervising my progress over the course of my Ph.D. studies. I really appreciate your help including but not limited to offering assistantship, revising journal articles, providing valuable ideas and suggestions, and sharing information and opportunities.

I have also received great help from Linda Larsen, who guided me from the beginning of my dissertation writing until the final product. She gave me many useful suggestions and tips on my dissertation writing. These skills I learned not only helped me to write the dissertation, but also are very instructive for my future writing.

Furthermore, I am very grateful to the people at the consulting group of ATLAS. Thanks ATLAS for offering assistantship to support my Ph.D. study in the past two years. I was very glad to work with my supervisors Maryalice Wu and Dawn Owens-Nicholson, and my colleagues Gina, John, Cameron, Laura, Kit, and Nick.

I am also thankful to my department Landscape Architecture and professors who taught me and guided me, and staff who helped me in all kind of ways. Thanks Professor Ruggles, Elen Deming, Rebecca Ginsburg, Jinki Kim and other professors for mentoring me in the past four years.

Also, tremendous thanks go to all my friends whom I share my laugh, confusions and excitements with and I learned from, people from the lab, Dongying, Bin, Tum, Douglas, and Xiangrong; people from the MLA programs, Yi Qu, Shuangshuang, Peng Chen and so many others; people in my apartment and house, Li Guan, Qingyao, Louis, Alfred, Marietta, and many others friends. This cannot be a complete list, but I would like to thank all the people who helped me during my course of Ph.D. study.
Most of all, I would like to thank my family for your infinite support and encouragement. Thanks for being with me in my journey to the end of my Ph.D. study!
TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. VIII
LIST OF FIGURES ................................................................................................................ IX

CHAPTER 1 INTRODUCTION ................................................................................................. 1
  Challenges in travel data collection ................................................................................... 2
  Challenges in contextualizing travel behavior ................................................................. 3
  Challenges in understanding the perceptions of travelers ............................................... 3
  Dissertation organization ................................................................................................. 4
  References ......................................................................................................................... 5

CHAPTER 2 DOES YOUR SMARTPHONE KNOW HOW YOU TRAVEL? .................. 7
  BACKGROUND .................................................................................................................. 8
    Advancement of travel behavior measures ...................................................................... 8
    Gaps in previous methods .............................................................................................. 9
  METHODS .......................................................................................................................... 12
    Application, study site, and participants ....................................................................... 12
    Minimizing GPS point drop-out and maximizing spatial accuracy .............................. 13
    Reference travel mode .................................................................................................... 13
    Predicting travel mode ..................................................................................................... 14
    Data integration and model predictors .......................................................................... 16
    Classification approach ................................................................................................. 16
  RESULTS ............................................................................................................................. 21
    Accuracy of learning based method .............................................................................. 22
    Accuracy of the combined method ................................................................................ 22
  DISCUSSION ....................................................................................................................... 23
    Advantage of sensor combination ................................................................................ 23
    Advantage of classification method combination ......................................................... 24
    Future research .............................................................................................................. 24
  CONCLUSION ...................................................................................................................... 25
  REFERENCES ....................................................................................................................... 26

CHAPTER 3 BUILT ENVIRONMENT ASSOCIATIONS WITH ACTIVE TRAVEL .... 28
  BACKGROUND .................................................................................................................. 29
  METHODS .......................................................................................................................... 31
    Sites ................................................................................................................................. 31
    Participants ..................................................................................................................... 31
    Procedure ....................................................................................................................... 34
    Built environment measures ......................................................................................... 35
  RESULTS ............................................................................................................................. 36
    Overall travel pattern .................................................................................................... 36
    Influence of environmental features on travel behavior .............................................. 37
    Influence of local variability ......................................................................................... 40
  DISCUSSION ....................................................................................................................... 41
    Main findings ................................................................................................................ 42
    Contributions ................................................................................................................. 42
    Policy implications ........................................................................................................ 44
    Future research ............................................................................................................. 45
    Conclusion ...................................................................................................................... 46
  REFERENCES ....................................................................................................................... 47
CHAPTER 4  WHY DO WE PREFER WALKING HERE? CONNECTING ACTIVE TRAVEL BEHAVIOR WITH ENVIRONMENTAL PERCEPTIONS.............................50
BACKGROUND ..................................................................................51
  Environmental framework for active living........................................51
  Structured and user-derived environmental perceptions ..................52
  Geo-tagged environmental perceptions .........................................55
METHODS .........................................................................................56
  Participants ......................................................................................56
  Travel behaviors ..............................................................................56
  Photo narrative ...............................................................................56
  Photo narrative analysis ..................................................................57
  Mapping environmental perceptions ..............................................57
RESULTS ..........................................................................................58
  Perceptions ....................................................................................58
  Questionnaire-measured and user perceptions ...............................59
  Geographic feature of the perceptions ..........................................62
DISCUSSION ......................................................................................68
  Contributions ..................................................................................69
  Practical implications .......................................................................70
  Future research ...............................................................................72
CONCLUSION .....................................................................................73
REFERENCES ......................................................................................74

CHAPTER 5  CONCLUSION ..................................................................77
  Summary of findings ........................................................................77
  Methodological contributions .........................................................79
  Practical implications .......................................................................81
  Future research ...............................................................................83
  Final remarks ..................................................................................85
  References .......................................................................................86
LIST OF TABLES
Table 1 Results of machine learning-based classification accuracy ...................... 22
Table 2 Result of combined classification and evaluation using machine
learning-based and rule-based method ................................................................... 23
Table 3 Demographic and social confounding variables of the participants in two
cities ......................................................................................................................... 33
Table 4 Adjusted odds ratio of recreational travel behaviors (four modes) associated
with the built environment around each GPS point in Chicago ......................... 39
Table 5 Adjusted odds ratio of utilitarian travel behaviors (four modes) associated
with the built environment around each GPS point in Chicago ......................... 39
Table 6 Adjusted odds ratio of recreational travel behaviors (in three modes)
associated with the built environment around each GPS points in Singapore .. 40
Table 7 Adjusted odds ratio of utilitarian travel behaviors (in three modes)
associated with the built environment around each GPS points in Singapore .. 40
Table 8 Comparison between the adjusted odds ratio of recreational walking
associated with the built environment around each GPS point in UIC and UC.
..................................................................................................................................... 41
Table 9 Summary of the NEWS questionnaire ....................................................... 54
LIST OF FIGURES
Figure 1 Travel route and travel mode plotted over the Google Map. .................. 12
Figure 2 (a) Project instantaneous acceleration to the gravity direction. (b) The step
detection demonstration, in which each red triangle represents a step. ........ 15
Figure 3 The workflow of this study................................................................. 18
Figure 4 Speed and vibration spectrums for walking, biking, running, and in vehicle.
.................................................................................................................. 20
Figure 5 Participants’ travel behavior from four universities............................ 36
Figure 6 Diagram of the environmental features identified by photo-narratives made
by participants from Chicago................................................................. 61
Figure 7 Diagram of the environmental features identified by photo-narratives made
by participants from Singapore.......................................................... 62
Figure 8 The distribution of environmental perceptions in Chicago derived from the
photo-narratives made by study participants ........................................ 64
Figure 9 The distribution of environmental perceptions in Singapore derived from
the photo-narratives made by study participants ..................................... 65
Figure 10 Photos taken from participants around UIC campus....................... 67
Figure 11. Photos taken from participants around NTU campus.................... 68
Figure 12. Photos taken from participants during the night............................ 73
CHAPTER 1
Introduction

Physical inactivity and obesity are major risk factors for chronic diseases, including cardiovascular disease, certain types of cancers, diabetes, and depression. Smoking, another risk factor for many chronic diseases, is a problem for only 20% of adults, but more than 70% of adults do not meet physical activity recommendations (USDHHS, 2000), and 34.9% of people were obese or overweight by 2012 (Ogden, Carroll, Kit and Flegal, 2014). The estimated annual medical cost of obesity in the U.S. was $147 billion in 2008 U.S. dollars and the medical costs for people who are obese were $1,429 higher per year than those of normal weight (Jindal et al., 2012).

Sedentary behaviors among children and young people are particularly troubling (Flegal et al., 2012; Rodríguez et al., 2012). Over 90% of American and Canadian youth do not meet the public health guideline of 60 minutes of daily moderate-to-vigorous physical activity (Troiano et al., 2008; Laxer & Ian, 2013). Habitual inactivity prevents young adults from developing healthy patterns of physical activity that will continue with them into adulthood. Inactivity in young people thus poses a critical, life-long health risk.

One promising approach to counter the trend of declining physical activity is to promote outdoor physical activity from active travel for commuting and recreation (Guell et al., 2012). A growing body of evidence indicates that certain environmental characteristics promote active lifestyles in general or promote specific kinds of physical activity such as walking or biking (Ogilvie et al., 2008). For example, people who live in neighborhoods which are close to stores, recreational places, green spaces, playgrounds, walkways, and bike paths participate in more active travel and have a lower body weight (Carroll-Scott et al., 2013; Duncan et al., 2010; Zenk et al., 2011).

Although a number of studies have investigated the built environment's ability to promote an active lifestyle, other studies have reported inconsistent and sometimes contradictory results. Van Cauwenberg et al. 2011 reviewed 31 papers about the relationship between the physical environment and physical activity and found weak or inconsistent relationships between walking for transportation and environmental features such as presence of walking facilities, traffic, crime-related safety, and aesthetics. In addition, neighborhood population density and street design have been associated with
active travel in some studies (Cradock et al., 2009; Rodríguez et al., 2012) but have had an inconsistent or weak correlation to physical activity in other studies (Evenson et al., 2010). Oyeyemi et al. (2013) reported inconsistent associations between street connectivity and physical activity outcomes compared to their previous studies. These inconsistencies and contradictions stem partly from the challenges of collecting valid data regarding the environmental features and of the locations where people actually travel (Troped et al., 2010).

Three challenges prevent us from understanding the strength of the association between environmental features and travel behaviors, and thereby limit the usefulness of research results in guiding urban design and planning. Below, I provide a brief overview of each of these challenges.

**Challenges in travel data collection**

Researchers need to know (a) the exact locations where people travel, (b) the time they travel, and (c) which travel mode they use (running, walking, biking, car, or transit). Unfortunately current ways of acquiring travel data seldom collect all three kinds of information at once, and many of these ways are cumbersome for both researchers and participants. I set out to create a tool that could address these challenges.

Monitoring active travel in individuals is challenging. Individual travel data are commonly acquired from interviews, observations, travel diaries, pedometers, and accelerometers. Self-reported travel logs can provide rich information about the travel attributes, such as travel purposes and travel modes. But collecting detailed and accurate data regarding travel routes, travel time, and activity intensity are nearly impossible to achieve using such travel logs (Milton et al., 2013). Pedometers and accelerometers can measure physical activity intensity, but they cannot contextualize the environment where the exercise takes place (Maddison & Ni Mhurchu, 2009). Although GPS devices are becoming cheaper, logistical challenges exist in carrying out experiments with a large sample size. Wearing additional devices such as an accelerometer or GPS is cumbersome for participants. Moreover, GPS systems do not provide direct information about the transportation modes. Hence, it would be useful to have a tool that could automatically classify travel modes based on data collected from GPS-enabled devices, and do so in a
way that does not burden research participants with technology. Chapter two of this dissertation takes up these challenges.

**Challenges in contextualizing travel behavior**

Another challenge for research examining healthy behavior is to contextualize people’s spatial behavior and activities (Kerr et al., 2011). Many studies limit the space being explored by examining predefined boundaries such as a census block (Riva et al., 2007; Koohsari et al., 2013), or through the use of various sized buffer areas around points of interest, such as home location (Feng et al., 2010). Yet, neither boundaries nor buffer areas around a point of interest capture the actual environment where physical activity takes place. In fact, most people’s physical activity happens in a broader space than their residential neighborhood, and the environmental features where physical activity takes place often differ from those in residential neighborhoods (Troped et al., 2010; Zenk et al., 2011). A recent review found that 90% of the studies in the area of physical activity and environment measured environmental characteristics of participants’ residential neighborhoods and that only 4% of studies examined nonresidential locations (Leal & Chaix, 2011). The difficulty of capturing individuals’ actual travel behavior environment using area-based attributes such as boundaries and buffers around predetermined points is called the Uncertain Geographic Context Problem (Kwan, 2012). To identify environmental characteristics that promote active behaviors, it is critical to overcome the Uncertain Geographic Context Problem (UGCP) by shifting use of boundaries and buffers oriented methods but gain actual location-based data. In doing so, we should be able to more reliably identify the environmental characteristics that are associated with active travel. Chapter three of this dissertation takes up these challenges.

**Challenges in understanding the perceptions of travelers**

Although it is well established that features of the built environment influence people’s active travel behavior, we do not know enough about how people perceive these environmental features. Most previous studies regarding people’s perceptions of the environmental features that influence active travel have used structured questionnaires. Although these questionnaires are widely used and have been validated, they ask participants to respond to a set of pre-determined features identified by researchers. That
is, participants have not been invited to identify and evaluate the features and attributes of
the built environment that they feel are most important in promoting active travel.

In addition, features or characteristics of the built environment in one neighborhood
or city may not be relevant in a different neighborhood or city. Allowing people to
identify and describe the features or attributes that make them more likely to engage in
active travel behavior in their own cities may provide richer details that can guide
planners, designers, and municipal officials who are trying to create more active cities.
Chapter four of this dissertation takes up these challenges.

**Dissertation organization**

This dissertation addresses these three challenges. First, in order to better measure an
individual’s travel behavior in a convenient and cost-effective manner, I developed a
smartphone application that simultaneously collects location, time, and accelerometer
data. In addition, I developed a method to automatically classify this data into different
travel modes. Chapter two introduces these methods and tests the accuracy of the travel
mode classification method. In chapter three, I use collected GPS data to contextualize
travel behavior within the local environment in two cities, Chicago and Singapore. I then
use mixed logistic regression to examine the relationships among characteristics of the
built environment and the travel modes that participants used as they moved through
these two cities. In chapter four, I use geo-tagged photo narratives to reveal Chicago and
Singapore residents’ perspectives about the characteristics of the environment that
support or discourage active travel. Chapter five concludes the dissertation by
summarizing the main findings, identifying the contributions and implications of this
work, and describing pathways and questions for future research.
References


CHAPTER 2
Does your smartphone know how you travel?

Understanding human transportation modes (walking, running, driving etc.) is critical in many areas of research (Biljecki et al., 2013). In transport planning and traffic management, understanding where and when people travel and their mode of travel (bike, walking) can help evaluate travel cost, predict public transport demand, identify spots where traffic congestion occurs, and optimize urban transport systems (Bohte & Maat, 2009). In urban design and public health, researchers must understand how and where people travel in order to understand the association between environmental features and people’s choice of travel modes (Sallis et al., 2006). In environmental epidemiology, being able to measure transportation modes is essential for studies concerning air pollution exposure because air pollutants vary significantly by location and travel modes (Wu et al., 2011). Thus, detecting travel mode, the location, and time of travel is crucial in many areas of research.

Most traditional methods to measure travel behavior do not collect all three kinds of data (location, time, and travel mode) at once. Publicly available datasets on travel behavior are usually aggregated datasets, static in both space and time (Jerrett et al., 2003). The self-reported travel log may be biased and hardly reflect specific travel time and paths. Pedometers and accelerometers also cannot provide the location where the exercise takes place (Maddison & Ni Mhurchu, 2009). Moreover, wearing additional devices such as accelerometer or GPS are cumbersome for participants. These limitations call for new approaches to collect rich information of travel behaviors.

In this context, smartphones may offer a new platform to overcome the challenges that previous measures face. In general, smartphones, which can be mobile sensors, are less cumbersome than GPS systems, and applications on smartphones can be customized to collect location, time, and types of movement. Hence, increasing number of studies start to look at automatically classifying travel mode using the combined sensors in smartphones. This study aims to enhance the collection of travel information by combining sensors in smartphones with an integrated travel mode classification algorithm. I address two specific questions: can we create a smartphone application that
accurately classifies travel modes? And, to what extent will combining a rule-based approach with a machine-learning approach improve the accuracy of the classification? Answering these questions will help researchers in a variety of fields know how to use smartphones to collect travel information in a fast, convenient, and accurate way.

**Background**

**Advancement of travel behavior measures**

Researchers need to know (a) the locations where people travel, (b) the time they travel, and (c) which travel mode they use (running, walking, car, or bike), but unfortunately current ways of acquiring travel mode data seldom collect all three kinds of data at once, and many of these ways are cumbersome for researchers and participants. Individual travel data are commonly acquired from interviews, observations, travel diaries, pedometers, and accelerometers. The self-reported travel log is able to provide rich information about the travel attributes, such as travel purposes and travel modes but the validity of data regarding travel routes, time, and activity intensity is often in question (Milton et al., 2013). Pedometers and accelerometers are capable of measuring physical activity intensity, but they cannot provide location information where the exercise takes place (Maddison & Ni Mhurchu, 2009). Such limitations cast some doubt on the findings from research that relies on one or more of these techniques.

The advance of detailed geographically disaggregated data and the advent of location-based techniques such as Global Positioning Systems (GPS) have great promise in this context (Higgs et al., 2012). GPS-based tracking presents many advantages over traditional methods, including high temporal-spatial resolution, minimum reporting burden from participants, and less effort to transcribe data to a digital format (Rainham et al., 2010; Wu et al., 2011). Although GPS devices are becoming cheaper, they can be logistically challenging for participants. In addition, GPS devices do not automatically classify the mode of travel.

The increasing prevalence of smartphone devices provides potential in measuring travel behavior (Franko & Tirrell, 2012). Smartphones with embedded sensors arm researchers with opportunities to deploy mobile sensing applications with an unprecedented efficiency and a much broader geographical context (Miluzzo, 2011). In a pilot study, Kerr et al. (2011) found that, in comparison with independent GPS devices,
assisted GPS devices in smartphones provide faster fixes and fewer participant dropouts. In addition, smartphones can alleviate the burden on participants of using unfamiliar research instruments (Kerr et al., 2011). Moreover, along with GPS, the smartphone’s embedded accelerometer provides an integrated sensor that can measure types of movement.

In spite of these practical features, GPS and smartphones do not directly retrieve information about the transportation modes as traditional travel diaries do. Hence, it would be helpful to have tools to automatically classify travel modes based on data collected from smartphones, and do so in a way that users are not burdened by the technology.

Efforts designed to automatically classify travel modes usually use two types of information: locational data captured by the GPS unit or vibration data obtained by an accelerometer or pedometer. Classification of travel mode based on GPS largely comes from the travel behavior in the field of transportation planning. Predictors derived from GPS unit include speed, acceleration, direction, and spatial accuracy. In general, the speed between two consecutive points is the main predictor in most classification (Bohte & Maat, 2009; Biljecki et al., 2013). Using these predictors from a GPS device to classify travel mode can achieve about 70% to 85% accuracy (Biljecki, 2013). Other studies, mostly from public health and kinesiology, use accelerometers to classify physical activity types. For instance, one study used a single tri-axial accelerometer placed on the waist to record the acceleration data. Five types of activity were classified with about 80% accuracy (Long et al., 2009).

**Gaps in previous methods**

*Sensor combination.* Although many studies have used the speed of GPS points to classify travel mode, using a speed cutoff on GPS points does not differentiate travel modes with high accuracy (Boruff et al., 2012). There are situations when biking, running, or even driving have a similar range of speed, which makes it very difficult to distinguish these different activities if only using speed as cutoff. For instance, a fast runner or a person driving slowly may have a similar speed to a biker. Wan & Lin (2013) pointed out the necessity to integrate accelerometer-derived activity information with GPS data to derive more comprehensive dimensions of travel behaviors. Using only an
accelerometer cannot provide geospatial information, which may limit the application’s ability to provide a location for the classified travel modes. Combining GPS and accelerometer data would be a powerful way to classify travel modes (Wan & Lin, 2013).

Data collection platform. Previous studies have developed methods to classify travel mode using GPS data; however, most of these studies collected data through a handheld GPS or a GPS data logger and some measurements were conducted at predefined routes (Chung & Shalaby, 2005; Wu et al., 2011; Gong et al., 2012). By comparison, smartphones have many advantages. Smartphones are in widespread use: in 2010, 4.6 billion mobile phones were in use worldwide, and an increasing number of them were technologically advanced smartphones (He et al., 2012). The number of smartphone applications available on Apple’s and Google’s web store has increased in recent years. Because of their widespread use and because people with smartphones carry their phones with them most of the time, smartphones provide significant opportunities to capture people’s travel behavior.

Although using smartphone applications to conduct research can be convenient and cost effective, few studies have examined the challenges of using smartphones as a tool to collect and classify travel information. In general, developing an application is a good way to measure transportation mode in terms of convenience and customization. However, approaches in travel mode classification are limited because Android applications run on different hardware platforms (Samsung, HTC, Motorola etc.) and different hardware platforms may have different sensor standards. The accelerometers’ sensitivity and sampling frequency are different from device to device, which may generate data at different scales and make it difficult to compare and classify.

Classification method. In addition to the challenge of different device platforms, research using smartphones has tried to find accurate classification methods. Data classification methods can be largely divided into procedural-based (rule-based) and machine learning-based approaches (Bolbol et al., 2012). Rule-based approaches usually classify data based on logical assumptions (Stopher et al., 2008), and use segments to classify GPS data. This method divides the sequentially acquired GPS points into relatively uniform segments (in terms of speed, orientation, and acceleration). The homogeneous segments are later aggregated together to form trips or journeys. Other
studies try to search for zero speed points and cut GPS points into segments delimited by zero speed points. The rule-based models may become problematic when processing data in certain occasions. For instance, even if the device is stationary, the collected GPS points may still “jump” around, which makes it hard to accurately identify the actual stationary location without using some arbitrary threshold. Additionally, rule-based classification requires an algorithm to take care of many special circumstances, which may reduce the classification accuracy in daily activity.

Machine learning based classifications usually make inferences based on learning from training data (Bolbol et al., 2012). Examples of these studies use Random Forest (Wu et al. 2011), Decision trees (Reddy et al., 2010), fuzzy expert system (Biljecki, 2013; Schüssler & Axhausen, 2009), and Support Vector Machines (SVMs) (Bolbol et al., 2012). These methods do not require identifying segments, but a certain amount of data is needed to make an inference from the training data. For machine learning based classification, although it requires minimum human interference and usually makes predictions quickly after training the data, it cannot discriminate special cases (such as abnormal points or brief stops between two driving segments), which are very common in travel activity. It may be possible to produce better classification results by combining learning-based and rule-based classification, thus integrating the strengths from the two classification methods.

To sum up, classifying travel mode should use combination of GPS and accelerometers data. More studies should test the applicability of smartphones to integrate both sensors (GPS and accelerometers) in classification. In addition, machine learning and rule-based methods for classifying travel mode, when used separately, cannot classify complicated travel activity with a high level of accuracy. In this study, I seek to answer two questions: How accurate will the classification of travel mode be using smartphone applications and machine-learning methods? Will combining a rule-based approach with a machine-learning approach improve the accuracy? Without a suitable method that runs on a smartphone platform, we may lose the opportunity to employ smartphone devices to collect travel mode information in a fast, accurate, and convenient fashion.
Methods

Application, study site, and participants

I developed the smartphone application for our experiment using the Java Android Application Programming Interface (API). In this study, two sensors – the GPS receiver and accelerometer – in the smartphone were used to collect data for measures of spatiotemporal attributes of travel behaviors. The sampling rate for location updates was set at five seconds. In other words, the GPS receiver obtained a GPS fix about every five-second. For accelerometer, I use the sampling frequency SENSOR_DELAY_NORMAL, which is the default sampling rate for accelerometer.

The accelerometer in smart phones measures the acceleration of the device on three axes. The x-axis points in the cross direction (from left to right) of the device, the y-axis points in the longitudinal direction (from down to up) and the z-axis is orthogonal to the display of the device (Huang et al., 2010). I developed algorithms to generate a vibration index based on the projected acceleration to the gravity direction. This index can also simulate step with a relatively high accuracy. The GPS and accelerometer data were stored in a PostGIS database. A program was developed to plot the locational data on top of Google Map, which provided spatial context for manually assigning the travel modes (Figure 1).

![Figure 1 Travel route and travel mode plotted over the Google Map. Circular marker with number represents point clusters, which are used to avoid too many plotting burdens and to increase drawing performance. Purple is walking; red is in-vehicle; blue is biking; and green is running.](image)
The experiment sites were in Singapore and Chicago, and 121 students from four universities participated (University of Illinois at Chicago, University of Chicago, Nanyang Technology University, and National University of Singapore). Four days of data from each participant were used in the evaluation. Participants were trained to use the application and asked to run the application during their daily activities. Participants also recorded their daily activity in a log in which they wrote down the mode of travel and the duration of travel for each of their trips during the day.

**Minimizing GPS point drop-out and maximizing spatial accuracy**

A major difference between smartphone GPS and the dedicated GPS concerns the manner in which signals are processed. Most smartphones employ a server-side component for processing GPS signals, which is referred as Assisted GPS (A-GPS). Compared to a dedicated GPS device, A-GPS in smartphones is understood to provide faster fixes and fewer drop-outs (Kerr et al., 2011). However, smartphone A-GPS may be less accurate than those from dedicated GPS units (Zandbergen, 2009). In order to harness the advantage of smartphones and to improve the accuracy of travel behavior detection, I adopted a “current best estimate” approach from the Android location strategy (Android Developer, n.d.). I set the location to be updated every two seconds, and set the sampling window within which I selected the best coordinates to be updated every five seconds. This procedure eliminated a great number of jumping points, but unfortunately consumed a considerable amount of the smartphone’s power. In a common android device, the battery can sustain 6-8 hours of experiment. To overcome the power demand, we asked the participants to turn on the application when they were moving outdoors, and turn it off when indoors. Participants could also choose a portable battery as a reward for participation in the experiment. This battery served as an additional power source, doubling the running time of most smartphones.

**Reference travel mode**

The daily activity log was used as a reference to check the accuracy of our smartphone-derived travel mode classification. Participants were asked to click the upload button in the app to send their daily data to a dedicated server. In addition to the travel diaries, I collected GPS points that included location and time information. Each segment was manually assigned to one of the five activity modes: walking, running,
biking, in-vehicle, and stationary by comparing the accelerometer and GPS data with participant’s travel diaries and its location superimposed onto Google Map. All points within each segment were assigned to the same mode. Based on these methods, the reference travel modes were evaluated and stored in the database. I divided data into two groups, training and testing. The testing set included 195,555 records, which were used to test the accuracy of the model.

**Predicting travel mode**

Next, I predicted travel mode using the accelerometer and GPS data from the smartphones in order to compare this classification to the reference travel modes. To use accelerometer data, the first question is how to use 3-axial acceleration. Prior studies usually use 3-axial linear acceleration data to classify travel mode (Anguita et al., 2012). I first used the raw acceleration on X, Y, Z, and the magnitude of the acceleration on the three axes to classify travel modes, but found that these variables were not the best features to predict travel modes. Placing the device horizontally or vertically will generate different readings on X, Y, Z-axes even with the same activity. For instance, if the device is put horizontally and shaken up and down, reading on the Z-axis would have significant variation. However, if the device is put vertically and shaken in the same direction, the Y-axis would have a lot of variation. The magnitude of the acceleration may blur out the directional information.

In addition, different Android phones have varying sensitivity. I found that the acceleration projected in the gravity direction had a better influence on classification than the magnitude indicator. The acceleration projected to the gravity direction represents the up and down vibration of the phone regardless of the orientation of the phone. Ideally, we could extract the acceleration parallel to the gravity direction even if the device is not held horizontally (that is, we could remove the influence of the horizontal acceleration). In order to do this, I first applied the moving average low-pass filter to each individual axis to find the gravity direction represented by the vector of the device coordinate system (Equation 1).

**Equation 1:** \[ a_g = \sum_{i=-30}^{0} \cos(\pi i / 60) \, a_{(t+i)} \]

where \( a_g \) is the instantaneous gravity direction acceleration, and \( a_{(t+i)} \) are the previous points to the current point. In this study, I used the adjacent 30 points to extract the low
frequency portion, which represents the gravity direction. I then projected the instantaneous acceleration onto the gravity direction.

Equation 2: \( a_{gr} = |a| \cos \theta \)

where \( \theta \) is the angle between current acceleration direction and gravity direction and \( a_{gr} \) is a scalar, the acceleration projected to the gravity direction (Figure 2a). By doing so, I derived the critical value—instantaneous acceleration in the gravity direction—from the device coordinate system. In addition, different android devices have different acceleration sampling rates. In order to remove the intra-device difference, I divided the acceleration in gravity direction by the number of our sampling duration.

Another predictor I used in the classification of travel modes is the simulated step number. I used the acceleration in the gravity direction to derive step number. Each inflection point on the oscillation curves was recorded. Based on many experiments, I found that the step occurs at the place when the difference between local maximum and local minimum of the oscillation curves exceed 3.5 m/s\(^2\) and the frequency was in the range of 1 Hz to 2Hz. These criteria can effectively simulate steps (Figure 2b). For instance, Figure 2b demonstrated an oscillation wave. When the difference between local maximum and minimum acceleration exceeded 3.5 m/s\(^2\), which indicated the vibration along the gravity direction was significant enough, and when the frequency was in the range of 1 Hz to 2 Hz, one step was detected.

![Figure 2](image)

Figure 2 (a) Project instantaneous acceleration to the gravity direction. (b) The step detection demonstration, in which each red triangle represents a step.
**Data integration and model predictors**

For most studies using separate GPS units and accelerometer devices to collect locational and vibrational data, data integration requires considerable effort. Timestamps on both devices need to be synchronized and some scripts need to be developed to generate the integrated data. In our application, GPS data and accelerometer data were intrinsically integrated when the sensors collected data. Each sampling record comprised a GPS coordinate (latitude and longitude), GPS accuracy, a timestamp, instantaneous speed, acceleration in the gravity direction, the angle between two consecutive segments, the distance difference between two consecutive line segments, the average estimated step, and the average acceleration on three axes. The estimated step per second was the number of steps between two adjacent points divided by the time between the measurements. Angle between two adjacent line segments was a supplementary predictor to account for jumping GPS points when the smartphone was stationary. Thus, the angle between two adjacent line segments in the stationary mode was usually more acute than the angle in other modes. GPS accuracy for each point was recorded by the Android device to evaluate the quality of the point. These variables were then evaluated in the classification models.

**Classification approach**

I selected Random Forest (RF) and Support Vector Machine (SVM) algorithms to classify travel mode. Each of these algorithms uses different strategies, and thus I wanted to see which algorithm would more accurately classify travel mode when compared to the travel diaries. In this study, I used R, a statistical computing and data analysis software package, coupled with Java, to do the classification. Data were pulled from the server using Java and sent to R to carry out the classification via rJava, which is a low-level R to Java interface.

Random forest is a learning method that generates multiple decision trees in the training process to predict the dependent variable in order to classify data. To classify a new case, random forest puts the variables into each of the trees in the forest. Each tree produces a classification result. The forest then chooses the classification with the most votes as the final classification. Random forest usually works better than the general
decision tree, and is considered one of the most accurate general-purpose learning techniques available (Zheng et al., 2008). Random forest uses a bagging approach, which reduces the probability of overfitting. Random forest is also relatively fast to train. In random forest, the model does not require original data to be shifted or scaled as gradient descent based methods (for instance: SVM or neural network) do. In the RF model, I grew 100 trees with 3 variables randomly sampled as candidates at each split. The sampling cases are done with replacement. I used the randomForest Package to conduct the analysis.

In contrast to Random Forest, which uses tree based structure to classify data, Support Vector Machine is another state-of-the-art algorithm, which looks for optimal decision planes that maximize the margin of the training data. SVM is a kernel-based algorithm. Different kernel functions transform the predictors into a higher dimensional space where more complicated patterns can be differentiated. The transformation will make data from a linearly inseparable space to a space linearly separable. The quality of generalization and ease of training of a large dataset with many predictors are the main advantages of SVM. I used SVM function in the e1071 package with the linear kernel function to train the data.

I also wanted to see how machine-learning (RF and SVM) approaches might compare to classification if both machine-learning and rule-based techniques were used. Will additional rule-based techniques enhance the accuracy of the classification? After classification using machine-learning techniques, I used rule-based techniques to regulate the results and eliminate some erroneous classifications that fail to recognize the continuity of travel mode. The rule-based techniques included a smoothing filter, short segment removal, and transitional segment adjustment.

Some very short travel segments were misclassified as walking in between two long in-vehicle modes. Similarly, some points embedded in walking segments were misclassified as biking. The misclassification may be caused by a sudden change of speed or acceleration, or because of abnormal records for either GPS or accelerometer. Hence, I first applied a smoothing filter with a window of 5 sequential points on the classified data. Most discrete erroneous classification can be eliminated in this way.
Second, if some short segments are less than 10 points or span less than one minute, and segments on each side of the segment have the same mode, these small segments were merged into the adjacent long segments. In the machine-learning method, these points in the middle would be classified as walking or stationary, but in fact they should be in-vehicle. So I used a rule-based technique to adjust these segments.

Third, based on our pre-tests, most classification errors occur at the transition from one mode to another. For instance, walking usually connects the change from in-vehicle mode to stationary mode. If the walking distance is short, the first two constraints may filter it out. Thus, the third rule-based technique was to detect the points around the transitional points and correct those points based on accelerometer readings. The overall method flow was shown in Figure 3.

![Figure 3](image)

Figure 3 The workflow of this study.

Accurate classification requires consideration on variable selection. Too many variables put in the model may reduce the performance of the classification model. Hence, I used some variable selection methods to find the necessary variable inputs. I first inputted all the variables into the model. The variable importance index was calculated for each input variable. Final variables were selected by comparing the variables in the index generated from the model diagnostics. After comparing the
variables, only the simulated step per second, speed, projected acceleration in the gravity direction, angles between two adjacent line segments were used in the prediction model. The rest of the variables were removed from the model because of low contribution or high correlations with other variables that remained in the model. Speed and simulated steps (Vibration) were the two variables that best distinguish different travel modes. Figure 4 shows the speed and vibration signal spectrums. As can be seen in Figure 4, although different modes had speed and vibration overlaps, there were some observable patterns among the four different modes. In-vehicle mode displayed the highest speed and lowest vibration. Running showed the highest vibration and the speed was between biking and walking. Biking generally had higher speed than running, but less vibration than walking. By using these variables we were able to classify the different travel modes.
Figure 4 Speed and vibration spectrums for walking, biking, running, and in vehicle.
Results

How accurate will the classification of travel modes be using smartphone applications and different classification methods? Will combining a rule-based approach with a machine-learning approach improve the accuracy? Analyses examining these questions are presented here. I compared the classification results with participant’s travel diaries and its location superimposed onto Google Map. The points with the pre-labeled modes were called reference group, and I compared the reference group with the automatically classified modes to evaluate the accuracy of the automatic classification.

To evaluate accuracy, I calculated the sensitivity (recall), specificity, and precision indices to reflect the model accuracy. The sensitivity and specificity are statistical measures of classification performance, which were widely used to evaluate classification models. Sensitivity measures the proportion of accurate positive estimations to the summation of accurate positive and inaccurate negative estimations, indicating how well the model can identify a specific travel mode (the percentage of real walking that is correctly identified as walking among all the real walking). Specificity is true negative divided by the summation of true negative and false positive, indicating how well a test avoids false alarms (the percentage of all travels modes but walking that are correctly not identified as walking). Precision is the division between true positive and the summation of true positive and false positive, reflecting how many of the positively classified modes were relevant (the percentage of real walking that are correctly identified as walking among all the classified walking). The calculation was conducted using the R caret statistical software package. I listed the equations to calculate these indicators here:

Equation 3: Sensitivity=TP/(TP + FN)
Equation 4: Specificity=TN/(FP + TN)
Equation 5: Precision=TP/(TP+FP)

In each equation, T is the true or right classification. F is the false or wrong classification. P is positive or identified as certain category. N is negative or rejected as certain category.

I applied this approach to evaluate the classification accuracy for (a) machine learning based method and (b) combined machine learning and rule-based methods. The following section presents the results for each of these groups.
Accuracy of learning based method

How accurate are the classifications of travel modes based on machine-learning methods? To answer this question, I compared the auto-classified modes using both RF and SVM with the reference travel modes. I then used sensitivity, specificity, and precision as the three variables to evaluate the model. Table 1 shows the classification results from the classification using the machine learning method. Both classifiers identified most of travel modes well. Both models predicted stationary and in-vehicle travel with the highest accuracy (for random forest, stationary: sensitivity = 93.3%, specificity = 92.8%, and precision = 86.7%; in-vehicle: sensitivity = 94.8%, specificity = 96.7%, and precision = 87.1%). Both models were less accurate for biking, running, and walking classification. In-vehicle mode was prone to be classified as biking or stationary when the participant had been in a slow-moving vehicle. Most misclassification happened in the transitional segments between two travel modes. I therefore used rule-based classification to see if we could improve the accuracy of the results. The following section shows the results for the combined model.

Table 1 Results of machine learning-based classification accuracy. Each row represents classification outcomes and each column represents the real condition. Numbers on the diagonal line represent the correctly assigned modes.

<table>
<thead>
<tr>
<th>Method</th>
<th>Stationary</th>
<th>Walk</th>
<th>Run</th>
<th>Bike</th>
<th>In-Vehicle</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>S</td>
<td>60583</td>
<td>4912</td>
<td>1381</td>
<td>955</td>
<td>2009</td>
<td>0.933</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>3593</td>
<td>32479</td>
<td>4092</td>
<td>792</td>
<td>288</td>
<td>0.785</td>
<td>0.945</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>181</td>
<td>1297</td>
<td>8326</td>
<td>643</td>
<td>167</td>
<td>0.546</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>371</td>
<td>1714</td>
<td>460</td>
<td>7630</td>
<td>442</td>
<td>0.658</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>3394</td>
<td>1104</td>
<td>1049</td>
<td>2596</td>
<td>55097</td>
<td>0.948</td>
<td>0.967</td>
</tr>
<tr>
<td>SVM</td>
<td>S</td>
<td>63722</td>
<td>3634</td>
<td>1334</td>
<td>1691</td>
<td>2393</td>
<td>0.935</td>
<td>0.929</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>3626</td>
<td>34651</td>
<td>5999</td>
<td>1092</td>
<td>197</td>
<td>0.835</td>
<td>0.929</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>141</td>
<td>1081</td>
<td>6476</td>
<td>507</td>
<td>134</td>
<td>0.423</td>
<td>0.990</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>285</td>
<td>1182</td>
<td>378</td>
<td>6840</td>
<td>268</td>
<td>0.542</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>348</td>
<td>958</td>
<td>1121</td>
<td>2486</td>
<td>55011</td>
<td>0.948</td>
<td>0.964</td>
</tr>
</tbody>
</table>

*RF stands for Random Forest and SVM stands for Support Vector Machine

Accuracy of the combined method

To what extent can the accuracy of the classification be improved by combining the machine-learning approach with a rule-based approach? To answer this question, I applied the rule-based model (a smoothing filter, transitional segment adjustment, and short segment removal) to adjust the results from the machine-learning based methods. Results of the classification that grew from this procedure were better than the classification without the rule-based adjustment. The adjustment took into account the continuity (e.g. one GPS point following 10 points that were classified as walking is more likely to be classified as walking as
well) to filter out the most discrete misclassification, and adjusting transitional points (e.g. five GPS points following 10 points classified as in-vehicle and following 10 points classified as stationary) based on actual accelerometer readings. By using the combined method, classification errors were reduced for most classifications. For random forest, on average, rule-based adjustment increased the accuracy by 6.4%, with an obvious increase in walking, running, and biking by 5.7%, 4.2%, and 12.0% respectively. SVM performed similarly to random forest after using the rule-based algorithm (Table 2).

<table>
<thead>
<tr>
<th>Method</th>
<th>S</th>
<th>W</th>
<th>R</th>
<th>B</th>
<th>IV</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>S</td>
<td>64551</td>
<td>2038</td>
<td>1288</td>
<td>823</td>
<td>2387</td>
<td>0.948</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>3091</td>
<td>36746</td>
<td>3144</td>
<td>423</td>
<td>116</td>
<td>0.885</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>172</td>
<td>881</td>
<td>9513</td>
<td>755</td>
<td>199</td>
<td>0.621</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>144</td>
<td>928</td>
<td>398</td>
<td>8673</td>
<td>197</td>
<td>0.687</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>164</td>
<td>913</td>
<td>965</td>
<td>1942</td>
<td>55104</td>
<td>0.950</td>
<td>0.971</td>
</tr>
<tr>
<td>SVM</td>
<td>S</td>
<td>64583</td>
<td>2035</td>
<td>1307</td>
<td>1400</td>
<td>2656</td>
<td>0.948</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>3081</td>
<td>37007</td>
<td>3487</td>
<td>431</td>
<td>100</td>
<td>0.892</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>175</td>
<td>790</td>
<td>9190</td>
<td>704</td>
<td>200</td>
<td>0.600</td>
<td>0.990</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>143</td>
<td>741</td>
<td>366</td>
<td>7906</td>
<td>234</td>
<td>0.627</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>140</td>
<td>933</td>
<td>958</td>
<td>2175</td>
<td>54813</td>
<td>0.945</td>
<td>0.969</td>
</tr>
</tbody>
</table>

*RF stands for Random Forest and SVM stands for Support Vector Machine

**Discussion**

In this research, I collected data from GPS and accelerometers embedded in Android smartphones and used these data to predict travel behavior. Results demonstrate that a smartphone device was capable of capturing data that can reveal how, where, and when we travel. The classification system I developed recognized various travel modes with accuracy greater than 80% for all the modes of travel in our study; most modes were estimated with accuracy greater than 85%. The rule-based adjustments refine the machine learning-based method and improve the accuracy by an average of 6.4%. SVM and random forest perform well in both cases.

In the paragraphs that follow, I discuss the strengths of this study and ideas for future research.

**Advantage of sensor combination**

One strength of our approach was to optimize both the GPS and accelerometer sensors to generate more stable and comprehensive predictors for classification. Prior research in travel
mode detection has primarily used GPS and derivative predictors such as speed, acceleration, direction, and acceleration ratio to predict travel modes. These GPS and derivative predictors are generally discriminating, but still have difficulty in distinguishing some modes with speed spectrum overlaps. Additionally, it is unavoidable to have some jumping points, which makes it difficult to distinguish stationary and slow speed walking. In contrast, a three axes accelerometer alone may differentiate different activity types, but it does not provide data in the spatial dimension. GPS devices can provide this data, which are critical in most location-based studies.

This study demonstrates a way to use smartphones to combine these two sensors in transportation studies. When applying both sensors in Android devices, I found that sensors with different specifications may generate data at different scales. For instance, I found that some android devices had a lower acceleration sampling rate than other devices even when both sampling rates were set at “SENSOR_DELAY_NORMAL” levels. To reduce the impact of different sampling rates on the acceleration predictor, acceleration was adjusted based on the sampling rate for different devices. For GPS fix, we used the current best estimate method to obtain GPS locations with less spatial error.

**Advantage of classification method combination**

Rule-based and machine learning methods have both been used in transportation studies to classify travel modes. The rule-based method was reported to be more flexible and easy to interpret, but required much more effort to tune to various conditions. Machine learning is easy to conduct and requires minimum human interference, but it may face potential over-fitting problems and can be less sensitive to special cases. In this study, I combined both methods to classify travel modes.

Both SVM and RF perform the classification process quickly and achieve relatively high accuracy. When I added the rule-based method to further adjust the classification results, I found the model had a better performance because rule-based adjustment takes into account the before-and-after relationship of points, eliminating most sparsely distributed points and adjusting transitional segments.

**Future research**

Future research can build upon this study in a number of ways. Currently, we have a project looking at the built environment characteristics that may promote active travel behavior. The approach presented in this study provides a way of monitoring the spatial distribution of active
travel behaviors. In epidemiological studies, many studies require the travel behavior distribution data to model air pollution exposure. With the increase use of smartphone devices, crowdsourcing based on combined sensors on mobile devices will provide invaluable data for such research. In our study, I experienced that power drainage is the downside of such kind of research. In future studies, better ways to save power or using dedicated devices, such as a smart watch, can be investigated.

Some studies classified travel mode based on both GPS and geographic information (Tsui & Shalaby, 2006; Gong et al., 2012). Chung and Shalaby (2005) used a map-matching algorithm based on geographic information systems (GIS) to identify transport links and a rule-based algorithm to identify transportation modes in Toronto. Their algorithm can detect walking, bicycling, busing, and driving a car at a relatively high accuracy. Combining GIS in the classification process provides additional information to distinguish travel patterns. However, it requires detailed transportation GIS data support, which may not be available for different places.

Conclusion

In this study, I classified travel modes based on a GPS and an accelerometer embedded in Android smartphones. I optimized the approaches to collect accurate locational data and vibration data as predictors for classification. These predictors discriminate travel modes. The combined rule-based and machine learning-based method was used to classify speed, step, acceleration in gravity direction, and segment angle for predictors into walking, biking, running, and in-vehicle travel modes. I evaluated the performance of these methods in predicting travel modes. Results reveal that a combined rule-based and machine-learning method can accurately classify different travel modes with accuracy over 80%.

The approach and results of this study can provide useful tools for research in public health, urban design, and transportation planning fields. Future work should use spatial reference data, such as OpenStreetMap (OSM) or other sources to provide more contextual information, such as the location of subway lines or bus stations in order to further classify travel into more detailed modes (i.e., bus, tram, driving). In future travel surveys, the method proposed in this paper can supplement travel diaries to facilitate the data collection and achieve more accurate results. Using the smartphone-based technology will advance studies in public health and transportation planning that require travel mode detection.
References


CHAPTER 3
Built environment associations with active travel

Engaging in regular physical activity has substantial health and social benefits for adolescents (Jackson et al., 2013; Janssen & LeBlanc 2010), especially in reducing the risk factors for chronic diseases such as type II diabetes and cardiovascular disease (Koohsari et al., 2013). For adolescents, many studies suggest engaging in physical activity maintains or even enhances academic performance (Dwyer et al., 2001). Despite the benefits of physical activity, overall engagement in an active lifestyle has declined over the last few decades (Brownson et al., 2005; Koohsari et al., 2013; Adams et al., 2013). Guthold et al. (2010) found that more than 80% of school children failed to meet the recommended physical activity level in 34 countries.

Fortunately, prior research suggests that increasing opportunities for active lifestyles by providing a well-designed built environment is a promising intervention (Ding & Gebel, 2012; Harris et al., 2013). A variety of studies suggest that characteristics of the built environment are associated with more active behaviors. For instance, easy access to urban green space is significantly associated with moderate to vigorous activity among young people (Lachowycz et al., 2012). Sustainable planning strategies focused on diverse, high-density, and activity-oriented neighborhoods help to promote active travel behavior (Wong et al., 2011).

Although a number of studies have investigated the built environment’s ability to promote an active lifestyle, other studies have reported inconsistent and sometimes contradictory results. For instance, neighborhood population density and street design have been associated with active travel in some studies (Cradock et al., 2009; Rodríguez et al., 2012) but have had an inconsistent or weak correlation to the physical activity in other studies (Evenson et al., 2010). Oyeyemi et al. (2013) reported inconsistent associations between street connectivity and physical activity outcomes compared to their previous studies. These inconsistencies and contradictions partly stem from the challenge of collecting valid and reliable data regarding the environmental features of the locations where people actually travel (Troped et al., 2010). For instance, many studies assume activity happens within a predefined area such as a census tract or residence buffer, and associate the travel behaviors with the environmental features in those predefined units (Oakes et al., 2007; Tsunoda et al., 2012). In this way, researchers are unable to capture the much broader context where physical activity actually occurs. These challenges prevent us from
understanding the strength of the association between environmental features and travel behaviors, and thereby limit the potential to use research results to guide evidenced-based urban planning.

In this study, I use smartphones to track people’s active travel behaviors and link these travel behaviors with objectively measured environment characteristics where active travel actually takes place. I begin by reviewing the gaps in our knowledge and then pose several questions that this study addresses.

**Background**

A variety of studies suggest that characteristics of the built environment are associated with more physical activity and active travel. Pikora, Giles-Corti, Bull, Jamrozik, and Donovan (2003) have described a conceptual framework in which the physical environment can promote physical activity through four aspects: by providing destinations, aesthetics features, functionality, and safety. Destination refers to the availability of service or commercial facilities, such as food, grocery, and hospitals. Aesthetics refers to characteristics that might be considered visually or experientially appealing, such as presence of trees or greenness. Functionality relates to the structural aspects of the local environment that enable people to get around easily or function well in the location, such as the population density, number of intersections, and street design. Safety reflects the sense of security a person has in a place.

A variety of studies have measured environmental features using items in Pikora’s framework. For instance, easy access to urban parks was significantly associated with active transportation and leisure time walking (Veitch et al., 2013). Diverse and dense neighborhoods were associated with more active travel (Wong et al., 2011). Fear of crime was found to be related to lower levels of outdoor recreation (Shinew et al., 2013).

Based on the results of previous studies and the availability of the data from Chicago and Singapore, I selected a group of environmental features under Pikora’s framework to be examined in this study. For destination, I selected restaurants, stores, bus stops, banks, hospitals, and parks, which were all common destinations in previous work (McCormack et al., 2008; Witten et al., 2011). For functionality, building density and road connectivity have been used as predictive variables (Frank et al., 2005). For aesthetics, greenness has been the most frequently used indicator associated with outdoor activity (Coombes et al., 2010). For safety, crime incidences have been widely used to predict outdoor recreation (Shinew et al., 2013).
In fact, many previous studies have used items under Pikora’s framework to examine the relationship between environmental features and travel behaviors. However, inconsistent or even contradictory results were generated from previous studies and the inconsistency is partly due to the method designs (Troped et al., 2010; Rodríguez et al., 2012).

Many studies limit the space examined by using predefined boundaries such as census block (Riva et al., 2007; Koohsari et al., 2013), or various sized buffer areas around points of interest, such as home location (Feng et al., 2010). Yet, neither boundaries nor buffer areas around a point of interest capture the actual environment where activity takes place. In fact, most people’s physical activity happens in a broader space than their residential neighborhood, and the environmental features where physical activity takes place often differ from those in residential neighborhoods (Troped et al., 2010; Zenk et al., 2011). This mismatch between methods to measure environmental features that use boundaries or buffers around predetermined points and the actual travel behavior of individuals has been called the Uncertain Geographic Context Problem (Kwan, 2012b). To identify environmental characteristics that promote active behaviors, it is critical to overcome the Uncertain Geographic Context Problem by switching from the use of boundaries and buffers to actual location-based data. In doing so, we should be able to more reliably identify the environmental characteristics that are associated with active travel.

The inconsistencies and contradictions in previous studies might also stem from the ways in which activity data were collected. In the vast majority of previous studies, activity data is obtained from self-reports, which are subject to time and location inaccuracy. The recent use of global positioning systems (GPS), which record locational information over time, has enabled researchers to more accurately contextualize travel behavior (Troped et al., 2010; Jones et al., 2009; Rodríguez et al., 2012). However, GPS units alone cannot detect activity type, so some studies use accelerometers in addition to GPS units to examine the vigor of the activity and thereby classify travel type. However, using a GPS device and an accelerometer can be cumbersome for participants, and synchronizing GPS data and accelerometer data requires a lot of extra effort.

In sum, obtaining more valid and reliable knowledge regarding the association of urban design features and active travel requires that we overcome the Uncertain Geographic Context Problem, and make technical advances in our ability to track types of travel through space and time.
To overcome the Uncertain Geographic Context Problem (UGCP), this study makes use of smartphones equipped with global positioning system (GPS) receivers and accelerometers that measure people’s movements through space and time. To objectively measure environmental features, I use GIS data to calculate scores for items related to destinations, aesthetics, functionality, and safety. This study addresses three specific questions:

1. What are the overall travel patterns in two cities, Chicago and Singapore?
2. To what extent are environmental features associated with travel behaviors in the two cities?
3. Within each city, do the associations between environmental features and behaviors vary by neighborhood?

Methods

Sites

This study examined the environmental context of the physical activity behavior of college students in two cities, Chicago and Singapore. Singapore is an island country in Southeast Asia and one of the world’s leading commercial hubs. It is renowned for its urban design and transport planning and is regarded as one of the most walkable cities in the world (Sanyal, 2009). In addition, its sustainable transport planning has been suggested as a model for Asian nations and rapidly growing cities (Han, 2010; Olszewski, 2007).

Chicago is the third most populous city in the United States (U.S. Census Bureau, 2010). Chicago has a national reputation for bicycling, a reputation achieved by investment in bicycling infrastructure. Chicago currently has more than 200 miles of on-street, protected, buffered, and shared bike lanes (Chicago Department of Transportation, 2013). Comparing these two cities of similar size but with different sustainable transportation strategies will explore the extent to which differences in the built environment and urban infrastructure impact travel behaviors.

Participants

Participants’ data were collected from June 2013 to October 2013. The study was advertised through post fliers, email invitations, active recruitment on campus, and Facebook messages. A total of 142 individuals university students who had no difficulty in conducting daily activity and who owned smartphones, were recruited from four universities (i.e. University of Illinois at Chicago (UIC), University of Chicago (UC), National University of Singapore (NUS), Nanyang
Technology University (NTU)) in two cities. To control for the potential race/ethnic influence, I only recruited Asian students in Chicago.

Out of the 142 participants, 11 participants did not complete either the travel diary form or the follow-up questionnaire. Another 10 participants did not have a sufficient amount of GPS data that matched with the travel diary. Therefore, the final study sample included 121 participants, with 54% female in Chicago and 31% female in Singapore. There was an average of 3.8 valid days (range: 3-5 days) of GPS data for all participants. The average body mass index (BMI) of the participants is within the normal spectrum (mean= 22.3 in Chicago and mean=21.0 in Singapore). About 8% of participants from Chicago own a car while no participants from Singapore own one. Participants’ families were mostly highly educated, with more than 80% possessing a college degree. In addition to demographic variables, this study also controlled for participants’ attitude of active travel and peer/family influence. Participants were asked to rate on a scale from 1-5 with regard to these two aspects. Most participants were aware of the benefits of physical activity, but about half of the participants reported their friends or family members were relatively inactive (Table 3).
Table 3 Demographic and social comfunding variables of the participants in two cities

<table>
<thead>
<tr>
<th></th>
<th>Chicago</th>
<th></th>
<th></th>
<th>Singapore</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>Mean</td>
<td>SD</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>29</td>
<td>46.0</td>
<td>40.0</td>
<td>2.7</td>
<td>40.0</td>
<td>69.0</td>
</tr>
<tr>
<td>Female</td>
<td>34</td>
<td>54.0</td>
<td>18.0</td>
<td>1.9</td>
<td>31.0</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td>22.3</td>
<td>2.7</td>
<td>21.0</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Own a car</td>
<td>5</td>
<td>8.0</td>
<td>0</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>28</td>
<td>52.1</td>
<td>31</td>
<td>53.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent education level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctoral or equivalent level</td>
<td>3</td>
<td>4.8</td>
<td>6</td>
<td>10.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master's or equivalent level</td>
<td>21</td>
<td>33.3</td>
<td>21</td>
<td>36.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor's or equivalent level</td>
<td>27</td>
<td>42.9</td>
<td>25</td>
<td>43.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between secondary level and university</td>
<td>9</td>
<td>14.3</td>
<td>3</td>
<td>5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary school</td>
<td>2</td>
<td>3.2</td>
<td>2</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school only (or less)</td>
<td>1</td>
<td>1.6</td>
<td>1</td>
<td>1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel healthier if I walk/bike regularly</td>
<td></td>
<td>4.0</td>
<td>1.4</td>
<td>3.7</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>I feel relaxed if I walk/bike regularly</td>
<td></td>
<td>3.7</td>
<td>1.3</td>
<td>3.9</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>I make less air pollution if I walk/bike</td>
<td></td>
<td>3.9</td>
<td>1.3</td>
<td>3.7</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>My friends always walk/bike</td>
<td></td>
<td>3.5</td>
<td>1.5</td>
<td>2.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>My parents always walk/bike</td>
<td></td>
<td>2.7</td>
<td>1.1</td>
<td>3.3</td>
<td>1.3</td>
<td></td>
</tr>
</tbody>
</table>
**Procedure**

In order to collect travel data, I developed an android application. This application used the embedded GPS and accelerometer to track and classify users’ travel behavior. The participants downloaded and installed this application from Google Play. The experiment was conducted throughout a four-day time window, including weekdays and weekends. On these four days, participants were asked to turn on the application and let it run in the background while moving outdoors. To save battery power, users were instructed to turn off the application when they were indoors. The application tracked location through the GPS unit and vibration through the accelerometer. Based on the GPS and accelerometer data, I used machine-learning and developed a rule-based model to classify the data into four travel modes (walking, biking, running, and in-vehicle). The classification accuracy was above 80% for the four modes. Participants were also asked to document their travel modes, time range, and travel purposes using a detailed travel diary. This diary was used to validate and complement the auto-classification. Upon the completion of the study, smartphone data was uploaded to a dedicated server and stored in PostGIS database, and the travel diary was returned to the researcher.

GPS data was downloaded from the server and plotted against Google Map to check for accuracy. I first cleaned GPS points by removing erroneous points and jumping points. Because this study focused on travel behaviors, static point clusters were removed from the dataset. GPS points were time-stamped. Each point was assigned a travel purpose based on the descriptions provided in the travel diary and the point feature interpretation against the map.

Participants’ demographics (shown in Table 3) were collected as part of the experiment. At the end of the study, participants filled out an online questionnaire which included items regarding gender, families' highest education level, weight and height (in order to calculate BMI), three items about physical activity attitudes (Health: I feel healthier if I walk/bike regularly; Relaxation: I feel relaxed if I walk/bike regularly; Environment: I make less air pollution if I walk/bike), and two items about family/peer influence. This demographic and attitude information was based on previous studies (Hino et al., 2011; Fermino et al., 2013; Reis et al., 2013; Rodríguez et al., 2012) and was used to examine potential confounding variables in the models. Each participant was given 20 dollars or a portable smartphone charger for their participation in this study.
**Built environment measures**

I measured environmental features including functionality (road connectivity and building density), destination, safety, and greenness. GIS data for Chicago were mostly available from the Chicago City Innovation and Technology Department and Data Portal for the City of Chicago. For Singapore, data sets were obtained from OpenStreetMap, Google Place API, and governmental web portals such as National Parks Board, Singapore Police Force, and data.gov.sg. Road network data were gathered from the TIGER database and OpenStreetMap. Landsat 8 data were used to extract green space in the two cities. Two sets of images for Chicago and Singapore respectively were captured on May 24, 2013 and July 27, 2013. Both images have good quality with little cloud coverage.

Destinations were measured for restaurants, stores, bus stops, banks, hospitals, and parks. Street connectivity was measured through the number of street intersections. Building density was measured by dividing the area of the building footprint by the total area. The crime index was calculated based on crime point incidences collected in the two cities. The amount of green space was derived from the supervised classification of Landsat 8 imagery. Greenness included canopy, grass, and shrubs. I merged these three types into one green space category. Both images were enhanced by pan-sharpening techniques to increase the multispectral resolution to 15m. The maximum likelihood classifier was used in land use classification. Classification was conducted using ERDAS IMAGINE.

Built environment indicators were represented by a set of 50m X 50m raster layers. For point-based data such as restaurant, road intersection, and crime, cell value represents the number of features within the cell. A 500m square buffer was selected to represent the focal cell neighborhood. Values of the focal cell for the point-based features were calculated by aggregating the adjacent cell values in the focal neighborhood. For example, to calculate restaurant score in one cell, the restaurant counts within 500m of the grid cell were added up as the score for the central cell’s value. I also used a Gaussian distance decay function, implying that counts for restaurants far away from the central cell had less weight in the sum calculation. For density and greenness (polygon features), the value of the central cell was calculated by averaging values of all cells in the focal neighborhood. I processed the computation using R and ArcGIS software package.
Results

**Overall travel pattern**

What are the overall travel patterns in the two cities? To address this question, I present two-dimensional maps of the travel paths and locations in Chicago and Singapore (Figure 5). These maps result from plotting the GPS points obtained from the smartphones on the maps of each city. In Chicago, the University of Illinois at Chicago (UIC) participants had a wider activity space spreading from the campus to surrounding areas than the University of Chicago (UC) participants, who were tightly clustered around their campus. In general, in Chicago, more activities spread to the north and west. Few activities extended to the south.

In Singapore, the patterns were slightly different. Unlike the grid movement pattern found in Chicago, activity paths in Singapore were more irregular. Participants from both universities in Singapore were active mostly in western and central Singapore. In addition to the two campuses, some clusters of activity occurred around Singapore River, Boon Lay, Orchard Road, and Clarke Quay. Participants from NTU had a little wider activity space than individuals from NUS. We wanted to know how the travel distribution in two cities was associated with the built environment characteristics. The next section addresses this question.

Figure 5 Participants’ travel behavior from four universities. Locations of each university are identified as red points. Green and blue colors represent participants’ travel paths from different universities.
**Influence of environmental features on travel behavior**

To what extent did the built environment shape travel behavior? To address this question, I conducted mixed logistic regression analyses (Tables 4 -7). For the statistical analyses, the unit of observation is the GPS point. A random intercept model using the generalized linear mixed models with a multinomial logit link function was used to test the statistical significance. This model appropriately accounted for the clustered data structure for each study participant (Rodríguez et al. 2012). The regression analyses tested whether the measured built environment attributes including greenness, destinations, functionality, and crime four aspects predicted more presence of active travel behaviors as opposed to in-vehicle. The analyses controlled for participants’ gender, BMI, car ownership (for Chicago), parents’ education level, and education program. The model also took into account social support, including family and friends’ participation in active travel and individual activity attitude including health, relax, and environment concerns. In the models, travel behavior was partitioned into two types, utilitarian or recreational. Recreational travel refers to activity undertaken for discretionary reasons such as relaxing or exercising in someone’s leisure time while utilitarian travels are undertaken to fulfill certain purpose such as to reach workplaces (Frank et al., 2003). Statistical analyses were conducted in statistical program R v. 2.15.3. I also used the constructed models to predict the walking versus driving environment in two cities.

The logistic regressions produced an odds ratio (OR) that I used to investigate the environmental features that were more or less related to active travel compared to in-vehicle travel. In these analyses, in-vehicle travel was treated as the reference group. The odds ratio represents the relative likelihood that active travel took place compared to in-vehicle travel. Thus, an OR of 1.5 indicates that there was a 50% greater likelihood that active travel took place rather than travel in a vehicle.

After exploring the indicators, values for access to hospital and access to banks were too low and had limited variability for the measured GPS points and thus were excluded from the regression model. Closeness to stores and restaurants had relatively high collinearity. Thus, the store variable was also excluded from the final model. Odds of active travel versus in-vehicle travel were treated as dependent variables and environmental features including access to parks, restaurants, and bus stops, road connectivity, building density, greenness, and crime distribution
were treated as independent variables. Below, I present results for each city focusing on features from greenness, destination, functionality, and crime.

**Chicago.** In Chicago, higher levels of greenness were consistently associated with more recreational walking (OR 1.36), biking (OR 1.13), and running (OR 1.30) versus in-vehicle travel. For utilitarian walking and biking, greenness was also associated with more active behavior, but the odds ratios were much lower. These results reinforce findings from recent studies focusing on the influence of green space on physical activity (Mytton et al., 2012; Schipperijn et al., 2013).

In terms of access to destinations, access to bus stops was significantly associated with utilitarian walking and biking. Access to parks was significantly associated with recreational running, but the odds ratio was close to one (OR 1.01), suggesting the influence is positive but not strong. Places with more restaurants in general are associated with more active travel versus in-vehicle travel. However, the influence is not particularly high (e.g. OR 1.03 for recreational running vs. in-vehicle).

For functionality, building density showed a negative influence on recreational travel behaviors, and a positive but not significant influence on utilitarian walking. High building density was more influential for utilitarian activities than recreational ones. Road connectivity did not significantly influence most modes of active travel.

Crime did not show a significant impact on most active travel behavior.

**Singapore.** In Singapore, biking occurred so infrequently that it was excluded from the regression model. Therefore, I analyzed running and walking versus in-vehicle activities for Singapore. Except for utilitarian running, greenness was significantly and positively related to walking and running (OR 1.03 and 1.14 for recreational walking and running). People tend to run more in a greener environment.

For destinations, access to restaurants was significantly and positively related to both recreational (OR 1.09) and utilitarian walking (OR 1.01). Easily accessible food courts in many neighborhoods may largely contribute to this influence. Access to park was not a strong predictor of active travel.

For functionality, a well-connected road system was an important facilitator for utilitarian walking (OR 1.03), running (OR 1.07), and recreational walking (OR 1.02). Building density
was significantly and negatively associated with recreational running. People tended to run for recreation in a less dense environment.

Crime in Singapore was not significantly related to any modes of active travel either.

To sum up the findings, greenness in both cities was strongly associated with more active travel. Access to destinations, especially to restaurants and bus stops was positively associated with active travel, but the association was relatively weak. Contrary to the previous findings that showed that high density of housing generates more active travel, I found that more recreational activities occurred in some of the less dense areas (around large green spaces or along the lakeshore). Crime was not significantly associated with active behaviors in either city.

Table 4 Adjusted odds ratio of recreational travel behaviors (four modes) associated with the built environment around each GPS point in Chicago

<table>
<thead>
<tr>
<th>Indicators</th>
<th>walking vs. in-vehicle</th>
<th>running vs. in-vehicle</th>
<th>biking vs. in-vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Access to park</td>
<td>0.97 0.96 0.97 *</td>
<td>1.01 1.00 1.01 *</td>
<td>1.51 0.82 2.81</td>
</tr>
<tr>
<td>Access to restaurant</td>
<td>1.20 0.99 1.36</td>
<td>1.03 1.03 1.04 *</td>
<td>0.62 0.46 0.83 **</td>
</tr>
<tr>
<td>Building density</td>
<td>0.72 0.62 0.90 **</td>
<td>0.70 0.38 1.30</td>
<td>0.80 0.66 0.93 **</td>
</tr>
<tr>
<td>Access to bus stop</td>
<td>1.12 1.09 1.56 **</td>
<td>1.10 1.01 1.19 **</td>
<td>0.33 0.24 0.46 **</td>
</tr>
<tr>
<td>Road connectivity</td>
<td>1.08 1.05 1.11 **</td>
<td>0.99 0.96 1.01</td>
<td>1.05 0.93 1.37</td>
</tr>
<tr>
<td>Crime distribution</td>
<td>1.00 1.00 1.01</td>
<td>0.94 0.94 0.95 *</td>
<td>1.23 1.00 1.52</td>
</tr>
<tr>
<td>Greenness</td>
<td>1.36 1.29 1.44 **</td>
<td>1.13 1.09 1.17 **</td>
<td>1.30 1.04 1.52 **</td>
</tr>
</tbody>
</table>

*Variable is significant at the 0.05 level (2-tailed); **Variable is significant at the 0.01 level (2-tailed).

Table 5 Adjusted odds ratio of utilitarian travel behaviors (four modes) associated with the built environment around each GPS point in Chicago

<table>
<thead>
<tr>
<th>Indicators</th>
<th>walking vs. in-vehicle</th>
<th>running vs. in-vehicle</th>
<th>biking vs. in-vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Access to park</td>
<td>1.00 1.00 1.00</td>
<td>1.01 1.00 1.01 *</td>
<td>1.01 1.00 1.01 *</td>
</tr>
<tr>
<td>Access to restaurant</td>
<td>1.00 1.00 1.00</td>
<td>0.93 0.92 0.94 *</td>
<td>1.01 1.01 1.02 *</td>
</tr>
<tr>
<td>Building density</td>
<td>1.08 0.81 1.47</td>
<td>1.15 0.42 3.13</td>
<td>0.80 0.40 1.61</td>
</tr>
<tr>
<td>Access to bus stop</td>
<td>1.13 1.10 1.17 **</td>
<td>1.33 1.05 2.15 *</td>
<td>1.13 1.04 1.23 *</td>
</tr>
<tr>
<td>Road connectivity</td>
<td>1.00 0.99 1.01</td>
<td>0.75 0.70 0.80 **</td>
<td>0.92 0.88 1.05</td>
</tr>
<tr>
<td>Crime distribution</td>
<td>1.00 1.00 1.00</td>
<td>1.02 1.01 1.03 *</td>
<td>1.02 0.91 1.02</td>
</tr>
<tr>
<td>Greenness</td>
<td>1.02 1.01 1.03 *</td>
<td>0.83 0.77 0.89 **</td>
<td>1.06 1.01 1.12 *</td>
</tr>
</tbody>
</table>

*Variable is significant at the 0.05 level (2-tailed); **Variable is significant at the 0.01 level (2-tailed).
Table 6 Adjusted odds ratio of recreational travel behaviors (in three modes) associated with the built environment around each GPS points in Singapore

<table>
<thead>
<tr>
<th>Indicators</th>
<th>walking vs. in-vehicle</th>
<th>running vs. in-vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95%CI</td>
<td>OR 95%CI</td>
</tr>
<tr>
<td></td>
<td>L  H  L  H</td>
<td>L  H  L  H</td>
</tr>
<tr>
<td>Access to park</td>
<td>1.04 0.93 1.16</td>
<td>0.77 0.55 1.08</td>
</tr>
<tr>
<td>Access to restaurant</td>
<td>1.09 1.07 1.11 *</td>
<td>1.03 1.00 1.05 *</td>
</tr>
<tr>
<td>Building density</td>
<td>1.00 1.00 1.00</td>
<td>0.95 0.93 0.96 *</td>
</tr>
<tr>
<td>Access to bus stop</td>
<td>1.09 1.03 1.15 *</td>
<td>0.92 0.83 1.03</td>
</tr>
<tr>
<td>Road connectivity</td>
<td>1.02 1.01 1.03 *</td>
<td>0.92 0.90 0.95 *</td>
</tr>
<tr>
<td>Crime distribution</td>
<td>1.00 0.91 1.11</td>
<td>1.18 0.93 1.49</td>
</tr>
<tr>
<td>Greenness</td>
<td>1.03 1.00 1.06 *</td>
<td>1.14 1.10 1.19 **</td>
</tr>
</tbody>
</table>
*Variable is significant at the 0.05 level (2-tailed); **Variable is significant at the 0.01 level (2-tailed).

Table 7 Adjusted odds ratio of utilitarian travel behaviors (in three modes) associated with the built environment around each GPS points in Singapore

<table>
<thead>
<tr>
<th>Indicators</th>
<th>walking vs. in-vehicle</th>
<th>running vs. in-vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95%CI</td>
<td>OR 95%CI</td>
</tr>
<tr>
<td></td>
<td>L  H  L  H</td>
<td>L  H  L  H</td>
</tr>
<tr>
<td>Access to park</td>
<td>0.97 0.93 1.01</td>
<td>0.96 0.69 1.34</td>
</tr>
<tr>
<td>Access to restaurant</td>
<td>1.01 1.00 1.01 *</td>
<td>0.82 0.72 1.00</td>
</tr>
<tr>
<td>Building density</td>
<td>1.01 1.01 1.02 *</td>
<td>0.93 0.91 0.95 *</td>
</tr>
<tr>
<td>Access to bus stop</td>
<td>0.89 0.87 1.11</td>
<td>1.07 0.94 1.21</td>
</tr>
<tr>
<td>Road connectivity</td>
<td>1.03 1.02 1.03 *</td>
<td>1.07 1.03 1.10 *</td>
</tr>
<tr>
<td>Crime distribution</td>
<td>0.95 0.91 1.01</td>
<td>0.76 0.51 1.13</td>
</tr>
<tr>
<td>Greenness</td>
<td>1.05 1.04 1.06 *</td>
<td>0.89 0.84 0.94 **</td>
</tr>
</tbody>
</table>
*Variable is significant at the 0.05 level (2-tailed); **Variable is significant at the 0.01 level (2-tailed).

**Influence of local variability**

Within each city, do the associations between environmental features and behaviors vary by neighborhood? To answer this question, we repeated the statistical procedures described above but broke down the analysis into two subgroups (one for each university in each city). Results demonstrate that the association between the built environment and travel behavior varied in different areas within the cities. For example, I compared recreational walking versus in-vehicle odds ratios around two universities in Chicago. Green space was associated with more recreational walking among students at UC than at UIC. The UC campus exhibited a significant odds ratio (OR = 1.39) while UIC campus had a positive but not significant correlation (OR = 1.01). This may be attributed to the closeness of the lakeshore park at UC campus (Table 8).
Closeness to restaurants predicted walking around UIC campus more than around UC because more participants from UIC tended to walk in downtown areas where businesses and restaurants were plentiful.

The relationship between density and active travel was also different between the Chicago campuses. You will remember that when I analyzed density for Chicago as a whole, the relationship to active travel was negative. But when I analyzed the two campuses separately, high building density around UIC became an important predictor of active travel while density was a negative predictor around the UC campus. These cross-neighborhood comparisons suggest that various combinations of features within the built environment interact with one another such that in some cases individual features are positively associated with active travel and in other cases they are not. It may be the interaction of these features is more important to study than the individual features by themselves.

Table 8 Comparison between the adjusted odds ratio of recreational walking associated with the built environment around each GPS point in UIC and UC.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>walking vs. in-vehicle for UC</th>
<th>walking vs. in-vehicle for UIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95%CI</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Access to park</td>
<td>0.93</td>
<td>0.67</td>
</tr>
<tr>
<td>Access to restaurant</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Building density</td>
<td>0.79</td>
<td>0.53</td>
</tr>
<tr>
<td>Access to bus stop</td>
<td>1.15</td>
<td>1.09</td>
</tr>
<tr>
<td>Road connectivity</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>Crime distribution</td>
<td>1.01</td>
<td>1.00</td>
</tr>
<tr>
<td>Greenness</td>
<td>1.39</td>
<td>1.02</td>
</tr>
</tbody>
</table>

*Variable is significant at the 0.05 level (2-tailed); **Variable is significant at the 0.01 level (2-tailed).

Discussion

This study investigated the association between built environment attributes and active travel behavior in Chicago and Singapore. I used smartphones to collect GPS and accelerometer data, and classified movements into four different travel modes. These travel modes were modeled against the objectively measured built environment characteristics.
Main findings

The overall travel pattern shows that UIC participants’ active travel behavior spread beyond campus to the surrounding areas, while the UC participants’ active travel remained mostly clustered around campus. In Singapore, activity paths were more irregular. Participants from both universities in Singapore were active mostly in the western and central portions of Singapore.

Greenness was consistently associated with more recreational active travel than in-vehicle travel in both cities. The odds ratio also indicated that greenness was a strong influential factor in active travel behavior. Destinations in general showed a positive relationship with active travel behavior, but the odds ratios suggest the relationship is relatively weak. In general, high building density was related to more utilitarian active travel but less recreational activity. For instance, I found that recreational running took place in environments with less building density in both cities. Crime did not predict active travel.

The association between the built environment and travel behavior varied in different neighborhoods within the cities. Greenness was associated with more recreational walking around UC surroundings than UIC surroundings. People walked for recreation more in places with higher building density near UIC while the opposite relationship was found in the UC areas.

In the paragraphs that follow, I describe the major contributions of this work, policy implications and possible interventions, and future research directions.

Contributions

There are two especially common methodological challenges in research that examine the effects of geographic variables on individual behaviors: the Modifiable Areal Unit Problem (MAUP) and the Uncertain Geographic Context Problem (UGCP) (Kwan, 2012a). The modifiable areal unit problem (MAUP) is a well-known methodological problem, occurs when different levels of area-based variables are used to compute environmental features, such as census tract, block groups, blocks, or school district (Openshaw, 1983). Using different area units may generate different or even contradictive results. Hence, in order to overcome the MAUP problem, most previous studies focused on deriving the best areal division or geographic scale for the data representation (Weiss et al., 2007). The UGCP problem, however, has only recently been identified and there is little record of it being considered in previous studies. The methods to test the optimal geographic scale for MAUP problem do not help much when dealing
with the UGCP problem (Kwan, 2012a). Addressing the UGCP problem requires precise geographic delineations of contextual units. This study demonstrated an approach to address the uncertain geographic context by using the actual GPS route and GIS data. By using GPS points, the measured environment was not a predefined areal unit (such as census tract), but an environment where travel actually happened. This approach addressed both MAUP and UGCP problems to a great extent.

The findings presented here confirm the notion that it is critical to measure travel behaviors with respect to the specific details of the surrounding environment. For instance, in this study, I found building density was not consistently related to active travel, as many other studies found (Hanibuchi et al., 2011). I found building density was not positively associated with recreational active travel, and that building density in different neighborhoods can have either positive or negative associations with active travel. At UIC campus, which is closer to downtown Chicago, building density was positively associated with recreational active travel. However, at UC, recreational activities actually happened away from the dense part of campus. Local environmental variability may shape people’s travel behavior. This indicates the importance of contextualizing spatial behavior and also warrants future study to take into account the UGCP and the possibility that the features of the built environment interact in various ways to promote or deter active travel.

In measuring travel behavior, most previous studies have combined GPS receiver and accelerometer data to measure both activity location and intensity. Synchronizing data from GPS receivers and accelerometers, however, requires multiple steps and risks losing data if timestamps do not match (Dessing et al., 2013; Dunton et al., 2014). Using smartphones to monitor travel behavior provides great potential for future healthy behavior studies. Broad use of smartphones and mobile apps enables convenient individual trip monitoring. In this study, I classified travel modes using data from a smartphone’s embedded accelerometer and GPS receiver. This method alleviates the burden of wearing additional devices for participants and avoids data loss because of data synchronization.

The comparison between two cities in this study also reveals interesting findings. First, I found that biking is more prevalent in Chicago than in Singapore. This implies the environmental features in Chicago are relatively more suitable for biking than in Singapore. Second, green environments in both cities were significantly associated with recreational active travel.
behaviors. This confirms previous research on the role of urban green space in promoting active lifestyles. Third, I observed that road connectivity is significantly associated with recreational walking in Singapore, while access to bus stops is more significantly related to different types of active travel behavior in Chicago.

I also found it is important to partition travel behavior data into different purposes. Built environment influences travel behavior differently for utilitarian and recreational activities. The analysis would not show a significant result if travel purposes were not differentiated.

**Policy implications**

The differences of travel behavior within each city suggest several policy implications. For biking, participants from UC biked more along the lakeshore than people from UIC (Figure 5), while people from UIC traveled more extensively in the business areas. The location of UIC is closer to the city center where more destinations and businesses are available. The UC campus is closer to Michigan Lake where a lot of green space is available. This study shows that different environments shape people’s activities differently. People tend to visit nearby destinations for recreation. Places with diverse destinations enrich active travel behaviors of the people nearby, while places with more green space or natural area pull people’s travel towards those directions. People tend to capitalize on the environment close to their neighborhood. Thus, planners should facilitate active travel in ways that are suited to the local environment. For instance, planners should consider building safe and well-connected trails to link major natural areas in neighborhoods close to natural spaces, while neighborhoods with few stores available may benefit from the addition of businesses.

The findings in this study suggest that the planning strategies in the two cities are working. Biking is more common in Chicago than in Singapore. The city planning department initiated several strategies to promote biking in Chicago. To enhance biking and the use of public transit, the government makes it easy to combine biking and transit travel. Bringing bikes on a train or bus is easy. To make biking possible, the planning department continues expanding biking facilities. A 645-mile network of biking trails and facilities is expected to be in place by 2020. In addition, planners offer more bike sharing programs where ridership is high, while focusing more on establishing biking infrastructure where ridership is currently lower (Chicago Department of Transportation, 2013). Results from this study indicate that these efforts are
working. Students are making good use of these facilities, and cities like Chicago should continue to fund these efforts.

In Singapore, ample attention is given to pedestrian movements on roads. Curbside walkways exist along most roadways. Furthermore, the Land Transport Authority (LTA) has created an environment conducive to walking by providing about 400 sheltered pedestrian overhead bridges. Walkways connecting nearby transport nodes such as transit stations, bus stops, and taxi stands are also sheltered to provide a better walking experience in sunshine or rain. In some situations, these sheltered walkways were even extended to nearby amenities, such as schools and public housing (Koh et al., 2011).

City planners may learn from the successes of other cities. In Singapore, the popularity of biking is still limited. To promote biking, Singapore could adopt a bike-sharing project similar to Chicago’s and create more biking infrastructure and a safer, convenient biking environment. Chicago could learn from Singapore’s efforts to increase walkway connectivity by connecting walking corridors to major destinations and parks or by creating more overhead bridges to ensure safety.

Future research

In this study, because of GIS data availability, I chose the typical environmental features commonly used in previous studies. Using this set of environmental features enabled us to compare our results with previous studies, but other environmental features may also impact active travel behavior. Some detailed environmental characteristics such as the amount of greenness and specific crime type distribution can be used to further explore the urban features that impact active travel. Future research could break down the crime indicator into different types and use LiDAR data to model the amount of greenness in an urban area.

In addition, the battery is the main restriction for smartphone usage. In our experiment, the battery can work continuously for 6-8 hours depending on the different phone models. Our participants turned off the app when they were stationary or indoor. Future research should attempt to capture less GPS points per minute to save power or use other dedicated devices like smart watches to collect travel data. This is a new pathway for physical activity or travel behavior measure and has great potential in studies in human health and environment.
**Conclusion**

The findings from this study suggest that certain components of the built environment increase the likelihood of active travel. Green spaces in both cities were consistently associated with more recreational active travel. Closeness to restaurants or food courts was associated with utilitarian walking in both cities. Surprisingly, higher density was not always associated with more active travels. Crime did not have a significant relationship with the different modes of active travel. This study has two main contributions. First, I use an ecologically sound study design to investigate the relationship between the built environment and travel behavior. The use of smartphones to measure travel behavior and contextualize the movement in the built environment allows us to investigate the impacts of direct environmental exposure on travel mode choices. Second, the comparative design across two cities accommodates more environmental variability. The comparison between the two major cities in this study reflects how different environmental components influence active travel behavior. The methods used in this study can be applied to other places and quasi-experimental study design to increase the generalizability of the results. Results of this study can be used by city and environmental planners to support young adults' active lifestyle and promote lifelong health.
References


CHAPTER 4
Why do we prefer walking here?
Connecting active travel behavior with environmental perceptions

In the past two decades, people in many parts of the world have become increasingly sedentary, which has led to soaring rates of obesity related disease and dramatically increased medical costs (Ogden, Carroll, Kit, & Flegal et, 2012). One promising population level approach to overcoming this challenge is to embed physical activity in people’s everyday lifestyle by encouraging people to walk, run, and cycle rather than drive when traveling to work or other activities (Rutt et al., 2008; Guell et al., 2012). A variety of studies argue that a well-designed built environment may be key to encouraging people to engage in active travel (Ding & Gebel, 2012; Harris et al., 2013).

We know that people’s perception of their neighborhood environment affects their willingness to be physically active (Hume et al., 2009). We do not, however, know much about people’s perceptions of the specific features that constitute a well-designed urban neighborhood.

What understanding we do have comes primarily from structured and standardized questionnaires used in previous studies. These instruments by their nature, ask people to respond to features in the built environment that researchers believe are important factors in active travel. To what extent do these features overlap with those that regular people – that is, non-researchers – identify as impacting their decisions to engage in active travel? What new knowledge might we obtain if we allowed people to identify and describe the features of the built environment that promote or discourage active travel (Mahmood et al., 2012)? What are the characteristics of the built environment people prefer to walk in? What features draw them outdoors? What characteristics of the city do they report as impacting their decisions to walk or cycle? How can we validate their reported perceptions with their actual travel behavior? Will the features and characteristics they identify be the same as the ones commonly used in questionnaires? Without this knowledge, it will be difficult for urban planners, designers, and municipal leaders to create cities that promote physical activity.

In addressing these questions, I begin by reviewing our understanding of the impact the built environment has on active travel and the methods used for obtaining people’s perspectives about such issues. Next, I examine the travel behavior of young adults in Chicago and Singapore and assess hundreds of photos of places and features that they report impact their willingness to
walk, run, or cycle. I then compare the findings of this assessment to existing models describing the relationship between the built environment and physical activity.

**Background**

**Environmental framework for active living**

Social environmental frameworks for understanding and describing active living have been proposed in previous studies (McNeill et al., 2006; Vrazel et al., 2008). Among these models, Pikora et al.’s (2003) model of four environmental dimensions has been widely used as a framework to categorize and describe the built environment features that promote physical activity. This model comprises four dimensions: aesthetics, functionality, destination, and safety.

Aesthetics includes characteristics that might be considered visually or experientially appealing. Greenness is the most frequently used aesthetic indicator associated with outdoor activity (Coombes et al., 2010). Functionality refers to the physical attributes of the local environmental structure, such as street connectivity, land use diversity, and density (Buck et al., 2011; Yan et al., 2010; Zhang et al., 2006). These indicators were widely reflected in the literature on the connections between the built environment and physical activity (Badland et al., 2008; Buck et al., 2011; Dygryn et al., 2010). Desirable places or service areas are defined as destinations. Destinations refer to restaurants, shopping, recreational, educational, financial, cultural, and healthcare facilities. Finally, safety refers to environmental characteristics that help people feel safe when they engage in active travel. Typical measures of safety include reported crimes, severity of crimes, and traffic accidents (Weiss et al., 2011; Zhang et al., 2006).

Although Pikora et al.’s model (2003) has been widely used in previous studies, we do not know the extent to which it overlaps with the categories that might grow from the inhabitants of, or visitors to, cities. If the environmental features people perceive conform to Pikora et al.’s model, we may be able to find out which specific features from the environmental constructs (e.g. aesthetics, destination, functionality, or safety) matter most to local people. If the desired environmental features do not conform to the theoretical framework, we may be able to identify the difference between the two and potentially adjust the model to accommodate more environmental features that promote active travel.

Most prior studies investigating environmental attributes that promote or discourage active travel have been conducted in the U.S., Europe, and Australia (Ding et al., 2011). Asian cities have rarely been investigated. The extent to which people’s perceptions of the built environment
in Asian cities conform to Pikora’s model is unclear. Thus, in this study, I collected data from people in multiple neighborhoods in Chicago and Singapore.

**Structured and user-derived environmental perception**

Understanding the characteristics of the built environment that lead to active travel is important, but how might we measure environmental perceptions from people? Many of previous studies have used survey instruments to examine the perceived environment. These surveys were usually developed based on theoretical models from academic experts and expertise from practitioners in the related areas (Cerin et al., 2006). Several questionnaires have been widely used. First, the Neighborhood Environment Walkability Scale (NEWS) and its derivatives, such as NEWS-A (an abbreviated version), NEWS-CFA (confirmatory factor analysis version) and NEWS-Y (version for youth) are among the most popular instruments for predicting the impact of the built environment on walking (Van Dyck et al., 2011). NEWS was developed in 2002 to assess people’s perception of neighborhood features such as residential density, street infrastructure, and neighborhood satisfaction. NEWS has been validated in several countries and has been widely used in previous studies (Cerin et al., 2006; Saelens et al., 2003; Sallis, 2011). Table 9 shows the overall structure and examples from a NEWS questionnaire.

The second widely used questionnaire is the Assessing Levels of Physical Activity and related Health Determinants (ALPHA) Measure of Environmental Perceptions. The ALPHA instrument is designed to assess the environmental features that support physical activity in European cities (Meusel et al., 2007). It includes 49 items and has been validated in Europe and used in a number of studies (Spittaels et al., 2009; Spittaels et al., 2010).

The third questionnaire is the Neighborhood Quality of Life Study (NQLS) survey, which includes a set of questions related perceptions of one’s neighborhood (Sallis et al., 2009). This survey also has a version for senior people (SNQLS).

In addition to these instruments, other scholars have developed questionnaire to suit the local environment and specific participant groups (Salmon et al., 2013; Gómez et al., 2010).

Although these questionnaires are widely used and have been validated, they ask participants to respond to pre-determined aspects of the built environment. That is, participants cannot identify and evaluate the features and attributes they feel are most important, but rather must respond to the features identified in the questionnaire. It is likely that items in the questionnaire do not fully capture participants’ perspectives regarding the preferred features of the
environment. In addition, features or characteristics that describe the built environment in one neighborhood or city may not be relevant in a different neighborhood or city. Allowing people to identify and describe the features or attributes that make them more likely to engage in active travel behavior in their own cities may provide richer details that can guide planners, designers, and municipal officials who are trying to create more active cities.
<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
<th>Example question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of residences</td>
<td>6</td>
<td>How common are detached single-family residences in your immediate neighborhood?</td>
</tr>
<tr>
<td>Facilities in the neighborhood</td>
<td>23</td>
<td>About how long would it take to get from your home to the nearest businesses or facilities (e.g. book store/ park)?</td>
</tr>
<tr>
<td>Access to services</td>
<td>7</td>
<td>To what extent do you agree there are many places to go within easy walking distance of my home?</td>
</tr>
<tr>
<td>Streets in my neighborhood</td>
<td>5</td>
<td>To what extent do you agree the streets in my neighborhood do not have many, or any, cul-de-sacs?</td>
</tr>
<tr>
<td>Places for walking and cycling</td>
<td>5</td>
<td>To what extent do you agree there are sidewalks on most of the streets in my neighborhood?</td>
</tr>
<tr>
<td>Neighborhood surroundings</td>
<td>6</td>
<td>To what extent do you agree there are trees along the streets in my neighborhood?</td>
</tr>
<tr>
<td>Safety from traffic</td>
<td>8</td>
<td>To what extent do you agree the speed of traffic on the street I live on is usually slow (30 mph or less)?</td>
</tr>
<tr>
<td>Safety from crime</td>
<td>6</td>
<td>To what extent do you agree my neighborhood streets are well lit at night?</td>
</tr>
<tr>
<td>Neighborhood satisfaction</td>
<td>14</td>
<td>How satisfied are you with the access to public transportation in your neighborhood?</td>
</tr>
</tbody>
</table>
Geo-tagged environmental perceptions

Most measures of environmental perceptions lack specific location information (Dennis et al., 2009). Although scholars have used questionnaires, interviews, self-administered drawings, photography, or a combination of these methods, they have rarely geo-tagged the settings (Anthamatten et al., 2013; Guell et al., 2012; Lee & Abbott, 2009). Locational information is important for some reasons. First, people may assume that certain features of an environment are present only in certain locations. For instances, aesthetic features may concentrate in parks. Feeling of crime or safety features may cluster in poor neighborhoods. Without locational information associated with the photos, we will not be able to verify if those assumptions hold true in different places. Second, without locational information, we cannot know which specific places promote or impede active travel. The lack of location-based perception prevents us from linking positive or negative perceptions to specific places, and reduces opportunities to conduct place-specific interventions that might improve local environments based on data from nearby residents. Third, geo-coordinates of the photos can be a bridge to connect the locations of environment perception with the locations of travel behavior. Using the location information of the photos, we will be able to examine the environmental features where most active travel actually occurs.

In this study, I seek to better understand the features of the built environment that young adults indicate to support active travel. I then assess the extent to which their perceptions conform to existing models describing the built environment’s influence on active travel. Another goal of this study is to investigate the distribution of people’s perceptions across the city to see what the desirable environmental features are at the places where active travel takes place. To capture a wide range of environmental features, I conducted this study in two cities, Chicago and Singapore. I used smartphones to collect spatial travel behaviors and geo-tagged photo-narratives to probe people’s perceptions of environments that promote or discourage active travel. I address three research questions:

1. What features of the built environment do users suggest promote or inhibit active travel?
2. To what extent do these features overlap with features identified in previous environmental model describing ways to promote active living, such as Pikora et al.’s framework for physical activity? What are the differences between features derived from user perceptions and the features commonly identified in research questionnaires?
3. How are the perceived features distributed across city space and what are the desirable environmental features at the places where higher levels of active travel take place?

**Methods**

**Participants**

Participants were recruited from Chicago and Singapore through fliers, email invitations, and Facebook messages. 121 participants who had no difficulty in conducting daily activities and who owned smartphones were recruited from four universities: University of Illinois at Chicago (UIC), University of Chicago (UC), National University of Singapore (NUS), Nanyang Technology University (NTU). Data were collected from June to October 2013.

**Travel behaviors**

To record travel behaviors, I developed an Android smartphone application. This application used the embedded GPS and accelerometer within a participant’s device to track and classify its users’ travel behavior. The experiment was conducted over a four-day time window, including weekdays and weekends. On these four days, participants were asked to turn on the smartphone application and let it run while they were moving outdoors. The application tracked their location through the phone’s GPS unit and classified travel modes based on an analysis of the data collected from the phone’s accelerometer and GPS unit. Participants were also asked to log their travel modes, time of travel, and travel purposes (e.g., recreational, utilitarian) so that I could validate the smartphone data. Each participant was given 20 dollars or a portable smartphone charger for their participation in this study.

**Photo narrative**

Photo narrative is a technique in which people take photos and attach descriptions that reflect their perception, preferences, or needs (Wang & Burris, 1997). Photo narrative is a flexible research strategy in which people create, describe, and discuss photographs as a means of expressing their perceptions (Wang et al., 1998). It is a procedure that allows individuals to identify environmental features they value and express their ideas regarding how specific neighborhood features impact their travel behaviors.

Participants used their smartphone to take about 10 pictures of places they traveled through regularly during weekdays and weekends. I asked that they take photos of environmental features that either promoted or inhibited active travel behavior (walking, running, biking as opposed to driving). Participants were provided a few photo examples in an effort to reinforce the purpose of
the study and were encouraged to take photos based on their own understanding of the environmental features that impacted their travel decisions.

For each photo they took, participants were asked to write a short description regarding the environmental features they were highlighting. Because the photos were taken by smartphone, a geographic location was attached to each photo. Participants could send the photos to us through email or post them on Twitter using their own account or the created accounts. I collected 962 photos from participants. 286 photos did not have appropriate geo-coordinates, and 194 photos had irrelevant scenes or descriptions. I used 225 photos from Singapore and 257 from Chicago along with the short descriptions for the analysis.

**Photo narrative analysis**

For each photograph, an identification number and the short description for the photo were entered into an excel spreadsheet. I used this information in a three-stage process to identify the characteristics of the photos that individuals said impacted their travel behavior. First, I examined each photo and description and identified one or more keywords or meaningful segments out of the description. If no keywords were selected or the keywords were not representative, I summarized the meanings based on the participant’s description and the content of the photo. I excluded photos and descriptions irrelevant to the built environment (e.g. weather, friends, or specific time of day). Second, keywords were grouped into emergent themes. For instance, photos of Michigan Lake and keywords including “Lakeview”, “yacht spot”, “great view”, “Michigan Lake” were summarized as “Lakeside View” theme. Third, I compared each photo and its associated theme with the four built environment constructs: aesthetics, functionality, destination, and safety (Pikora et al., 2003).

**Mapping environmental perceptions**

For each construct, I geocoded all photo-narratives using ArcMap 10.2. A kernel density map was created to illustrate the density of the narratives across the city. By doing so, I was able to compare the distribution of environmental perceptions across the city. I also linked each photo to the places where it was taken. Travel patterns and locations of the photos were overlaid to investigate the detailed environmental features where active travel clustered. I selected two areas (UIC and NTU campuses, and surrounding areas) with many activities took places, and displayed the environmental features related to active travel people reported.
Results

Perceptions

What features in the built environment did participants identify as either promoting or inhibiting their active travel? To answer this question, I examined the photographs participants took along their daily travel routes and their written descriptions of the photographs. The photographs and descriptions were coded to find recurring themes. Themes were compared to the four constructs in the Pikora’s framework: aesthetics, functionality, destination, and safety. As Figures 6 and 7 show, there was considerable overlap between the categories that emerged from our participants and the categories in Pikora’s model.

Aesthetics. Aesthetics accounts for 44.7% and 37.8% of all the photo narratives in Chicago and Singapore respectively. Photos and comments related to aesthetics were divided into six subcategories: green, water, open space, road, negative, and other features. Among these six categories, green space features attracted the most comments in both cities (38.6% in Chicago and 36.4% in Singapore). Green features such as tree shade, lush neighborhood spaces, and spacious grass and lawns were reported as positive factors attracting participants outdoors. One participant wrote: “we had a lot of fun of watching band performance here” (on a grassy field). Another participant wrote “jogging here everyday” (a green park). Water features were the second most mentioned aesthetic feature for active travel, accounting for 14.8% in Chicago and 25.0% in Singapore of the photo narratives related to aesthetics. Many participants mentioned that the lake view, waterfront walk, and beach were conducive for walking and playing. In addition, other features such as lighting in the evening, stylish architecture, colorful plants and facilities, and statues or sculptures were also identified as having a positive impact on active travel. Negative comments included barren, treeless landscape, rundown buildings, and littering. More negative features were reported in Chicago than in Singapore.

Functionality. Photos and comments related to functionality cluster into four subcategories: connectivity, facilities, accessibility, and rest places. Many participants from Singapore mention that connected walkways facilitate walking. Nearly 40% of comments about functionality were related to connectivity features in Singapore. Participants wrote, “Like walking in this wide sheltered sidewalk”, and “walkways connect to bus stops”, to describe their preference for the sheltered walkways. In Chicago, fewer comments were about connectivity, but more comments were related to the bicycle infrastructure. The Divvy bike-sharing (blue bike) project received
considerable attention. The bike infrastructure photos accounted for 35.6% of all the photos related to functionality in Chicago. Rest places and accessibility were two additional features identified as promoting active travel in Chicago. Roadside chairs and kiosks provided rest places for walkers while ramps with handrails were appreciated by others.

**Destinations.** A variety of destinations were mentioned by participants in both cities. Access to destinations accounted for 17.9% and 20.4% of all the comments from Chicago and Singapore respectively. In both cities, easy access to food courts, restaurants, and food booths were identified as features that promote walking. In Singapore, photos and comments about the mass transit station and bus stops were frequently identified as promoting walking. The local shopping mall and grocery store received repeated comments. Participants wrote: “It is very easy to walk to NTUC for grocery shopping” (National Trades Union Congress, a shopping market)

**Safety.** In both cities, participants made multiple comments about safety. Safety features accounted for 14.4% in Chicago and 18.9% in Singapore. In Chicago, most safety-related comments concerned biking and walking infrastructure. For walking, sidewalks, crosswalks, and dedicated lanes were identified as features that help ensure safe walking. For biking, road markings to prioritize biking at the road junctions and dedicated bike lanes were mentioned. In Singapore, participants identified a greater number of safety features than did participants in Chicago, such as anti-collision barriers, structures that prevent falling, fences around constructions sites, and video cameras in outdoor spaces. The perceptions about physical features in the environment that promote and inhibit active travel largely conform to Pikora et al.’s environmental model, but are these perceptions consistent with items commonly used questionnaires? I address this question below.

**Questionnaire-measured and user perceptions**

To what extent do the reported perceptions of people in Chicago and Singapore overlap with the features commonly examined in questionnaires examining relationships between the built environment and active travel? To answer this question, I compared items in the commonly used NEWS questionnaire (Table 9) to the reported perceptions of our participants. Overall, I found a good deal of overlap but also a number of gaps. In terms of the overlap, many items from the NEWS questionnaire concerning destinations were similar to comments made by our participants. In terms of the gaps, I also noticed two aspects. The first discrepancy concerns the emphasis placed on some items in the NEWS compared to the reported perceptions of our
participants. For instance, the NEWS includes six items related to aesthetics: trees, attractive buildings, and areas free from litter. Our participants, however, repeatedly mentioned that water features are important factors pulling them outdoors. Another example in emphasis concerns the attention placed on residential density. In the NEWS questionnaire, there are six items concerning residential density, with questions such as “how common are the single-family residences in the neighborhood?” Participants from Chicago and Singapore, however, included few comments regarding types of houses or residential density.

The second gap concerns the level of detail regarding elements within the built environment. For instance, the NEWS questionnaire asks if there are interesting things to look at while walking. Reports from individuals in Chicago and Singapore, however, included descriptions of the actual items of interest, such as colorful facilities such as playgrounds or equipment, plants, water features, funny statues. For another example, reports are more related to local special features. The questionnaire asked about cycling trails, but bike-sharing projects actually attracted more attention in Chicago. The questionnaire touched on general street connectivity, but the connected and sheltered walkways with many other facilities are more valuable to people in Singapore.
### Aesthetics

**Road features**
- Road decoration (3)
- Road art (3)
- Trash can (2)
- Cleanliness (4)
- Roadside flower bed (4)

**Open space**
- Nice urban square (3)
- City/Street garden (5)

**Water features**
- Beach (2)
- Riverside walk (5)
- Lakeside view (4)
- River tours (1)
- Fountain (1)
- Clean water (1)
- Lakeside paths for activities (3)

**Negative Features**
- Barren land (4)
- Random buildings/neighborhood (3)
- Dirty wall (1)
- Graffiti (1)
- Littering (3)

**Green features**
- Presence of road trees (5)
- Seats in the green (4)
- Tree shade (8)
- Green courtyard (5)
- Lush green neighborhoods (5)
- Green space for people and pet (1)
- Spacious grass/lawn (8)
- Road along the green space (6)
- Well-trimmed bush (1)
- Mowed lawn (1)

**Other features**
- Building styles (2)
- Building shade (1)
- Statue (4)
- Colorful plants (4)
- Swing (1)
- Decorations for Halloweens (1)
- Quietness (5)

### Functionality/Facility

**Connectivity**
- Connected bikeway (1)
- Connected bridge (1)
- Long and connected run/walk way (3)
- Tunnel (3)
- Mixed run and bikeways (1)

**Accessibility**
- Handrails (2)
- Slope for wheelchairs (3)

**Rest areas**
- Rest station (2)
- Sitting curb (4)
- Bench (14)
- Flowerbed for sitting (3)

**Facilities**
- Exercise facilities (2)
- Bike repair booth (1)
- Bike rack (10)
- Drivvy bike sharing (7)
- Meter for bike (1)
- Drinking fountain (1)

### Destination

<table>
<thead>
<tr>
<th>Downtown tour (1)</th>
<th>Work place (2)</th>
<th>Bus stop (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookstore (1)</td>
<td>Swimming pool (1)</td>
<td>Tennis court (3)</td>
</tr>
<tr>
<td>Shopping business (3)</td>
<td>Play event (2)</td>
<td>Restaurant (9)</td>
</tr>
<tr>
<td>Yoga (1)</td>
<td>Pet shop (1)</td>
<td>Church (2)</td>
</tr>
<tr>
<td>Student center (1)</td>
<td>Library (1)</td>
<td></td>
</tr>
<tr>
<td>Garage sale (1)</td>
<td>Food booth (2)</td>
<td></td>
</tr>
<tr>
<td>Chinatown (5)</td>
<td>Grocery (2)</td>
<td></td>
</tr>
<tr>
<td>Park (1)</td>
<td>Playground (6)</td>
<td></td>
</tr>
</tbody>
</table>

### Safety

**Positive features**
- Dedicated lanes (8)
- Flat road surface (1)
- Presence of sidewalk (3)
- Overhead walkway (2)
- Pedestrian sign (2)
- Bike box/priority (1)
- Emergency call booth (1)
- Speed limit (1)
- Crosswalk (5)
- Road lights (1)
- Wide/straight walkways (4)
- Construction fence (1)

**Negative features**
- Sidewalk need maintenance (2)
- Road construction (5)

---

Figure 6 Diagram of the environmental features identified by photo-narratives made by participants from Chicago. Numbers represent the frequency for each theme.
**Geographic feature of the perceptions**

What is the spatial distribution of the features that participants say promote or inhibit their active travel? To answer this question, I mapped photo narratives of each construct within each city.

Figures 8 and 9 show the spatial distribution of environmental perceptions for aesthetics, functionality, destination, and safety. In Chicago, positive narratives about aesthetics concentrated along the lakeshore, downtown areas, and University of Chicago. I also found some negative perceptions clustered along the west side of the University of Chicago. The aesthetics

**Aesthetics**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open space</td>
<td>12</td>
</tr>
<tr>
<td>Urban gardens</td>
<td>7</td>
</tr>
<tr>
<td>Open space for activities</td>
<td>5</td>
</tr>
<tr>
<td>Water features</td>
<td>24</td>
</tr>
<tr>
<td>Stream/Pond in the garden</td>
<td>2</td>
</tr>
<tr>
<td>Lake view</td>
<td>5</td>
</tr>
<tr>
<td>Fish pond</td>
<td>2</td>
</tr>
<tr>
<td>Lakeside deck</td>
<td>2</td>
</tr>
<tr>
<td>Bay front walk</td>
<td>3</td>
</tr>
<tr>
<td>Riverside restaurant</td>
<td>1</td>
</tr>
<tr>
<td>Beach/Ocean/Lake side view</td>
<td>9</td>
</tr>
<tr>
<td>Statue</td>
<td>5</td>
</tr>
<tr>
<td>Creatures in the urban environment</td>
<td>2</td>
</tr>
<tr>
<td>Cleanliness and trash can</td>
<td>5</td>
</tr>
<tr>
<td>Natural decoration</td>
<td>1</td>
</tr>
<tr>
<td>Architectural style</td>
<td>3</td>
</tr>
<tr>
<td>Colorful plants/facilities</td>
<td>2</td>
</tr>
<tr>
<td>Art ornaments</td>
<td>1</td>
</tr>
<tr>
<td>Fountain</td>
<td>1</td>
</tr>
</tbody>
</table>

**Green features**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree shade</td>
<td>6</td>
</tr>
<tr>
<td>Trails in the green/plants</td>
<td>4</td>
</tr>
<tr>
<td>Green neighborhood</td>
<td>5</td>
</tr>
<tr>
<td>Green wall</td>
<td>2</td>
</tr>
<tr>
<td>Wide tree crowns</td>
<td>4</td>
</tr>
<tr>
<td>Hilly green area</td>
<td>1</td>
</tr>
<tr>
<td>Fresh air in green space</td>
<td>2</td>
</tr>
<tr>
<td>Plant diversity</td>
<td>3</td>
</tr>
<tr>
<td>Lush trees</td>
<td>5</td>
</tr>
<tr>
<td>Spacious grassland/lawn</td>
<td>2</td>
</tr>
<tr>
<td>Apartments in the woods</td>
<td>1</td>
</tr>
</tbody>
</table>

**Road feature**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean street</td>
<td>2</td>
</tr>
<tr>
<td>Nice lighting along the roads</td>
<td>2</td>
</tr>
</tbody>
</table>

**Negative features**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>2</td>
</tr>
</tbody>
</table>

**Functionality/Facility**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilities</td>
<td>12</td>
</tr>
<tr>
<td>Vending machine</td>
<td>4</td>
</tr>
<tr>
<td>Bike rack/parking</td>
<td>4</td>
</tr>
<tr>
<td>Drinking fountain</td>
<td>1</td>
</tr>
<tr>
<td>Postbox</td>
<td>1</td>
</tr>
<tr>
<td>Barbecue pit</td>
<td>1</td>
</tr>
<tr>
<td>Fitness facilities</td>
<td>1</td>
</tr>
<tr>
<td>Connectivity</td>
<td>23</td>
</tr>
<tr>
<td>Walkways connect to bus stop</td>
<td>1</td>
</tr>
<tr>
<td>Sheltered corridor</td>
<td>15</td>
</tr>
<tr>
<td>Connecting bridge</td>
<td>4</td>
</tr>
<tr>
<td>Tunnel avoids traffic</td>
<td>2</td>
</tr>
<tr>
<td>Short distance between malls and MRT</td>
<td>1</td>
</tr>
<tr>
<td>Rest areas</td>
<td>16</td>
</tr>
<tr>
<td>Sitting places along corridor</td>
<td>2</td>
</tr>
<tr>
<td>Street/Park bench</td>
<td>5</td>
</tr>
<tr>
<td>Kiosk for rest</td>
<td>3</td>
</tr>
<tr>
<td>Picnic or tent area</td>
<td>2</td>
</tr>
<tr>
<td>Wood maze</td>
<td>1</td>
</tr>
<tr>
<td>Sitting places at bus stops</td>
<td>3</td>
</tr>
<tr>
<td>Accessibility</td>
<td>7</td>
</tr>
<tr>
<td>Ways for disabled</td>
<td>2</td>
</tr>
<tr>
<td>Slope for easy access</td>
<td>1</td>
</tr>
<tr>
<td>Handrails along trails</td>
<td>2</td>
</tr>
<tr>
<td>Ramps for people with stroller</td>
<td>1</td>
</tr>
<tr>
<td>Facilities for the blind</td>
<td>1</td>
</tr>
</tbody>
</table>

**Destination**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gym</td>
<td>1</td>
</tr>
<tr>
<td>Cafe</td>
<td>3</td>
</tr>
<tr>
<td>Barbecue</td>
<td>1</td>
</tr>
<tr>
<td>MRT Station</td>
<td>3</td>
</tr>
<tr>
<td>Club</td>
<td>1</td>
</tr>
<tr>
<td>Exhibition center</td>
<td>1</td>
</tr>
<tr>
<td>Ice cream booth</td>
<td>1</td>
</tr>
<tr>
<td>Park</td>
<td>5</td>
</tr>
<tr>
<td>Seashore</td>
<td>1</td>
</tr>
<tr>
<td>Swimming pool</td>
<td>2</td>
</tr>
<tr>
<td>Fruit market</td>
<td>2</td>
</tr>
<tr>
<td>Food court</td>
<td>9</td>
</tr>
<tr>
<td>Tennis</td>
<td>2</td>
</tr>
<tr>
<td>Grocery</td>
<td>3</td>
</tr>
<tr>
<td>Outdoor study places</td>
<td>2</td>
</tr>
<tr>
<td>School bus/Bus stop</td>
<td>5</td>
</tr>
<tr>
<td>Trails</td>
<td>1</td>
</tr>
<tr>
<td>Playground</td>
<td>5</td>
</tr>
<tr>
<td>Shopping mall</td>
<td>5</td>
</tr>
<tr>
<td>Fitness facilities</td>
<td>1</td>
</tr>
</tbody>
</table>

**Safety**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated trails</td>
<td>6</td>
</tr>
<tr>
<td>Overhead walkway</td>
<td>1</td>
</tr>
<tr>
<td>Road light</td>
<td>3</td>
</tr>
<tr>
<td>Pedestrian sign</td>
<td>2</td>
</tr>
<tr>
<td>Materials prevent falling</td>
<td>3</td>
</tr>
<tr>
<td>Highlighted curb/step lines</td>
<td>4</td>
</tr>
<tr>
<td>Direction sign on the ground</td>
<td>1</td>
</tr>
<tr>
<td>Safety sign at the train doors</td>
<td>1</td>
</tr>
<tr>
<td>Fences around construction</td>
<td>4</td>
</tr>
<tr>
<td>Anti-collision concrete barriers</td>
<td>2</td>
</tr>
<tr>
<td>Handrails</td>
<td>3</td>
</tr>
<tr>
<td>Safe run/walk/bike ways</td>
<td>8</td>
</tr>
<tr>
<td>Swimming ring</td>
<td>1</td>
</tr>
<tr>
<td>Speed Signs</td>
<td>1</td>
</tr>
<tr>
<td>CCTV</td>
<td>2</td>
</tr>
<tr>
<td>Crosswalk</td>
<td>6</td>
</tr>
</tbody>
</table>

Figure 7 Diagram of the environmental features identified by photo-narratives made by participants from Singapore. Numbers represent the frequency for each theme.
features are not just distrusted in natural areas, but spread extensively. For functionality, positive comments concentrated in the downtown area. Photos and comments about the Divvy Bike Sharing project were scattered across the city. Functionality features along Lake Michigan were mostly about connected walkways and rest places. Most destinations were concentrated around the two campuses. Access to restaurants and bus stops were mentioned multiple times and in multiple locations. Safety features generally followed the road network system because traffic safety concern related to the biking infrastructure and walkways came up most often. A few comments were related to crime, but they did not display particular spatial patterns.

In Singapore, the four themes did not exhibit obvious spatial patterns. Aesthetics features have a wide distribution. Other features clustered in and around both campuses and central Singapore (e.g. Orchard, Botanical garden and Chinatown). Photos and comments about aesthetics were scattered in many locations. In the city center, more comments were related to architectural styles, cleanness, and road decorations. Along the water features, comments about sand beaches, lakeside deck, and lake views were described as promoting walking. The most popular comments on functionality were the connected and covered walkways. This feature was reported around both campuses and many residential neighborhoods. More safety features were mentioned in the Nanyang Technology University campus neighborhood and downtown areas.
Figure 8 The distribution of environmental perceptions in Chicago derived from the photo-narratives made by study participants. We mapped the perceptions by constructs.
Figure 9: The distribution of environmental perceptions in Singapore derived from the photo-narratives made by study participants. I mapped the perceptions by constructs.
In addition to mapping the location of the features that promote and inhibit active travel, I provide two snapshots, one from each city, that provide a more fine-grained look at the features that impact active travel. The snapshots are from the University of Illinois at Chicago campus and NTU campus neighborhood, where many activities took place.

Figure 10 includes photos taken around the University of Illinois at Chicago campus where many active travel photos and comments were clustered. Pikora’s four constructs (aesthetics, functionality, destination, and safety) were all represented in the photo narratives in and around the UIC campus. Regarding aesthetics, participants pointed out both positive and negative features of the environment. At the northwest corner of campus, shown in Figure 10, where less active travel occurred, dirty walls and streets were reported as discouraging walking. However, some positive features such as “beautiful and quiet neighborhood” were identified in the neighborhoods around campus. In many of the photo narratives, participants commented on the Divvy bike-sharing program. One participant wrote: “It is very convenient to ride (the Divvy Bike), especially for visitors.” For destinations, both recreational and utilitarian destinations were reported from participants. One Walmart branch located close to a student residence was reported as a convenient place to buy groceries. Sport facilities such as tennis court also attracted some activities. Safety was another important feature reported many times in downtown and around UIC. One participant noted the traffic signs around campus, and wrote “speed limit sign to slow down traffic.”
Figure 10 Photos taken from participants around UIC campus. (a) “Dirty wall and street” (b) “speed limit sign to slow down traffic” (c) “walk to buy stuff” (d) “Benches in campus park” (e) “beautiful and quiet neighborhood” (f) “It is very convenient to ride, especially for visitors” (g) “Place to play tennis”. Points on the map were the actual movements.

Figure 11 shows a sample of the photos close to the NTU campus. I found that places where people frequently engaged in active travel were associated with many positive environmental comments. Running activities were observable around the Chinese Garden. Many participants reported lake views, green space, and statues the same places. Along bus stops, anti-collision barriers, and yellow curb lines were reported as features to keep people safe who were waiting buses. Around the NTU campus, connected walkways and green views were the most reported features.
Discussion

I investigated the characteristics of built environments that young adults from Chicago and Singapore perceive as either encouraging or discouraging their participation in forms of active travel. I then compared their reports to both Pikora et al.’s framework and items in the NEWS questionnaire, one of the most popular tools for assessing the impact of the built environment on physical activity. I found that characteristics described in participants’ photo narratives largely conformed to the four categories in Pikora et al.’s framework. In the aesthetics category, greenness and water features were consistently and widely reported from both cities. For functionality, connectivity, accessibility, and rest places were frequently mentioned. In the safety category, people mostly mentioned issues relating to traffic, pedestrian, and bike safety. Finally, in the destination category, people in both cities appreciated easy access to destinations such as restaurants and bus stops.
Although there was considerable overlap between the photo narratives and the items in the NEWS questionnaire, our findings suggest that future iterations of the NEWS questionnaire should include more items related to aesthetics and safety.

With respect to the spatial distribution of physical features that promote or inhibit active travel I found some variation in the two cities. In Chicago, aesthetics features were reported across the city, and destination features mainly clustered around the two campuses and city center. Safety features were mostly concentrated on traffic safety and generally followed the road network. In Singapore, features scattered more widely across the city, but concentrated more around the two campuses, central Singapore, and a few places close to MRT stations.

Below, I consider the major contributions of this work, policy implications and possible interventions, and future research directions.

**Contributions**

This study makes three contributions: it provides a unique validation of Pikora et al.’s model, suggests some modification to the NEWS questionnaire, and presents a new methodology that may be widely applicable in other areas of scholarship in which the spatial distribution of features in the built environment are important to measure.

Reassuringly, the features of the built environment that people perceived as supporting or inhibiting active travel corresponded well to Pikora et al.’s model. This is reassuring because our participants likely had no knowledge of the model and yet they spontaneously produced photo narratives that largely corresponded to its categories. These finding validate Pikora et al.’s model and suggest that it should continue to be used in studies not only in the West, but also in Asia.

The photo narratives also largely overlap with items in the NEWS questionnaire. This was especially the case for items related to destinations. Observations from the photo narratives, however, did diverge from the items in the NEWS questionnaire. One example of this was the diversity and specificity of observations related to aesthetics, where water features were mentioned a number of times in the photo narratives but were not included among the items in the NEWS questionnaire. Another example was that the photo narratives made little mention of housing density, while the NEWS questionnaire includes six items related to housing density. For safety, NEWS asked environmental features from both traffic and crime aspects, but photo narratives provide more detailed description of traffic concerns, such as subway boarding and alighting and safe zones at bus stops.
Based on these results, researchers who use the NEWS questionnaire may want to include more items about aesthetics and safety features of the built environment. Questionnaires about local environment uniqueness are also needed. Or, because the features people notice and value may vary from city to city, researchers may also consider using more open-ended research methods, such as the photo narrative, to find out what features people notice and value in particular cities.

The use of smartphones to track active travel behavior and geo-tagged photo narratives to gather people’s environmental perceptions could be useful for many other studies. I found that collecting the geo-tagged photo narrative using smartphones was a convenient and cost-effective approach. The photo narrative method used in this study also provided location information about people’s perception, which helped to map the positive and negative features across spaces and to identify opportunities for future planning interventions. For example, I found that aesthetics perceptions are not only distributed around natural areas, such as parks or lakeside, but also spread to residential or commercial areas. People frequently mentioned decorations along the street, clean environments, and statues in the built-up areas. These features in a less green environment are non-trivial in promoting active travel.

This study also suggests that current tracking technology is effective in collecting spatial data for travel behavior monitoring and modeling. Future research on health environment studies such as obesogenic settings can benefit from the methods introduced in this paper. For instance, photo narratives can be used to analyze the perceived food environment in the neighborhood, pinpointing the healthy and unhealthy places to eat and letting people express their desired environmental interventions for healthier diet (Dennis et al., 2009).

**Practical implications**

Perceptions from city inhabitants’ geo-tagged photo narratives can provide a wealth of information about the urban design features that effectively promote active travel and the spaces that need to be improved. Importantly, these perceptions come from ordinary residents, not planners, and thus may reflect how residents perceive these features and function in a space. These photo narratives can help city planners in Chicago and Singapore to know which features to develop and maintain, and which areas still need improvement. These practical implications are likely generalizable to other cities seeking to create active living environments. City planners
in different cities can learn from Chicago and Singapore’s strengths and weaknesses in promoting active travel.

Participants from Singapore identified a number of environmental features they believe are associated with active travel, and Singapore should continue to maintain these effective features. Weather in Singapore is normally hot, humid, and rainy—conditions that are unfriendly to active travel. Still, Singapore is ranked as one of the most walkable cities in the world (Sanyal, 2009). Singapore seems to support high level of walking through good planning. Many of our participants mentioned the sheltered and connected walkways, easy access to facilities (e.g., food kiosks, vending machines, sitting places), direct links between residential blocks and vehicular transport (e.g., transit stations, bus stops, taxi stands), which encourage walking in sunshine or rain. Good street connectivity was mentioned by people as a feature of the city that promotes walking. Subway stations are usually connected to major commercial and residential hubs, which provide convenient walking environment. People also wrote about easy access to different types of destinations. Many participants reported easy access to food kiosks, food courts, and vending machines. Daily errands such as dining or grocery shopping can be achieved within walking distance or a few bus stops in many residential neighborhoods. In Singapore, a variety of design details related to safety were reported by people, such as wide distribution of closed-circuit television camera (CCTV) monitors, anti-collision barriers at bus stops, and overhead walkways that allow pedestrians to avoid traffic. Wide application of these features may explain the relatively low rate of crime and annual traffic accident in Singapore (WHO, 2009).

Convenient, safe, and connected walkways and easy access to a variety of destinations are promoting active travel in Singapore, and Singapore should continue to develop and maintain these features. Chicago and other cities could learn from Singapore’s strengths by providing more walkways and easy access to destinations. Although the climate in Chicago is very different from Singapore, both often have weather conditions that are unfriendly to active travel. Thus, city planners in cities like Chicago might encourage more walking by adding covered walkways that connect popular destinations.

In contrast, biking in Chicago is more prevalent. People reported many biking related features. “Blue bikes” (the Divvy bike sharing projects) in Chicago were photographed many times. People also frequently noted the safety of dedicated bikeways. These comments indicate that Chicago’s efforts to promote biking are working. To enhance biking and public transit, the
city makes it easy to bring bikes on subways and buses. To further encourage biking, planners offer more bike-sharing where ridership is high, while establishing a strong backbone of infrastructure where ridership is currently low (Chicago Department of Transportation, 2013). People notice and value the biking infrastructure and safety features.

Chicago should continue to develop and maintain their biking infrastructure to encourage biking. In addition, the well-developed biking infrastructure in Chicago may be helpful for cities such as Singapore that seek to promote cycling.

In addition, the well-developed biking infrastructure in Chicago may be helpful for cities such as Singapore that seek to promote cycling.

In addition to the features identified in these two cities, the research methods used in this study can be replicated in other cities where scholars want to explore a wider range of design features that might support active living. Those features, of course, must be directly associated with local residents. Planning strategies involving the perceptions of local residents will help generate new knowledge about the kind of environmental features that promote more active travel.

**Future research**

One extension of this work is to more thoroughly examine the impact of the built environment on active travel that takes place, or might take place, at night. The built environment likely has an enormous impact on the extent to which physical activities take place after dark. In this study, a few photos were taken in the evening. Photos taken during the night also represented each of the four aspects of the environmental characteristics (Figure 12). For aesthetics, participants expressed their preference for walking in a well-decorated night environment. For safety, participants wrote “Safe to walk under road light” to express the importance of road light to prevent crime. Areas with clustered shopping malls, restaurants, and clubs were more popular than other places during the evening. Facilities such as barbecue pits also facilitated activities after sunset. Future research should explore night features of the urban environment that promote and inhibit active travel after dark. For instance, would there be opportunities for people to walk after dark? What are the environmental and social facilitators or restrictions for activity after dark?
Conclusion

This study used smartphones to record the movement of people across space and time while also asking them to take photographs of places or features in their city that either promoted or inhibited active travel. Analysis of the travel patterns and hundreds of photographs taken in Chicago and Singapore conform to Pikora et al’s model of active travel. I also found the content of the photo narratives overlapped a good deal with items in the NEWS questionnaire, although the questionnaire paid less attention to aesthetic aspects of the built environment and more attention to building density than did our participants.

In Chicago, participants identified aesthetics features and biking infrastructure as the main catalysts of active travel. In Singapore, the connected and covered walkways, proximity to food courts, green environment, and wide variety of safety features were the main catalysts of active travel.

The user-based methods employed in this study may prove to be a powerful new technique for understanding how inhabitants perceive the impact of the built environment on their behaviors. I hope this approach will be used by designers and planners to tailor environmental design in ways that promote not only active living but also a variety of outcomes that can help produce healthier lives.
References


Ding, D. & Gebel, K. (2012). Built environment, physical activity, and obesity: What have we learned from reviewing the literature? *Health and Place, 18*(1): 100-105.


CHAPTER 5

Conclusion

In this dissertation, I examined how the built environment influences people’s active travel behavior in three ways. First, in an effort to better measure people’s travel behavior and overcome multiple problems associated with obtaining accurate travel data, I developed a convenient and cost-effective Android smartphone application. I used this application to simultaneously collect location, time, and travel mode data from 121 participants at four university communities in two cities: Chicago and Singapore. Second, to overcome the Uncertain Geographic Context Problem, I collected environmental characteristics data at the places where people actually traveled to find out which environmental characteristics were associated with more or less active travel and modeled the travel modes based on these characteristics. Third, to complement the statistical modeling, I captured ordinary individuals’ perceptions of the environmental features that promote or inhibit active travel through geo-tagged photographs and narratives.

In this chapter, I provide a summary of the major findings related to each of these issues, consider the contributions to our understanding of how the built environment impacts active travel, identify implications for cities, and suggest several ideas for future research.

Summary of findings

To what extent did the smartphone software application accurately measure and classify travel behavior? To answer this question, I developed algorithms to auto-classify smartphone data into different travel modes (e.g., walking, running, biking, and in-vehicle). I then used a set of evaluation criteria to test the classification accuracy. Results from smartphone data classification demonstrated that smartphone devices were capable of capturing data that can reveal how, where, and when people travel. The
classification system I developed recognized various travel modes with accuracy greater than 80% for all the modes of travel; most modes were estimated with accuracy greater than 85%.

The classification system has two components, the machine-learning based classifier and the rule-based classifier. The rule-based classifier comprises a set of predefined criteria to assign points with specific attributes into one mode. I used the rule-based model to adjust the results from the machine-learning based methods. Results of the classification after adjustment were about 6% more accurate than the classification without the rule-based adjustment. This suggests that a smartphone can be used as an alternative to the commonly used self-report as to collect travel behavior data by facilitating data collection and improving travel mode classification accuracy. The details of this portion of the study are described in Chapter 2.

Do different environmental features significantly promote active travel behavior? To address this question, I conducted mixed logistic regression to model the travel behavior based on environmental characteristics. I used odds ratio (OR) to investigate the environmental features that were more or less related to active travel than in-vehicle travel after adjusting demographics and physical activity attitude variables. Results showed that greenness was consistently and positively associated with more recreational active travel than vehicle travel in both cities. Destinations in general showed a positive relationship with active travel behavior, but odds ratios suggest the relationship is relatively weak. Crime did not show a significant relationship with different modes of active travel.

I also found that the association between the built environment and travel behavior varied in different areas within one city. High building density was related to more utilitarian active travel in the urban areas, but in the suburb, a negative relationship between recreational activities and high building density was obvious. More details can be found in Chapter 3.
In addition to built environment features themselves, individuals’ perceptions of the environment certainly shape their choice of travel mode. Hence, I explored from users’ perspectives the features of the built environment that promote or inhibit active living. I examined the photographs people took when they traveled and their written descriptions of the photographs. The photographs and descriptions were coded to find recurring themes. Results from the geo-tagged photo narratives showed that features in aesthetics, functionality, destination, and safety all make a difference in promoting an active lifestyle. In Chicago, participants identified aesthetics features including greenscapes, street decoration, and lake views as the main catalysts of active travel. Participants also indicated that Chicago’s bike infrastructure supports active travel. In Singapore, the connected and covered walkways, proximity to food courts, green environment, and wide variety of safety features were the main catalysts of active travel. These results described in chapter 4 suggest that planners should consider people’s perspectives in designing environments that promote active living.

**Methodological contributions**

Our findings have three main methodological contributions. For travel data monitoring, most previous studies combined GPS receiver and accelerometer data to measure both activity location and intensity. However, synchronizing data from GPS receivers and accelerometers requires additional steps, and data is lost if the timestamp does not match (Dessing et al., 2013; Dunton et al., 2014). Using smartphones to monitor travel behavior provides great potential for future healthy behavior studies. Broad use of smartphones and mobile apps enables convenient individual trip monitoring. Using smartphones to collect travel data will alleviate participants’ burden to wear additional devices and avoid data loss because of data synchronization. In addition, auto-classification of travel mode based on collected smartphone data provides a new avenue to collect travel activity data. The direct travel mode detection provides fast and
convenient behavioral information, which can be used in a wide variety of fields, such as public health, urban planning, and transportation.

With regard to modeling the environmental features for active travel, I confirmed that close surrounding environmental features do influence people’s choice of travel mode. In addition, I confirmed that active travel behaviors are widely distributed across the city. Previous attempts to capture travel behavior that used buffers or census tract-based approaches did not capture the wider environment where active travel behavior occurred. I suggest that future studies in environment and physical activity should use more location-based approaches to investigate where active travel actually occurs and thereby more fully explore the association between travel behavior and the environment.

This research also emphasizes the importance of understanding people’s perceptions in order to create spaces that promote active travel. I argue that using geo-tagged photo narratives to gather people’s environmental perceptions is a convenient and cost-effective approach, and could be applied to other study areas. These geo-tagged narratives provide locational information for people’s perceptions, which can help to map the positive and negative features across space and to identify opportunities for future improvement. More importantly, the information is gathered from users, indicating how users of a space might respond to and interact with an environment. In addition, perceptions from users also suggest potential revisions of widely-used environmental questionnaires, such as the Neighborhood Environment Walkability Scale (NEWS). For instance, items in the commonly used questionnaires are not diverse and specific enough to capture the features in the local environment that people perceive and value. Questionnaires can be modified to include features that are unique to the local environment.

The comparative framework employed in this study increases the generalizability of the results, and our confidence in the association between environmental features and active travel behavior among university students from different regions. By comparing
western and eastern university contexts, we gained confidence in the findings because the presence of green space was consistently associated with more activities. Providing more urban green space in various forms close to campuses will likely enhance recreational active travel at the population level. In addition, creating more destinations such as transportation hubs or various food opportunities will likely increase walking in both cities.

In addition to findings that were consistent across both cities, the methods used in this study are also generalizable to a larger context. The smartphone-based travel behavior monitor can be easily applied in different cities. The photo narrative approach can be implemented anywhere one needs to collect impressions about the local environment from local people. The combined methods used in this study provide both statistical confidence and user generated information about the environmental influences on active travel behavior. This research framework can be well extended to other sites for future exploration.

**Practical implications**

From this study, we gained more confidence in our understanding of the influence of the built environment on active living and expanded our knowledge of what environmental features promote active living by looking at two cities, Chicago and Singapore. By conducting this study in two contexts, we realized green features are consistently and significantly associated with more active walking and running. This confirms previous research on the role of urban green space in promoting active lifestyles. However, there were also some differences between the two cities in the features that promoted active travel. By examining these differences, cities can learn from each other and develop environments more conducive to active travel.

Chicago is more biking friendly. Biking infrastructure features, such as the Divvy bike sharing projects, the dedicated bikeways, priority of bikers in the traffic rules, and bike facilities really make this city bikable. To further encourage biking, planners built
more bike-sharing where ridership is high, while establishing a strong backbone of infrastructure where ridership is currently low (Chicago Department of Transportation, 2013). In contrast, Singapore seems to support a high level of walking through good planning. Many features such as the sheltered and connected walkways, easy access to facilities (e.g., food kiosks, vending machines, sitting places), and direct links between residential blocks and vehicular transport (e.g. transit stations, bus stops, taxi stands) encourage walking even in a hot and humid environment. In addition, Singapore seems to create a safe environment for people to engage in physical activity. A variety of design details related to safety explain the relatively low rate of crime and annual traffic accident in Singapore. These features include wide distribution of closed-circuit television camera (CCTV) monitors, anti-collision barriers at bus stops, and overhead walkways that allow pedestrians to avoid traffic.

City planners in Chicago and Singapore can learn from the strengths of each other and apply similar strategies. This research can provide city planners with information about the features that effectively promote active travel and the features that need improvement. For instance, Singapore can learn from Chicago as they seek to implement their 2013 “Cycling-For-All” master plan. In addition, these planning strategies can be generalizable to other cities seeking to create an active living environment.

Results from our model of built environment and active travel suggest that people tend to capitalize on the elements of the environment close to their neighborhood. Thus, planners should develop active travel amenities that utilize the unique features of the nearby local environment. For instance, planners should consider building safe and well-connected trails to link major natural areas to nearby neighborhoods or shopping areas. Neighborhoods with few stores might benefit from additional businesses to encourage people to walk. Although urban designers can learn from cities such as Chicago and Singapore in developing active living environments, they should also
consider their unique local context and the natural features that might encourage active travel.

**Future research**

Future research can build upon this study in a number of ways. For smartphone data collection and classification research, future research can test the possibility of estimating energy consumption based on smartphone devices. In addition, future research could further improve the accuracy of travel mode classification and distinguish more travel modes. Previous studies have suggested geographic information, such as bus lanes, metro stations, and bikeways, can assist in travel mode detection, in terms of both higher accuracy and more types of modes (such as bus, train, and driving) (Tsui & Shalaby 2006; Gong et al., 2012). For instance, by using the metro station and route data, in-vehicle mode can be further classified as public transit or private cars. Future research can build upon the current algorithm to incorporate geographic information in travel mode detection.

There are other ideas that can be examined with the data from this study. In this dissertation, I investigated the travel mode of each trip. But there are other important attributes that can be extracted from GPS data. For instance, we can estimate the places where people visit in order to understand the purpose of the trip, and thereafter obtain a better idea about what types of destinations are more important to active living. GPS data provide location and duration information about where people move and stop. By developing a smart algorithm combined with good GIS data, it is possible to detect where people visit and how long people stay in one place. Future work can capitalize on GPS travel data to extract more useful travel information.

For the association between environment and active travel, in this study, I only chose the general environmental features, because of GIS data availability and conformity to previous studies. To provide more concrete design guidelines to promote healthier behaviors, in future research, I plan to look into specific environmental features
in detail and investigate relationships that might exist. For instance, the crime indicator can be broken down into different types of crimes and time when the crime happened to see if crime type, number, or time is associated with active travel. In addition, we found greener space is associated with more physical activity, but how much green do we need or is there a threshold of greenness for people to be active? To answer these specific questions and give specific design guidelines, future studies can build upon Light Detection and Ranging (LiDAR) data to model the green volume in the urban area and to see how the dose of greenness is associated with physical activity levels. In addition, when it is possible, future research can launch a prospective study to evaluate how a change in an environment may influence people’s travel behavior. For instance, studies can compare people’s biking behavior before and after the launch of bike sharing project to examine the extent the bike infrastructure fosters active living.

Another pathway for future study is to examine the association of each specific travel route to its surrounding environment. For instance, people may detour from their origin to destination because of a more attractive green environment or because of perceived traffic safety. Investigating the reasons why people make a detour is a good way to examine the micro environmental features that potentially influence active travel. Google Map can calculate the hypothesized route from origin to destination, along with an estimate of the shortest distance, while the actual travel can be measured by the GPS points. By comparing the difference between two routes and questionnaire answers from participants, we can discover some potential environmental features good or bad for active travels.

One extension of the photo narrative work is to more thoroughly examine the impact of the built environment on active travel that takes place, or might take place, at night. The built environment likely has an enormous impact on the extent to which physical activities take place after dark. Research could explore opportunities for people to walk after dark or the environmental and social facilitators for activity after dark.
Final remarks

It is critical to have a better understanding of the relationship between the built environment and human health in order to improve place-based experiences and promote lifelong health. By using the innovative smartphone-tracked travel data coupled with the self-report travel log method and by combining statistical and spatial analysis with qualitative photo narratives, we increased the understanding of the extent to which local environmental characteristics are related to active travel of young adults in two different settings.

I also employed interdisciplinary methods from geography, public health, and urban planning. I attempted to integrate realms of urban planning and public health by using innovative technologies such as GIS and smartphone sensors to examine social behaviors. These tools can be widely applicable across different disciplines.

This project also probed individuals’ perceptions on the environmental characteristics that promote active travel behavior. This project provides valuable information for city planners seeking to improve active living environments and help citizens lead healthier lives.

I am also excited about this work because the methods used in this paper can be widely replicated in other cities and generally applied to environmental and behavioral research.

Finally, I hope the results of this study will help to combat the obesity and physical inactivity problems that plague cities. The methods and findings from this study can help cities take steps to promote more active travel, thereby increasing physical activity and keeping citizens healthier.
References


