A Generalized Model of Activity Space: Is Geographic Context Lifestyle Specific?

Spielman, S.E. 1\*, Singleton, A.D.2

1 – University of Colorado Boulder, Boulder, USA

2 – University of Liverpool, Liverpool, UK

\* - seth.spielman@colorado.edu

A Generalized Model of Activity Space: Is Geographic Context Lifestyle Specific?

This paper introduces the generalized activity space as a way to bridge area based and activity based representations of geographic context. A generalized activity space is an empirically derived delineation of an individual’s geographic context that blends, and attempts to reconcile, insights from multiple disciplines. This paper argues that micro-scale space-time paths fail to account for important social determinants of behavior, because they emphasize “contacts” over “contexts, a problem that may be solved, in part, by using a broader “generalized” representation of geographic context. Support for the existence of these generalized activity spaces is found. Analysis of 34,500 trips by 7550 individuals in Atlanta indicate that lifestyle and residential location jointly condition the configuration of a person’s activity space. However, residential location, by itself, is not an effective descriptor of the configuration of a person’s activity space. We argue that recent innovations of computational movement dynamics have potential to inform the discovery of generalized activity spaces through a more robust consideration of the socio-economic characteristics of moving entities. In addition, this paper presents evidence that activity spaces tend to anisotropic, this finding has some implications for applied spatial analysis.

Keywords: Lifestyle; Activity Space; Clustering; Circular Statistics

A Generalized Model of Activity Space: Is Geographic Context Lifestyle Specific?

# Introduction

In a 1969 address to regional scientists Torsten Hagerstrand outlined the principles of time geography. Time geography is a representational system which recognizes that human activities are embedded in space and time and thus are constrained by them. Hagerstrand motivated this idea by reference to a “twilight zone” in the social sciences between biographical analysis (of the kind conducted by historians) and the analysis of areal data (as is done by regional scientists). Time geography was envisioned as a system to establish “coherence between the two ends of the scale” (Hagerstrand 1970, 9). Since Hagerstrand’s initial formulation, enormous progress has been made extending and operationalizing its concepts (Kwan 1998; Kwan 2000; Miller 1991; Miller 2003; Mennis et al. 2013; Patterson and Farber, 2015). The widespread availability of Global Positioning Systems (GPS) and mobile devices has increased the contemporary relevance of Hargerstrand’s ideas. However, 50 years later researchers still bemoan that fundamental problem Hagerstrand raised remains, Schapfler et a; (2021) in *Nature* note, “the link between this microscopic behaviour and the temporal spectrum of recurrent mobility fluxes arising from an entire population is missing” (p.522). While the literature that directly engages with time-geographic concepts has remained largely contained to the discipline of Geography the basic idea of looking at individual movement in space and time has spread well beyond the discipline.

This broad body of work on human movement contains within it fundamental tensions. Some of these tensions relate to epistemic differences between disciplines, for example, the orientation of Physicists toward generalization based on statistical properties of phenomena compared to Geographers (and other Social Scientists) orientation toward theory built upon the observed particulars of place (O’Sullivan and Manson, 2015). However, many of these tensions are genuine problems in which well-grounded theories and/or observational studies disagree. For example, decades of work in transportation research find that the characteristics of individuals (location, wealth, stage in the life course) are fundamental drivers of human spatial behavior but recent work in Physics finds that “high degree of spatial and temporal regularity” regardless of these demographic characteristics (Gonzalez et al, 2008). Furthermore, within the social sciences there is disagreement between those who believe social context is individually constructed through spatio-temporal activity versus those who see it as collectively constructed through networks, institutions, and structural forces (we refer to these respectively as the “individualist” and “collectivist” perspective).

In this paper we introduce the concept of a generalized activity space in an attempt to reconcile these tensions. As a concept generalized activity spaces provide a way to refine the parameters of the generalized movement models identified in Physics and in so doing link them with the social scientific literature. Generalized activity spaces also provide a way to reconcile the tension between the individualist and collectivist views of social context. This paper describes these tensions in more detail, introduces the concept of a Generalized Activity Space, and then attempts to develop empirical evidence for its viability as a concept. This paper is by no means a proof of Generalized Activity Spaces, our goal is to introduce the idea and begin a conversation on how to reconcile tensions in the study of activity spaces. We believe such a reconciliation is necessary to advance our understanding how human behavior shapes and is shaped by the environment.

Decades of work in Public Health, Sociology, and other social sciences which finds that area of residence, often delimited by some type of census geography, can have important impacts on well-being and human development. This interdisciplinary literature is sometimes referred to as “neighborhood effects.” The neighborhood effects literature asks to what extent are geographic patterns in health and/or social outcomes due to area/neighborhood-level characteristics as opposed to individual or family-level characteristics. This question can be difficult to answer because people are often clustered (or segregated) geographically be race, ethnicity, and/or socio-economic status. The close linkage between demographic factors and area of residence makes it difficult to statistically separate “compositional effects” from “contextual effects”. Moreover, Kwan (2010?) and others have noted that the identification of contextual effects is complicated by uncertainty in the appropriate definition of social context. Richardson et al. (2013) highlight this in *Science* arguing that, “Life paths of individuals collected with GPS/GIS methods can provide more accurate assessment of exposures to environmental or social risk factors” (Richardson et al. 2013, 1391).

The Richardson et al. (2013) argument that space-time paths provide a “more accurate” assessment of individual environmental contexts is what we refer to as the “individualist” perspective on activity spaces. This stands in contrast to the collectivist view of social space which focuses institutional and social forces that structure and shape life, this does not divorce the collectivist framework from space. Collectivists would argue social fact are situated within space, surrounded, and influenced, by other social facts. The literature on space-time paths highlights the individualist-collectivist divide. Hagerstrand’s conception of time geography was deeply connected to thinking about how space and time structure interpersonal interactions. However, Giddens (1985) criticizes the time geographic framework because it stresses the physical location of a person and hence neglects to account for broader social influences on behavior. Giddens (1985, 270) notes that Hagerstrand’s time geography emphasizes “the corporeality of the human being, in structured space-time contexts” and treats individuals “independently of the social settings they confront in their day-to-day lives.” Giddens’ point is that location matters not only because of the direct physical exposures it provides, location is a marker of key social determinants of health and behavior (Diez-Roux 1998; Macintyre et al. 2002; Browning and Cagney, 2003). Social norms and culture may be place specific but not transmitted through geographic mechanisms such as exposure via proximity.

The application of a time geographic perspective to the measurement of socio-spatial contexts does not jibe with the dominant conceptualization in the social sciences. Historically, in the social sciences, context has been seen as important because it is collectively constructed and not “personal.” Abbot (1997, p.1152) in describing the influential Chicago School of Sociology, emphasizes the collective nature of space, "no social fact makes any sense abstracted from its context in social (and often geographic) space and social time...Every social fact is situated, surrounded by other contextual facts and brought into being by a process relating it to past contexts." Contemporary theoretical concepts, like social capital and collective efficacy are viewed by Sampson et al (2012, 2002) as properties of collectively defined neighborhood contexts that have positive spatial externalities, noting that for example the willingness of people to intervene on behalf of others can have broad impacts. Browning (2006) found that unequal distributions of community resources produced conditions that made residents of some areas more vulnerable than others to the 2003 Chicago Heat Wave. Traditional place-based social scientific analysis, for all of its shortcomings, uses area of residence (place) as a proxy for a complex set of forces, including cultural, economic and political. Schulz and Northridge (2004) see physical exposures as one element in a complex multi-scale eco-social system that connects place to health outcomes. Radil et al. (2010, 308) argue that "The ability of spatial analysis to incorporate the relative location of social actors, and the linkages between them, can, paradoxically, atomize actors being studied through a “spatial fetishism” that ignores or is unable to address the social relations that construct the spaces within which actors operate.” An exclusive focus on an individual’s location in space fails to capture broader social context.

Stated simply, Giddens and others suggest that micro-scale time geographic analysis of individual behavior is vulnerable to the conflation of location and context. Location structures social and environmental context but does not fully reveal it. This critique of time geography can seem counterintuitive, is inherently concerned with social relations and how they constrain activity spaces. Stated differently, Time Geography emphasizes “contacts” (direct social interactions such as coupling constraints) over “contexts” (neighborhoods, place-based norms, culture, social supports). The aim of this paper is to introduce and explore empirically a concept that may present a middle ground by exploring Hagerstrand’s (1970, 9) assertion that "on the continuum between biography and aggregate statistics, there is a twilight zone to be explored, an area where the fundamental notion is that people retain their identity over time, where the life of an individual is [their] foremost project, and where aggregate behavior cannot escape these facts.” In the spirit of Hagerstrand’s idea this paper seeks to identify a substantively relevant “twilight zone” between the analysis of micro-scale locational biographies and metropolitan scale circulation patterns. Specifically, we attempt to develop a meso-scale concept positioned between the analysis of individual locational biographies (GPS traces) and their aggregate description by invoking a geodemographic perspective. As such, this paper develops a new framework to group individual activity traces into *generalized activity spaces,* so that rather than describing a specific individual, we leverage information about many individuals inclusive of their socio-economic characteristics.

Space-time paths emphasize the corporeality of human experience and neglect the fact that locations which people do not enter can influence behavior; a child who lives near an open air drug market may not go near the drug dealing but living near danger may influence his/her behavior and/or well being. Mennis et al. (2013) have gone a long way toward addressing this concern through the incorporation of affective information about locations within a person’s activity space. Nonetheless, Gidden’s critique challenges Richardson et al.’s (2013) assertion that GPS provides “a more accurate assessment” of exposure to social risk factors than traditional place based approaches. We believe that both location based and place based study of social systems have value, and develop the concept of a generalized activity space in an attempt to bridge these two analytic paradigms.

In spite of the growing interest in computational movement analytics in a variety of disciplines, the connections between individual characteristics and movement patterns remains underexplored; a comprehensive review titled *Human mobility: Models and applications* Barabosa et al (2018) provide little insight on how to marry trajectories and the demographic characteristics of moving persons. We believe that such models can (and should) accommodate demographic information. We believe that Generalized Activity Spaces provide a mechanism for doing so.

# Generalized Activity Spaces

A generalized activity space expands on the individual-centric notion of an observed path to consider how groups of similar people behave spatially. It is “generalized” in that the concept attempts to abstract away some of the physical geographic and individual specific constraints that might shape movement at a micro scale in order look for patterns or similarities at the meso-scale. The concept is also generalized in the sense that it tries to integrate individual (activity) and communal (place) based definitions of socio-spatial context.

Generalization of activity spaces involves developing broad group specific statements about spatial behavior from the observation of individual behaviors. The concept looks beyond generalizations based purely on the geometry of movement to consider both demographic characteristics and spatial behavior simultaneously. That is, a generalized activity space reasons over observations of spatial behavior and lifestyle. In some sense the concept feels problematic, why develop generalized representations if the raw individual-level information is available? As we argue throughout this paper there is reason to suspect that raw representations of behavior fail to accurately capture social and environmental contexts within which people live their lives.

We believe that the concept of a *generalized activity space*, an empirically derived area that delineates a person’s context based on both individual attributes and the space time paths of other individuals might provide a unit of analysis that overcomes the trap of spatial fetishism described by Giddens (1985) and Radil et al (2010). Generalized activity spaces are aggregations of space-time paths into meso-scale areal units of analysis that are rooted in both observed behavior and demographics. Hagerstrand (1970, 14) notes, “over a lifetime he steers his path through a string of daily prisms, growing in radius during early years of his life and shrinking at an advanced age.” The concept of a generalized activity space accommodates changes in the sphere of spatial activity over the life course, as Hagerstrand suggests but does so in way that attempts to overcome the shortcomings of focusing only on an individual’s location in space and time.

Perhaps there are “types” of activity traces, similar people, regardless of location use space in a similar way. For example, the “dog bone” in figure 1 is a generalized activity space that might represent a person with a single stationary place of employment. In a personal conversation, an executive of a mobile phone company referred to the activity traces produced by individual mobile phone users as “dog bones.” For those with a single employer and a single residence, over time, activity traces begin to look like figure 1. Activities are geographically distributed around home and work, these expansive areas form the ends of the bone, a person’s commuting path forms the shank of the bone. The dog bone is just a conceptual prototype, it is a meso-scale concept that does not refer to a specific individual or a particular place.

**Figure 1. An example prototype activity space: “Dog Bone” (H represents home location, W represents work location)**



 There is already substantial evidence of the existence of generalized activity spaces. Gonzalez et al in *Nature* (2008) found regular patterns of activity in a data set describing 100,000 mobile phone users over six months. Song et al (2010similarly found that human movement patterns were regular and predictable using 50,000 anonymized call records. However, because these data are anonymized this work neglects decades of research on travel behavior which tell us that individual characteristics such as age, gender, employment, vehicle ownership, and family structure are all important drivers of travel behavior (Pisarski 2006, Goodchild and Janelle 1984, Hanson and Hanson 1981, Pas 1982, Janelle et al. 1988). While individual and family characteristics shape travel behavior it is not yet clear how they shape activity spaces. Work by Sila-Nowicka et al. (2015) takes a step in this direction, they find that gender and age conditioned activity spaces using a small sample of non-anonymized GPS traces covering multiple locations in Scotland. For all of its methodological sophistication the recent work in computational movement dynamics fails to account for known demographic, social, and economic drivers of mobility.

We believe that the concept of Generalized Activity Spaces could be easily integretated into work on movement from physics and other disciplines. For example, Schalfler et al 2021 develop an argument that that the probability of visiting a particular location, in this case a location is grid cell on a map, can be predicted by $p\\_i(r,f)\ =\sfrac\{\mu\\_i\}\{(rf)\^\eta\}\ $. Where *r* is the distance of a person’s home from that location and *f* is the number of previously observed visits, and $\mu\\_i$ is a parameter describing the attractiveness of a place. From a social scientific perspective this formulation is challenging because of its total disregard of the characteristics of a person. It implies that the attractiveness of a place is universal – that for example, race and/or socio-economic status do not effect the attractiveness of a place – which from a social scientific perspective is false on its face. Generalized activity spaces provide a way to get socio-economic status into these and other equations, for example by making p(r, f) or $/mu$ conditional on *g*, where *g* is an observed type of generalized activity space.

In spite of all of the arguments in favor a “generalized” approach to activity, it is not at all clear how these generalized representations of spatial behavior might be identified empirically, scaled-up from individual characteristics and trajectories. One approach could be to group people based on the neighborhood (or census zone) in which they live and to use such groupings to identify a sphere of activity for the residents of a particular area. However, Miller (2007) and Wellman et al (2002) have argued that the nature of place is changing; new technologies and the subsequent intermingling of real and virtual spaces have altered the meaning of residential location such that people living near each other can have very different daily experiences. By this argument residential location is of declining importance, Forrest and Kearns (2001, 2129) note, "it would seem that as a source of social identity the neighborhood is being progressively eroded with the emergence of a more fluid, individualized way of life. Social networks are city-wide, national, international and increasingly virtual." Miller (2007) argues that the declining influence of place on human experience as a result of space adjusting technologies will increasingly lead to the potential for a “place-based fallacy.” The place-based fallacy occurs when one incorrectly infers the attributes, activities, or experiences of people based on the places they live. While it is possible to group space-time paths based on residential location, the place-based fallacy suggests that this may lead to spurious conclusions about activity spaces because it would mix heterogenous groups.

Another type of approach might be to focus on characteristics of the movement itself via the techniques emerging from computational movement analysis. For example, McArdle et al (2014), Laube et al (2005), and Dodge et al (2008) develop ways to create taxonomies or clusters of movement patterns based on this geometric characteristics; Buchin et al (2014) extend this idea by developing ways to measure the similarity of trajectories accounting for the context in which the movement takes place, similarly Jaegal and Miller (2019) develop wats to measure the similarity of space-time prisms; Shoval and Isaacson (2007) use sequences of visiting locations to identify types of trips; Demšar et al, 2018 use aggregated to show urban circulation patterns via the density of paths in a place-time. However, these approaches neglect the characteristics of the moving entities, in part because human movement database tend to be attribute poor (Miller et al. 2019). Inherent in Hagerstrand’s formulation of twilight zone is the idea that aggregate patterns can be decomposed and/or individual space-time paths can be aggregated to form meaningful units of analysis which respect an individual’s characteristics.

Demographic characteristics and movement are intertwined. The COVID-19 pandemic clearly highlighted this, Weill et al (2020) show how area-level socio-economic status is clearly related to movement patterns before, and during, the pandemic suggesting that wealth creates a kind of elasticity of movement; wealthier areas were more able to reduce movement during the pandemic than lower income areas. However, this work is unfortunately limited by data availability. Demographic and economic information is only available at the area-level, that is, while the individual movement of phones is tracked, the characteristics of the person transporting it is unknown so socio-economic trends are established by ascribing the ecological characteristics of the area where the phone spends the night to the person. The concept of a generalized activity space more directly links movement to individual characteristics. The kinds of activities in which one engages both shapes, and is shaped by, a person’s lifestyle. Variations in lifestyles lead to variations in types of activity spaces. Dog bones describe one lifestyle, where a person has a single home and a single job and commutes regularly between them. Perhaps there are other prototypical activity spaces?

## Lifestyles

The concept of lifestyles starts “with the assumption that tastes are neither completely determined by economic status, as was implied by Marx, nor totally individualized. Tastes are determined in part by relative position in the markets for wealth and prestige, in part by individual choice informed by education and experience, and in part by voluntarily chosen, collectively held standards that determine lifestyles. Lifestyle differentiation takes place both inside and outside the markets for wealth and prestige and hence crosscuts them” (Zablocki and Kanter 1976, 269). Salomon and Ben-Akiva (1982) define lifestyle relative to an individual’s choice in three domains: (a) formation of a household, (b) participation in the labor force, and (c) orientation toward leisure. Lifestyles are determined by a person’s behavior in family, work, consumption, and leisure. Therefore by observing these variables it should be possible to gain insight into a person’s lifestyle. Thus, lifestyles exist in the middle space between micro-scale individual decisions and macro-scale social and economic forces that constrain choice.

Kipnis (2003) notes that lifestyle can be considered both an independent and dependent variable. Marketers typically take the latter position, viewing lifestyle as a product of consumption. However, in the transportation literature lifestyle is viewed as an independent variable (Dong et al, 2015; Dielman et al, 2002; Hanson and Hanson, 1981; Kressner and Garrow, 2012; Salomon and Ben-Akiva 1983). Kearns and Forrest (2001, 2130) see activity spaces as an important determinant of individual identity, “the local neighbourhood remains important as a source of social identity but there are many other sources partly dependent upon our individual and collective time-geographies and action-spaces within the urban arena.” Identity and lifestyle are distinct concepts however this quite illustrates the endogeneity inherent in linking these concept to activity spaces. This paper treats lifestyle principally as a determinant of spatial behavior, not an outcome of it. However, it is important to note that that there are other possible approaches to the association between lifestyle (or identity) and activity space.

To investigate the meso-scale between the analysis of individuals space-time paths and aggregate city-scale circulation patterns it is necessary to identify a unit of analysis: using place based units of analysis may be problematic because of the place-based fallacy identified by Miller (2007). On the other hand geographic units of analysis may describe shared social, cultural and political experiences that translate into shared drivers of behavior. Lifestyle might be a viable alternative to place. That is, grouping people with similar lifestyles might provide a basis for the generalization of activity spaces. Meso-scale units of analysis are necessary to create generalized activity spaces, such as the dog bone, but do lifestyle or location provide a better basis for aggregation?

# Data

Efforts to integrate rich demographics and human movement are severely data constrained. GPS traces of mobile phone users are locationally rich but demographically poor- they provide a high spatial and temporal resolution data about location but almost always lack meaningful demographic and economic data about the user. The only data set that we were able to obtain that contained both rich descriptions of spatial behavior coupled with demographic and economic data about respondents was a 2001 travel survey in Atlanta, Georgia which consisted of three survey instruments. Although these are dated, the purpose of this is analysis to explore the viability of a concept, rather than provide an empirical understanding of these specific sets of collective behaviors. As such, for this example, we are not interested in the urban geography of Atlanta per se; rather that this offers a conveniently workable dataset with limited access constraint. The sample is large and balanced by socio-economic status and county of residence (the Atlanta Metro has 13 counties), it includes 7552 heads of household. Household heads completed a travel diary describing their travel patterns (origin, destination, time, purpose, and mode of travel). All trip origins and destinations were geocoded. The database includes 34,582 trips. Each household’s trips were recorded by phone interview and were reported over a two day period. The data is suited to the problem at hand because it includes both detailed individual level demographic characteristics and detailed records of spatial behavior. The data is publicly available on-line at the University of Minnesota Travel Survey Data Archive.

# Methods

TO assess the viability of generalized representation of activity we ask the following questions: Do people who live in similar locations have similar activity traces? Independent of location, do similarities in lifestyle translate into similarities in the use of space? As such, a key contribution of this paper is to look for evidence of group specific patterns across two-dimensional projections of activity traces. These methods are exploratory, aimed at suggesting further directions for the identification of generalized activity spaces. Three different ways of grouping paths are explored; First, location-based aggregations of space-time paths are created, then non-spatial aggregates based on the notion of lifestyles are constructed, and finally we explore the joint combination of location and lifestyle. This paper does not explicitly consider the temporal component of space-time path but this may be an important differentiator of generalized activity spaces in future extensions to this work.

## Activity Space Standardization via polar reprojection

The principal challenge to the identification of generalizable patterns of paths across space and time, is that these detailed geolocation data tend to encode a great deal of information about the configuration the built environment; that may obscure generalizable trends. Space-time paths are very sensitive to local constraints such that two substantively identical paths in different places would appear different because they are shaped by infrastructure and location in the urban field. For example, a delivery driver who works in the central business district and lives in a middle class suburb on the northern edge of the city would have a different activity space than a delivery driver who lives on the western edge of the city virtue of both configurational differences in the built environment in different parts of the city and the simple fact that one has a South-North commute and the other has an East-West commute. These two delivery drivers, while they have different activity spaces because they live in different places are otherwise very similar. At one level of abstraction the delivery drivers lead very different spatial lives, yet at another they are quite similar.

To directly compare the activity spaces of the two delivery drivers, the paths have to be expressed in a standardized form that abstracts them from their particular locational constraints. A common way in which a variable can be standardized is by expressing it in terms of deviations from the mean (Freedman, et al. 2007); and a corollary is that activity spaces can be standardized by expressing them as deviations from a common reference point. Thus, as the standardization of numbers makes them unitless, the standardization of activity spaces makes them “placeless”. Calkins and Marble (1980) developed polar transformations for geographic data and Saxena (1997) first used this technique to standardize travel patterns using a home-work axis. Kwan (2000) also used this technique to produce GIS style overlays of spatio-temporal activity patterns. Standardization to the home work-axis potentially allows one to statistically compare activity spaces, however it only works for people who are employed in a single job. For those who do not work or who have multiple jobs standardization of paths to a home work axis is problematic. One of the persistent challenges to the idea of generalized activity spaces is therefore the difficulty of standardizing space-time paths.

In this example we standardize activity spaces using a polar coordinate system that was constructed for each of the 7552 households in the analysis. Polar coordinate systems look like a dartboard, the bulls-eye, is the origin of the projection. Degrees are measured as departures from some arbitrary azimuth, conventionally this is due east or due north. If the azimuth is due east, zero degrees is due east, 90 degrees is south, and so on. Distance from the origin, when combined with the angular departure from the azimuth provides a unique location on a polar coordinate system. On planar coordinate systems locations are described by an *x, y* pair where the *x* coordinate and the *y* coordinate represent distance from some arbitrary reference line, polar coordinates are given by their angle of displacement from an arbitrary reference and distance from the origin (ρ and θ respectively). We defined the origin of each person’s coordinate system as their residential address, and the reference axis (i.e. 0 degrees) was a line drawn between the address and the Atlanta City Hall (Figure 2). Although Atlanta is a large polycentric area, employment density Downtown is almost twice (55.21 people / acre) that of the next nearest center within Emory (29.39 people / acre); and the residential-workplace flow that one might expect because of such geography are bore out in our sample. As such, although a bearing such as North might have been utilized, this would have the disadvantage being of arbitrary directionality, and therefore limit the comparability across common activity spaces. For each person trips were projected onto their unique, personal, polar coordinate system. Figure 2 graphically illustrates the outcome of the projection procedure. There are people living within two houses, i and j, the city hall (indicated by the building with a flag on top), and a supermarket (indicated by the shopping cart). For each person a unique polar coordinate system is defined from their household location. The supermarket can be described by a single coordinate vector in Euclidean coordinate system, after the projection the coordinates of the supermarket are defined relative to each houses’ unique coordinate system. By projecting paths to an individual specific polar coordinate system, the objective is to allow the generalization/standardization of paths.

We fully recognize that our method of standardization may be problematic, Atlanta is a polycentric city making a choice of a reference direction somewhat problematic. In spite of this shortcoming we believe we show it to be effective for our expository purposes. We hope that others will innovate and develop new ways of standardizing activity. One interesting potential approach to standardization lies in the work of Pappalardo and Simini (2017) who use semantic information about activity, the kinds of places people go, as captured in a travel diary, to generatively simulate travel patterns. The idea being that generalized activity *sequences* might have some potential basis for synthetically creating generalized activity *spaces*.

**Figure 2. A Simplified Example of the Polar Reprojection**



## Identifying Lifestyles

Lifestyle segmentation systems divide populations into discrete groups based on observed similarities in behavioral or demographic characteristics; and are widely used in the commercial sector because of their proven use in the profiling of consumption patterns (Webber and Burrows, 2018). These systems are also useful for profiling general travel behavior (Birkin, 2019; Martin et al, 2018; Hinks et al, 2018). Constructing lifestyle groups is a technical exercise and there are a variety of different options available to analysts including cluster analysis via K-means or other clustering algorithms (Singleton and Spielman, 2014), data reduction techniques such as self organizing maps (SOM: Spielman and Thill, 2008; Liu et al, 2019) or latent class analysis which is commonly prescribed to individual level data (Swait, 1994). A latent class represents a variable that is not directly observable (in this case, “lifestyle”); with the probability that people belong to a group given their observed characteristics. In this instance, latent class analysis has some important advantages over k-means type cluster analysis. Firstly, it accepts a variety of input measurement types: and while initially the technique was conceived for categorical data, in its modern implementation it allows for input of mixed data including nominal, ordinal, count, and continuous data. Since classifications are model based one can evaluate how well a given classification scheme fits the data using likelihood ratio chi-square statistics and other information criteria such as the Bayesian Information Criterion (BIC) or Akaike information criterion (AIC).

## Directional Statistics

Directional statistics are used to analyze angular observations and are particularly useful when working with polar coordinate systems. The angular component of a set of points in a polar coordinate system can be described in terms of their directional mean and degree of dispersion, dispersion measures how evenly distributed observations are around the mean (Mardia, 1972; Rohde and Corcoran, 2015). With directional data conventional statistical procedures are not appropriate, for example taking the arithmetic mean of trajectories in the 1 and 359 degree direction would be incorrect. There are a limited number of hypothesis tests for directional distributions, these test: 1) the uniformity of a distribution (Rayleigh Test) which measures the extent to which the vectors are distributed evenly in all directions around the circle; 2) the equality of a set of directional distributions (circular ANOVA); and 3) the equality of a pair of distributions (Watson two sample test) (Lund and Agostinelli 2007). We employ these tests to compare standardized trajectories for geographic and lifestyle-based groups of people.

## Research Design

We argue that the standardization of space-time paths make it possible to compare travel patterns without regard for the effects of location. As such, hypothesis tests implemented within this directional perspective are useful because we can examine if lifestyle or location have statistically significant differences in directional means and/or degrees of dispersions. Activity spaces with different directional means and amounts of dispersion, if standardized and superimposed, will show a different use of space; and as such, makes directional statistics a simple, parsimonious, and useful mechanism for finding differences in complex spatial patterns. However, the use of direction statistics is also limiting. It allows the identification of group level-differences but what it does not permit is identification of shapes, or areas, like the dog bone, that might constitute a generalized activity space.

As such, this paper evaluates a series of simple null hypotheses to test the efficacy of the standardization and identify both lifestyle and place-based geographic regularities in activity spaces. The null hypotheses are:

1. *H0:* The mean direction of travel is equal in all residential locations.
2. *H0:* The amount of variability in trip direction (trip dispersion) is equal in all residential locations.
3. *H0:* The mean direction of travel is equal for all Lifestyle groups.
4. *H0:* The amount of variability in trip direction is equal in all Lifestyle groups.
5. *H0:* The mean direction of travel is the same for all groups defined by the interaction of lifestyle and location.
6. *H0:* The amount of variability in trip direction is the same for all groups defined by the interaction of lifestyle and location.
7. *H0:* The mean distance travelled is the same for all lifestyle groups.
8. *H0:* The mean distance travelled is the same for all locations.

Hypothesis tests 1, 2 and 5 examine the efficacy of the standardization, and if effective, these null hypotheses will be rejected with the direction and dispersion of space-time paths not being geographically differentiated. Hypotheses 3, 4, and 6 examine the impact of lifestyle on space-time paths, and these hypotheses will be rejected if lifestyle is not associated with the direction or variability in travel patterns. Hypotheses 4 and 5 examine the interaction of lifestyle and location, and instead of defining groups geographically or based upon their lifestyle; for these tests groups, cases are defined through the cross-product of lifestyle and location; a person with lifestyle *A* in location *B* is in a different group from a person with lifestyle *A* in location *C*. Hypothesis 7 and 8 examine the distance travelled.

# Results

## Lifestyle Analysis

A total of 7,552 people were included in the lifestyle analysis, who in total made 34,582 trips during the survey period. The analysis was restricted only to adults (over the age of 18) who were the primary survey respondents. Individuals with missing data were omitted from the classification. We used all available demographic and economic variables. The groups are primarily differentiated by employment, household size, and home ownership (see Table 1 Figure 3).

The latent class analysis was implemented, and a 4-class model selected which optimized AIC and BIC; and accounted for 76% of the variance in the data set, meaning that the negative log likelihood from a 1 class model decreased 76% with the addition of 3 classes to the model. The results are shown in Figure 3 where the x axis shows each of the variables included in the analysis standardized to 0-1 range, with the lines representing the mean value of each variable for each class.

**Table 1. Latent Class Analysis Input Variables**

|  |  |
| --- | --- |
| **Variable**  | **Description** |
| Age  | Ordinal variable with 5 age ranges |
| Home Ownership  | Indicator variable for home ownership |
| Vehicle Ownership  | Number of vehicles per household |
| Ethnicity  | Ethnicity (White, Black, Hispanic, Asian/other) |
| Employment  | Indicator variable for full time employment (35 hrs/wk at 1 job) |
| Professional  | Indicator variable for those who identify as “professionals” or “managers” |

**Figure 3. Lifestyle Class Profiles**

Two groups (class 1 and 3) consist of people who work full time, 40 hours/week, a large proportion of whom call themselves professional or managerial workers. Two groups (class 2 and 4) include people who either work part time or do not work (i.e. are unemployed, disabled, retired). Almost half of the people (44%) are in groups with a high proportion of full time workers, who self-identified as professionals or managers, owned their own home, lived in multi-person households, and owned multiple cars. The second group of full time workers (class 3) represented 19% of the people, by contrast they lived in small households, had a much higher tendency to rent their home, were younger, owned fewer cars per household, and while largely Caucasian had higher probability of being African-American, Hispanic, or Asian then members of the 1st group. The two groups where part time workers or unemployed were dominant represented 35% of the dataset and on average were older than members of the group containing full time workers. The first group (class 2), representing 22% of the people lived in large households, tended to own their home, owned multiple cars, and had a very high probability of being Caucasian. The second group (class 4) representing 13% of the data was the most racially diverse group, and had the highest probability of renting their home, they owned the lowest number of vehicles per household.

 Figure 4 shows the aggregation of all activity traces for members of each lifestyle group, paths are displayed using the polar coordinates from each person’s coordinate system. The origin of the projection is each person’s home, the zero degrees is at the top of each panel.

**Table 2. Lifestyle Groups**

|  |  |
| --- | --- |
| **Class** | **Description** |
| Class 1: Professional families | Young professional families who own their home and multiple cars. Predominantly white. |
| Class 2: Part-time, non-professional families. | Older families. Own multiple cars. Non-professional, do not work full time. |
| Class 3: Professional singles | Young professional singles who are equally likely to own or rent their home. Usually own one car. Predominantly white. |
| Class 4: Older non-full time workers. | Older racially diverse renters who are unemployed, retired, or work part time. Lowest vehicle ownership. |

**Figure 4. Spatiotemporal activity traces aggregated by lifestyle.**



If members of a particular lifestyle group are geographically clustered the effect of location and lifestyle on spatial behavior may be confounded. That is, it becomes impossible to separate lifestyle and location effects. The difference of Ripley’s K-functions comparing each of the lifestyle groups to the entire sample indicate that two of the lifestyle groups are significantly geographically clustered when compared to the geographic distribution of the entire sample. Statistical evaluation of the difference between two k-functions is difficult (see Diggle and Chetwynd 1991 for a discussion), a difference near zero is seen as evidence that the two k-functions describe similar patterns, a positive value can be interpreted as evidence of clustering. At all scales, classes 3 and 4 (Professional Singles and Older Part-timers) appear to be geographically clustered. To reduce the impact of geographic clustering locations were described using quadrants; the city was divided using North-South and East-West axes with City Hall at the center. Each person was associated with one of 4 quadrants; location thus became a categorical variable in the analysis. The objective was to minimize the effect that particular locations being strongly correlated with particular lifestyles. Moreover, the number and size of zones was constrained by sample size; the cross-tabulation of geographic and lifestyle groups would dilute the sample if too many geographic groups were used. The data we have available is imperfect but we believe sufficiently robust for our purposes.

## Distance and Directional Patterns

Generally, the number of trips in each lifestyle group was proportional to the number of individuals in the group. Figure 5 shows directional kernel density plots of trip direction for each lifestyle group. The number of trips in each direction determines the height of the kernel. When looking at the directional patterns interesting trends emerge, particularly when the groups representing full time workers are compared to groups representing unemployed or part-time workers. In the groups representing full time workers there are a large number of trips in the 180 degree range these appear as bumps in the kernel density plots of trip frequency by direction for each group. This bump occurs because people have a tendency to run errands on their way home. The morning commute (in the zero degree direction) represents a single trip but the trip home, if it includes multiple stops, appears as multiple trips yielding a spike in the distribution.

The profiles in figure 5 represent only the directional component of trips, short trips and long trips are not differentiated. If lifestyle groups have different directional profiles it seems, as a matter of course, that they will have different activity spaces. Without regard for distance if groups tend to travel in different directions within the standardized space they are using the space around their home in a different way. It is possible however, that many short distance trips, with varied directions, could yield a generalized activity space similar to one with fewer long trips even though it would have a different direction distribution.

The analysis found that lifestyle, location, and the interaction of lifestyle and location were not significant predictors of the mean travel direction for the standardized data (Table 3). Location was not associated with the variability of directional patterns, but lifestyle was associated with variability. This finding also held for an interaction of lifestyle and location, that is, people of the same lifestyle in different locations had different amounts of variance in the directions they travelled. This provides some support for the efficacy of the standardization, it shows that the standardization washes out the directional effects of location. Lifestyle, regardless of where a person lives, is associated with the directional variability of travel (but not the actual direction).

**Figure 5. Lifestyle Group Directional Profiles**



**Table 3. Summary of Hypothesis Tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Location**  | **Lifestyle** | **Location\*lifestyle** |
| Mean Direction  | Hypothesis 1: Not Significant (p-value .1185)  | Hypothesis 3: Not Significant (p-value .6576)  | Hypothesis 5: Not Significant (p-value .4311) |
| Dispersion  | Hypothesis 2:Not Significant (p-value .6167)  | Hypothesis 4: Significant (p-value .00599)  | Hypothesis 6: Significant (p-value .00544) |

To further explore the differences and similarities among groups a series of pairwise comparisons were conducted. A Watson two sample test of homogeneity was used to test the null hypothesis that a pair of lifestyle groups has the same directional travel patterns. The same null hypothesis was examined for 4 randomly generated groups. For all pairwise comparisons of lifestyle groups there were significant differences in directional distributions (Table 4). These results agree with previous studies and find statistically significant differences in travel between analytically derived population groups (Goulias et al. 2007; Pas 1982; Salomon and Ben-Akiva 1982) and are also supported by the tests in Table 3. For randomly generated groups there was no difference in directional distribution (Table 4), this test was done to verify the utility of the method. We recognize the potential multiple testing problems here but are unaware of an alternative for directional distributions.

**Table 4. Distributional Hypothesis Tests**

|  |  |  |  |
| --- | --- | --- | --- |
| **Null Hypothesis**  | **Test**  | **Result**  | **Significance** |
| The lifestyles have the same directional pattern  | Watson Two Sample Test  | Reject (for all pairwise comparisons)  | p-value < .001 |
| Randomly selected subsets of the data have the same directional pattern  | Watson Two Sample Test  | Accept (for all pairwise comparisons)  | p-value > .10 |
| Travel isotropic (entire sample)  | Rayleigh Test for Uniformity  | Reject  | p-value < .001 |
| Travel isotropic (latent classes)  | Rayleigh Test for Uniformity  | Reject (for all groups)  | p-value < .001 |

Hypotheses 7 and 8 examine the distance travelled and were tested using a standard ANOVA and Tukey’s Honestly Significant Difference (HSD) Test. For all pairs of lifestyles and quadrants the difference in mean trip length is significant (p-value < .0001 with the exception of the comparison between Greyer Part-Timers and Part-Time Families which had a p-value of 0.003).

Finally, the null hypothesis that travel has a uniform directional distribution was examined; that is, do travel patterns tend to be isotropic relative to a person’s home? This is particularly pertinent given that lags in spatial analysis often assume isotropy. However, these results challenge this assertion. The Rayleigh test of uniformity rejected the null hypothesis that travel is isotropic both for the entire sample, and for each lifestyle group (Table 4, and Figures 4 and 5).

## Discussion

The paper examines the feasibility of identifying generalized activity spaces as a way to link individual and area based definitions of socio-spatial context and to provide a framework for adding a social scientific perspective to recent large scale studies of human movement. In particular, lifestyle, residential location, and the combination of lifestyle and location are examined as a basis for defining generalized activity spaces. The goal was to test the tenability of the concept of generalized activity spaces; and to identify potentially fruitful avenues for further work on generalized activity spaces.

This inquiry is motivated by the idea that a person’s context is not determined simply by the locations in which they spend time. The aim here was therefore to gain insight into Hagerstrand (1970, 9) assertion that, "on the continuum between biography and aggregate statistics, there is a twilight zone to be explored.” The conceptual framework of a generalized activity space links an individual space-time paths to broader trends in the use of the urban landscape while maintaining an individual’s identity.

Our analysis finds that lifestyle seems to be associated with the amount of variability in the direction that a person travels but not the actual direction: people with similar lifestyles exhibit similar variance in travel direction. This is not true of geography, location is not associated with the variability or the mean direction of travel, and this finding provides evidence for supports Miller’s (2007) place-based fallacy. People with similar lifestyles but residing in different locations seem to have different amounts of variability in travel. The separation of lifestyle and location was not as neat as one would have liked, some lifestyles are geographically clustered. These hypothesis tests, because of the sample size, are sensitive to small inter-group differences, thus the differences identified, while statistically significant may be substantively small. The similarities and differences identified are based on an analysis at a rather coarse granularity, however this level of abstraction was necessary to prevent the confounding of lifestyle and location, using finer units of analysis would have caused particular locations to be dominated by particular lifestyles. Furthermore, Stopher (2007) notes that 12 days of tracking information may be necessary to gain a full understanding of the an individuals activity patterns, beyond 12 days the patterns tend to get repetitive. This analysis used a large sample of short duration activity paths. It is necessary to repeat this analysis with a set of longer durations paths.

Standardized space-time paths are not isotropic (as indicated by the Rayleigh Tests in Table4), a finding confirmed by Gonzalez et al. (2008). This has some implications for the spatial social sciences. First, anisotropy in spatial behavior raises questions about the use of radial buffers or simple contiguity-based weights matrices in spatial analysis of behavior. Radial buffers, commonly constructed “as the crow flies” or using the street network, are not supported by this analysis. Even when location is standardized travel patterns are not uniformly distributed in all directions. Second, this raises some interesting questions about how best characterize the shape of human activity spaces. For example, Sebastian et al. (2002) develop a shape similarity metric that can be used to query databases of shapes to identify objects with a similar morphology. An interesting area of future work might be the incorporating such metrics into movement databases or identifying the extent to which similarity in the morphology of movement patterns relates to individual characteristics and residential location. One might hypothesize that within location-demographic groups shapes are more similar than across groups, that is, living in the same census tract and having similar demographic characteristics leads to similarly shaped activity spaces. Our preliminary work here suggests that both who you are and where you live may interact to shape the morphology of activity spaces. Doi et al (2020) take an interesting approach by inverting the problem, trying to estimate individual characteristics from semantically enriched GPS traces. By using GPS traces to understand the nature of the places that people stopped; they estimated gender and age with some mixed success.

Generative approaches seem an extremely promising alternative to focusing on the geometric characteristics of observed movement. Pappalardo and Simini (2017) develop ways to generate synthetic activity patterns based on travel diaries. Extending this idea, if one created a lifestyle specific sequence of activities one could synthetically generate a potential activity space for a specific type of person in a specific location. This would side step the need to empirically estimating the shape of a generalized activity space allowing a researcher to generate one (or many) based the characteristics of a person.

With a robust dataset describing human activity, that covered many types of places, one might be able to determine if there are generalizable shapes – informed by demographics and locational characteristics. Such insight would allow more detailed understanding of social and environmental exposures in the absence of invasive GPS based tracking of people. The COVID-19 pandemic highlights the need to understand how spatial behavior is conditioned by socio-economic characteristics. While studies such as Weill et a. (2020) make it clear that area-level characteristics are associated with spatial behavior, generalized activity spaces add precision to this idea by providing a generic framework for understanding how location and lifestyle shape behavior. If developed to fruition generalized activity spaces would allow one to proactively estimate exposure risk for types of people in types of places without a reliance on *ex post facto* analysis.

## Conclusions

Taken together these hypothesis tests suggest potential for the concept of a generalized activity space. There is significantly more work to be done develop the concept and a need for better understanding of how lifestyle conditions the relationship between the environment and behavior. Here preliminary evidence in support of the idea that different types of people have different prototypical activity patterns is presented. We argue that generalized representation of activity spaces overcome the spatial/locational fetishism inherent in time-geography and the neglect of individual behavior in research on neighborhood effects. We find that people who live near each other do not have similar space-time paths but nearby people with similar lifestyles do. Statistically significant differences were found for 16 discrete lifestyle-location categories, suggesting that activity spaces are simultaneously conditioned by both *who you are* and *where you live*.

We present the concept of a generalized activity space in the hope that in the future the fusion in individual characteristics (or lifestyles) and movement patterns can be more elegantly fused. Neither the data nor the methods to do exists currently, thus in this manuscript we use the best data set we could find to explore the viability of this theoretical concept. The dog bone was introduced as a prototype generalized activity space. What are the other types? Since the 1980s it has been clear that lifestyle is associated with both trip length (distance) and the overall amount of travel (Hanson and Hanson 1981). This analysis advanced this work by showing that lifestyle defined across multiple dimensions affects the spatial configuration of activities. The tractability of the concept of the generalized activity spaces was explored, and it seems to have promise as way to advance the use of GPS collected time geographic data in ways that are consistent with and sensitive to social-scientific theory. A key next step to this analysis is the exploration of the spatio-temporal shape of space-time paths, not just directional distributions. The identification of lifestyle specific directional distributions provides some hope that that it will be possible to identify meaningful generalized activity spaces.

# References

Abbott, A. 1997. Of time and space: the contemporary relevance of the Chicago School. *Social Forces* 75 (4): 1149-1182. doi: 10.1093/sf/75.4.1149.

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., ... & Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, *734*, 1-74.

Birenboim, A., and N Shoval. 2016. Mobility research in the age of the smartphone. *Annals of the American Association of Geographers* 106(2): 283-291. doi: 10.1080/00045608.2015.1100058.

Birkin, M. 2019. Spatial data analytics of mobility with consumer data. *Journal of Transport Geography*, 76: 245-253. doi: 10.1016/j.jtrangeo.2018.04.012.

Browning, C. R., and K.A. Cagney. (2003). Moving beyond poverty: neighborhood structure, social process, and health. *Journal of Health and Social Behavior* 44 (4): 552-571. doi: 10.2307/1519799.

Buchin, M., Dodge, S., & Speckmann, B. (2014). Similarity of trajectories taking into account geographic context. *Journal of Spatial Information Science*, *2014*(9), 101-124.

Calkins, H., and D.F. Marble. 1980. *Computer software for Spatial Data Handling: volume 1, full Geographic Information Systems*. Ottawa, Ontario: International Geographical Union Commission on Geographical Data Sensing and Processing.

Cornwell, E.Y., and K.A. Cagney. 2017. Aging in activity space: results from smartphone-based GPS-tracking of urban seniors. *The Journals of Gerontology: Series B 72 (5):*  864–875. doi: 10.1093/geronb/gbx063.

Dash Nelson, G., and A. Rae. 2016. An economic geography of the United States: from commutes to megaregions. *PLoS ONE* 11 (11): e0166083. doi: 10.1371/journal.pone.0166083.

Demšar, U., and K. Virrantaus. 2010. Space–time density of trajectories: exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science* 24 (10): 1527-1542. doi: 10.1080/13658816.2010.511223.

Demšar, U., Reades, J., Manley, E., and M. Batty. 2018. Revisiting the past: replicating fifty-year-old flow analysis using contemporary taxi flow data. *Annals of the American Association of Geographers* 108 (3): 811-828. doi: 10.1080/24694452.2017.1374164.

 Dieleman, F. M., Dijst, M., and G. Burghouwt. 2002. Urban form and travel behaviour: micro-level household attributes and residential context. *Urban Studies* 39 (3): 507–527. doi: 10.1080/00420980220112801.

Diez-Roux, A. 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. American journal of public health 88 (2): 216–222.

Diggle, P.J., and Chetwynd, A.G. 1991. Second order analysis of spatial clustering for inhomogeneous populations. *Biometrics* 47 (3):1155-1163.

Dodge, S., Weibel, R. and A.-K. Lautenschütz. 2008. Towards a taxonomy of movement patterns. *Information Visualization* 7 (3–4): 240–252. doi: 10.1057/PALGRAVE.IVS.9500182.

Doi, S., Mizuno, T. & Fujiwara, N. Estimation of socioeconomic attributes from location information. *J Comput Soc Sc* (2020). https://doi.org/10.1007/s42001-020-00073-w

Dong L., Allan, A., and J. Cui. 2015. The impacts of urban spatial structure and socio-economic factors on patterns of commuting: a review. *International Journal of Urban Sciences* 19 (2): 238-255. doi: 10.1080/12265934.2015.1016092.

Forrest, R., and A. Kearns. 2001. Social cohesion, social capital and the neighbourhood. *Urban Studies.* 38 (12): 2125-2143. doi: 10.1080/00420980120087081.

Giddens, A. 1985. Time, space, and regionalization. In *Social relations and spatial structures* ed. D. Greggory and J. Urry, 265-295. New York: St. Martin's Press.

Goulias, K.G., Blain, L., Kilgren, N., Michalowski, T., and E. Murakami. 2007. Catching the next big wave: do observed behavioral dynamics of baby boomers force rethinking of regional travel demand models?. *Transportation Research Record* 2014 (1): 67-75. doi: 10.3141/2014-09.

Goodchild, M.F., and Janelle, D.G. 1984. The city around the clock: space-time patterns of urban ecological structure. Environment and Planning A (16): 807-820.

Hagerstrand, T. 1970. What about people in regional science. Papers of the Regional Science Association 24 (1): 7–21.

Hanson, S., and P. Hanson. 1981. The travel activity patterns of urban residents: dimensions and relationship to sociodemographic characteristics. Economic Geography 57 (4): 332–347.

Hincks, S., Kingston, R., Webb, B., and C Wong. 2018. A new geodemographic classification of commuting flows for England and Wales. *International Journal of Geographical Information Science* 32 (4): 663-684. doi: 10.1080/13658816.2017.1407416.

Jaegal, Y., & Miller, H. J. (2020). Measuring the structural similarity of network time prisms using temporal signatures with graph indices. *Transactions in GIS*, *24*(1), 3-26.

Janelle, D.G., Goodchild, M.F., and B. Klinkenberg. 1988. Space-time diaries and travel characteristics for different levels of respondant aggregation. Environment and Planning A, 20 (7): 891–906.

Jia-Ching, J., Hsueh-Chan,. E., Wang-Chien, L. Tz-Chiao, W., and V.S. Tseng. 2010. Mining user similarity from semantic trajectories. In Proceedings of the *2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks (LBSN '10)*. ACM, New York, NY, USA, 19-26. doi: 10.1145/1867699.1867703.

Katarzyna Siła-Nowicka, Jan Vandrol, Taylor Oshan, Jed A. Long, Urška Demšar & A. Stewart Fotheringham (2016) Analysis of human mobility patterns from GPS trajectories and contextual information, *International Journal of Geographical Information Science*, 30:5, 881-906, DOI: [10.1080/13658816.2015.1100731](https://doi.org/10.1080/13658816.2015.1100731)

Kipnis, B.A. 2004. The value stretch model-tooling for long-time horizon lifestyle studies. *Geography Research Forum* 23: 100-114.

 Kressner, J. D. and Garrow, L. A. 2012. Lifestyle segmentation variables as predictors of home-based trips for Atlanta, Georgia, airport. *Transportation Research Record* 2266 (1): 20–30. doi: 10.3141/2266-03.

Kwan, M. 2012. The uncertain geographic context problem. *Annals of the Association of American Geographers* 102 (5): 958-968. doi: 10.1080/00045608.2012.687349.

Kwan, M. 1998. Space-time and integral measures of accessibility: A comparative analysis using a point-based framework. Geographical Analysis 30 (3): 191–216. doi: 10.1111/j.1538-4632.1998.tb00396.x.

Kwan, M. 2000. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. Transportation Research C 8 (1-6): 185–203. doi: 10.1016/S0968-090X(00)00017-6.

Laube, P., Imfeld, S., and R. Weibel. 2005. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographic Information Science* 19 (6):639-668. doi: 10.1080/13658810500105572.

Liu, K., Gao, S., Qiu, P., Liu, X., Yan, B. and F. Lu. 2017. Road2vec: measuring traffic interactions in urban road system from massive travel routes. *ISPRS International Journal of Geo-Information* 6 (11): 321. doi: 10.3390/ijgi6110321.

Liu, Y., Singleton, A., and D. Arribas-Bel. 2019. A principal component analysis (PCA)-based framework for automated variable selection in geodemographic classification. *Geo-spatial Information Science* 1-14. doi: 10.1080/10095020.2019.1621549.

Lund, Ulric, & Agostinelli, Claudio. 2007. Circular: an r package for circular statistics. Singapore : World Scientific.

Macintyre, S., Ellaway, A., and S. Cummins. 2002. Place eﬀects on health: how can we conceptualise, operationalise and measure them? *Social Science and Medicine*, 55 (1): 139-155. doi: 10.1016/S0277-9536(01)00214-3.

Mardia, K.V. 1972. Statistics of directional data. London: Academic Press.

Martin, D., Gale, C., Cockings, S., and A. Harfoot. 2018. Origin-destination geodemographics for analysis of travel to work flows. *Computers, Environment and Urban Systems* 67, 68-79. doi: 10.1016/j.compenvurbsys.2017.09.002.

McArdle, G., Demšar U., van der Spek, S., and S. McLoone. 2014. Classifying pedestrian movement behaviour from GPS trajectories using visualization and clustering. *Annals of GIS* 20 (2) 85-98. doi: 10.1080/19475683.2014.904560.

Mennis, J., Mason, M.J., and T. Cao. 2013. Qualitative GIS and the visualization of narrative activity space data. *International Journal of Geographical Information Science* 27 (2): 267-291. doi: 10.1080/13658816.2012.678362.

Miller, H. 1991. Modeling accessibility using space-time prism concepts within geographical information systems. International Journal of Geographical Information Systems, 5(3): 287–301. doi: 10.1080/02693799108927856.

Miller, H. 2003. What about people in geographic information science. Computers, Environment and Urban Systems 27 (5): 447–453. doi: 10.1016/S0198-9715(03)00059-0.

Miller, H. 2007. Place-based versus people-based geographic information science. *Geography Compass* 1 (3): 503-535. doi: 10.1111/j.1749-8198.2007.00025.x.

Pappalardo, L., Simini, F. Data-driven generation of spatio-temporal routines in human mobility. *Data Min Knowl Disc* **32,**787–829 (2018). https://doi.org/10.1007/s10618-017-0548-4

Pas, E.I. 1982. Analytically derived classifications of daily travel-activity behavior: description, evaluation, interpretation. Transportation research record 879: 9–15.

Patterson, Z., and S. Farber. 2015. Potential path areas and activity spaces in application: a review. *Transport Reviews* 35 (6): 679-700. doi: 10.1080/01441647.2015.1042944.

Pisarski, A. 2006. Commuting in America III. Transportation Research Board. Available online: https://onlinepubs.trb.org/onlinepubs/nchrp/ciaiii.pdf [Accessed 14/6/19].

Radil, S., Flint, C., and G. Tita. 2010. Spatializing social networks: using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Annals of the Association of American Geographers* 100 (2): 307-326. doi: 10.1080/00045600903550428.

 Ratti, C., Frenchman, D., Pulselli, R.M., and S. Williams. 2006. Mobile landscapes: using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design* 33 (5): 727–48. doi: 10.1068/b32047.

Richardson, D.B., Volkow, N.D., Kwan, M., Kaplan, R.M., Goodchild, M.F., and R.T. Croyle. 2013. Spatial turn in health research. *Science* 339 (6126): 1390-1392. doi: 10.1126/science.1232257.

Rohde, D., and J. Corcoran. 2015. Circular Statistics. *Geocomputation: A Practical Primer*. London: Sage, 110-135.

Salomon, I., and M. Ben-Akiva. 1982. Lifestyle segmentation in travel demand analysis. Transportation Research Record, 879: 37–45.

Sampson, R.J., 2012. Great American city: Chicago and the enduring neighborhood effect. Chicago: University of Chicago Press.

Sampson, R.J., Morenoff, J.D., and T. Gannon-Rowley. 2002. Assessing "neighborhood effects": social processes and new directions in research. Annual Review of Sociology 28 (1): 443–478. doi: 10.1146/annurev.soc.28.110601.141114.

Saxena, S., and P. Mokhtarian. 1997. Impact of telecommuting on the activity spaces of participants. Geographical Analysis29 (2)**:** 124–144. doi: 10.1111/j.1538-4632.1997.tb00952.x

Schulz, A., and M.E. Northridge. 2004. Social determinants of health: implications for environmental health promotion. *Health Education & Behavior* 31 (4): 455-471. doi: 10.1177/1090198104265598.

Sebastian T.B., Klein P.N., Kimia B.B. (2002) Shock-Based Indexing into Large Shape Databases. In: Heyden A., Sparr G., Nielsen M., Johansen P. (eds) Computer Vision — ECCV 2002. ECCV 2002. Lecture Notes in Computer Science, vol 2352. Springer, Berlin, Heidelberg

Shih-Lung Shaw & Daniel Sui (2020) Understanding the New Human Dynamics in Smart Spaces and Places: Toward a Splatial Framework, *Annals of the American Association of Geographers*, 110:2, 339-348, DOI: 10.1080/24694452.2019.1631145

Singleton, A. D., and S.E. Spielman. 2014. The past, present, and future of geodemographic research in the United States and United Kingdom. *The Professional Geographer* 66 (4): 558-567. doi: 10.1080/00330124.2013.848764.

Song, C., Qu, Z., Blumm, N., & Barabási, A. L. (2010). Limits of predictability in human mobility. *Science*, *327*(5968), 1018-1021.

Spielman, S. E., and J. C. Thill. 2008. Social area analysis, data mining, and GIS. *Computers, Environment and Urban Systems* 32(2): 110-122. doi: 10.1016/j.compenvurbsys.2007.11.004.

Stopher, P.F.,, Camden, F., Zhang, J., and T. Bretin. 2007. Analysis of 28-day global positioning system panel survey: findings on the variability of travel. Transportation Research Record 2014: 17–26. doi: 10.3141/2014-03.

Swait, J. 1994. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. Journal of Retailing and Consumer Services 1 (2): 77-89. doi: 10.1016/0969-6989(94)90002-7.

Webber, R., and Burrows, R. 2018. The predictive postcode: the geodemographic classification of British society. London: SAGE.

Weber, J., and M. Kwan. 2003. Evaluating the effects of geographic contexts on individual accessibility: a multilevel approach. *Urban Geography* 24 (8): 647-671. doi: 10.2747/0272-3638.24.8.647.

Wellman, B., Boase, J., and W. Chen. 2002. The networked nature of community: online and offline. *IT & Society* 1(1): 151-165.

Zablocki, B.D., and R.M. Kanter. 1976. The differentiation of lifestyles. Annual Review of Sociology, 2 (1): 269–298. doi: /10.1146/annurev.so.02.080176.001413.