Annals of GIS
Vol. 18, No. 4, December 2012, 245–255

INVITED ARTICLE

How GIS can help address the uncertain geographic context problem in social science research

Mei-Po Kwan[a,b]*

[a]*Department of Geography, University of California, Berkeley, CA, USA;* [b]*Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China*

(*Received 26 July 2012; final version received 4 September 2012*)

The uncertain geographic context problem (UGCoP), first articulated by Kwan (2012; The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102 (5), 958–968), refers to the problem that findings about the effects of area-based contextual variables on individual behaviors or outcomes may be affected by how contextual units (e.g., neighborhoods) are geographically delineated and the extent to which these areal units deviate from the true geographic context. It is a significant methodological problem because it means that analytical results can differ for different delineations of contextual units even if everything else is the same. Drawing upon Kwan (2012) and recent social science studies (especially environmental health and neighborhood effects research), this article further elaborates on the nature of the UGCoP and explores how recent advances in geographical information system (GIS) and geospatial technologies can help address the problem. It discusses possible means for mitigating the UGCoP, especially with regard to the collection of detailed individual space–time data with global positioning systems, construction of individual activity spaces, and the use of qualitative and web-based GIS to capture people's activity locations and everyday experiences. It also discusses the challenges for future research that seeks to address the UGCoP.

Keywords: geographic context; UGCoP; environmental health; neighborhood effects; uncertain geographic context problem

1. Introduction

An important question in social science research is how the characteristics of place affect people's behaviors or experiences (e.g., voting behavior, racial or ethnic segregation, and utilization of health services). When a study uses area-based variables to explain or predict individual behaviors or outcomes, as in cases where data based on census tracts or blocks are used, it faces two fundamental methodological problems. One of these problems is much better understood than the other, which is the modifiable areal unit problem (MAUP). This well-known problem refers to the fact that analytical results about the effects of area-based variables may be affected by the zoning scheme and/or geographic scale of the areal units used (Openshaw 1984). The MAUP has received much attention by social science researchers to date, and considerable work has been conducted to address it (Fotheringham and Wong 1991, Cressie 1996, Holt *et al*. 1996, Hipp 2007, Oliver and Hayes 2007). Past studies were largely motivated by Openshaw's (1996) view that the MAUP should be dealt with through identifying and adopting the best zoning scheme and/or geographic scale at which the processes being studied operate. They focused on deriving the best areal division, neighborhood size, or geographic scale for the data and study area (MacAllister *et al*. 2001, Johnston *et al*. 2005, 2007, Weiss *et al*. 2007, Flowerdew *et al*. 2008, Mu and Wang 2008, Riva *et al*. 2009, Root 2012).

The other fundamental methodological problem for any social science study that examines the effects of geographic variables on individual behaviors or outcomes has received much less attention to date. While there have been concerns about this second problem in research on the MAUP (e.g., Manley *et al*. 2006) and neighborhood effects on health (e.g., Diez Roux 2001), explicit discussion on its nature and methodological implications has been rather limited. This second problem refers to the fact the analytical results about the effects of geographic variables on the outcome variables may be affected by the precise geographic delineations of contextual units or neighborhoods and the deviation of the contextual units from the true geographic context.

I provided the first explicit discussion about this problem in Kwan (2012) and referred to it as the uncertain geographic context problem (UGCoP). In that paper, I argued that the UGCoP arises because of the spatial uncertainty in the actual areas that exert relevant contextual influences on each of the individuals under study and the temporal uncertainty in the timing and duration in which

*Email: mpk654@gmail.com

ISSN 1947-5683 print/ISSN 1947-5691 online

http://dx.doi.org/10.1080/19475683.2012.727867
http://www.tandfonline.com

these individuals experienced the contextual influences (also see Lam (2012) for a detailed discussion on the various types of uncertainties encountered in environmental health research). I suggested that the UGCoP is a problem as fundamental as the MAUP for any study that uses area-based contextual variables, but it is a different kind of problem because it is not due to the use of different zonal schemes or spatial scales for area-based variables. Instead it is due to the fact that the precise spatial and temporal configuration of the true contextual area is not known in most, if not all, social science studies, and that analyses that use data based on administrative units (e.g., census tracts) have no *a priori* reason to be able to accurately capture an individual's true geographic context or to accurately estimate values of the contextual variables. I emphasized that methods for addressing the MAUP such as finding the best areal division, neighborhood size, aggregation scheme, or geographic scale for the study area and population do not automatically solve the UGCoP (which arises because of our limited knowledge about the precise spatial and temporal configuration of each individual's true geographic context, not because of the use of a particular scheme of areal division, zonal aggregation, or spatial scale). I called for explicit attention to be paid to the potential confounding effects of the problem on research results and to methods for mitigating it in future research.

Drawing upon Kwan (2012) and recent social science studies (especially environmental health and neighborhood effects research), this article further elaborates on the nature of the UGCoP and explores how recent advances in geographical information system (GIS) and geospatial technologies can help address the problem. It discusses possible means for mitigating the UGCoP, especially with regard to the collection of detailed individual space–time data with global positioning systems (GPS), construction of individual activity spaces, and the use of qualitative and web-based GIS to capture people's activity locations and everyday experiences. It also discusses the challenges for future research that seeks to address the UGCoP. However, it is important to emphasize that the use of GIS and geospatial technologies in addressing the UGCoP is necessitated by the very nature of the problem, which arises because of the complex spatial and temporal configuration of people's true geographic context. As shown in recent studies (e.g., Gulliver and Briggs 2005, Wiehe *et al*. 2008, Duncan *et al*. 2009), these technologies can help researchers to better capture the complex movements of people and the spatiotemporal dynamics of environmental influences, which constitute the theoretically sound basis for delineating context units in many types of studies (e.g., how the physical features of neighborhoods affect people's physical activities). No previous methods can accomplish this to the extent that is allowed by GIS and geospatial technologies. Proposing to use and the actual use of advanced geospatial technologies to address the UGCoP thus does not mean privileging technical solutions over sound conceptual or theoretical justifications concerning why particular delineations are good approximations of the true geographic context.

2. Geographic context and the UGCoP

An important focus in social science research is the effects of place, neighborhood, or geographic context on people's behaviors or experiences. This literature is by now enormous. Some examples are the studies of MacAllister *et al*. (2001) and Johnston *et al*. (2007) on voting behavior, Hipp (2007) and Browning *et al*. (2010) on crime, and Sampson *et al*. (1997) on social organization. Much of this literature has been articulated in the rubric of neighborhood effects (e.g., Galster 2001, Dietz 2002, Sampson *et al*. 2002, Kawachi and Berkman 2003, van Ham *et al*. 2012). For instance, Johnston *et al*. (2007) examined contextual influences at various spatial scales (individuals, households, neighborhoods, constituencies, and regions) on people's voting behavior in the United Kingdom. The study observed that when all of the smaller-scale level (individual, household, constituency) characteristics are taken into account, regional variations ceased to be significant. It concluded that contextual influences on people's voting behavior operate at much smaller spatial scales, and regional patterns of voting behavior only emerge as a result of analyses that used highly aggregate data.

A considerable amount of the literature on the effects of neighborhood and geographic context on people's behaviors or experiences concerns their influences on health (e.g., Curtis and Jones 1998, Diez Roux 1998, 2001, Kawachi and Berkman 2003). Studies by health researchers and geographers in the past two decades or so has shed new light on many significant conceptual and methodological issues pertaining to the role of place, context, and neighborhood effects on health. Moving beyond the notion that health behaviors and outcomes can be explained exclusively in terms of the characteristics of individuals or their households, researchers have re-established the centrality of place or neighborhood as a significant component in explanations of health behaviors and outcomes. It is now widely recognized that geographic variations in health cannot be explained exclusively in terms of the characteristics of individuals, as specific characteristics of place or neighborhoods also exert significant influence on health.

Past research on environmental health has identified various characteristics of place that are important in affecting individual behavior or outcomes (e.g., Macintyre 1997, Pickett and Pearl 2001). These characteristics can be described in terms of two broad groups of factors: (1) physical or environmental features, and (2) social, cultural, and institutional characteristics of neighborhoods. Physical features that affect health include the availability of health-promoting environments (such as sidewalks, trails, and

parks) and the accessibility of health services and grocery stores selling healthful foods at affordable prices. These are largely material resources or 'opportunity structures' that are socially and spatially patterned features of the environment which may 'promote or damage health either directly, or indirectly through the possibilities they provide for people to live healthy lives' (Macintyre *et al*. 2002, p. 132). Social and cultural features of neighborhoods that affect health include the prevailing attitudes toward health and health-related behaviors, local social ties and cohesion, collective efficacy, neighborhood institutions, crime, and structural dimensions of neighborhood disadvantage such as residential segregation and social isolation of particular population groups. These characteristics may be referred to as 'collective social functioning and practices' (Macintyre *et al*. 2002, p. 132).

In past research, the effects of these physical and social environmental factors on a variety of behaviors and phenomena are conceptualized as neighborhood or contextual effects (e.g., Sampson *et al*. 2002, Kawachi and Berkman 2003). However, results about the influences of contextual effects in these studies are often inconsistent (e.g., Inagmi *et al*. 2007, Wilks *et al*. 2010). One important reason for this is that past studies on the same issue (e.g., obesity or physical activity) often used different contextual units. For instance, studies on the effect of neighborhood features such as land-use mix and residential density on people's physical activity or body weight have defined neighborhood around each participant's home as a 1-km road network buffer (Frank *et al*. 2005), as a 1- or 3-km circular zone (Berke *et al*. 2007), as a 0.5-mile radius or a 10-min walk from the respondent's home for some variables, and as a 10-mile radius or a 20-min drive from the respondent's home for several other variables (Brownson *et al*. 2004). But it is far from clear which of these areal units appropriately represent the areal extent and spatial configuration of the true geographic context. The mixed results of past studies on neighborhood effects (e.g., neighborhood income inequality and racial composition) on health (e.g., obesity) may thus be partly due to the different neighborhood delineations used.

It is important to note that researchers normally have little prior knowledge about the precise spatial configuration and boundary of the geographic area that exerts significant social and physical influence on the phenomenon under study. Since the 'true causally relevant' geographic context is unknown in most studies, a common practice is to use residential neighborhoods as contextual units (Diez Roux and Mair 2010, p. 134). These units are often operationalized as static administrative areas such as census tracts or postcode areas, or buffer areas around individuals' home addresses or centroids of their home census tracts.

However, residential neighborhoods may not accurately represent the actual areas that exert contextual influences on the individual behavior or experience under study (Cook 2003, Cummins 2007, Matthews 2008, 2011, Chaix 2009, Kwan 2009). For example, adolescent risk behavior like substance use may be affected not only by socioeconomic deprivation in the residential neighborhood but also by interactions with friends and peers in various non-residential contexts (e.g., schools and places for various leisure activities). "The boundaries of these multiple contexts are often difficult to clearly delineate. Even when they can, some of them may not be continuous in geographic space (i.e., one contextual unit may consist of several discrete geographic areas) and thus cannot be represented or analyzed in any simple manner" (Wiehe *et al*. 2008, Kwan 2012, p. 960). These multiple contexts also may not be organized hierarchically or have a nested structure and thus cannot be dealt with using multilevel models. Further, social contexts and networks such as families, friends, or peers are not explicitly defined in geographic terms and thus cannot be easily delineated as geographic areas with precise boundaries (Diez Roux 2001). In others cases, "neighborhoods defined on the basis of people's perceptions may be more relevant. However, the perceived neighborhood for different individuals may not coincide with or may even deviate significantly from the administratively defined home neighborhood or people's activity space" (e.g., Vallée *et al*. 2010, Kwan 2012, p. 960).

These difficulties mainly contribute to the spatial uncertainty in the actual areas that exert contextual influences on the phenomenon under study. The dynamics of contextual influences and the movement of people, on the other hand, lead to considerable temporal uncertainty in contextual influences (Chaix 2009, Kwan 2009, 2012, Gatrell 2011). For instance, certain contextual influences may vary over space and time in a highly complex manner (e.g., traffic-related air pollution). They may vary with different temporal patterns or time frames. As people move through the changing pollution field over time during the day, for instance, their exposures to these environmental influences also change (Gulliver and Briggs 2005; see Figure 1). Further, some environmental influences may change over the 24-h period of a day (e.g., pollutants from truck traffic), while some may change over the seasons. The physical and social characteristics of neighborhoods may also change over time (Entwisle 2007). Population composition and local social ties may change as a result of residential mobility and migration. When environmental or neighborhood influences have considerable spatial and temporal variability, their influences on people often cannot be adequately assessed using data for just one time point (Setton *et al*. 2010).

Further, past geographic and activity–travel behavior research has shown that people move around to undertake their daily activities (Hanson and Hanson 1981, Burnett and Hanson 1982, Kwan 1999, 2000). They often traverse the boundaries of multiple neighborhoods during the course of a day and come under the influence of many

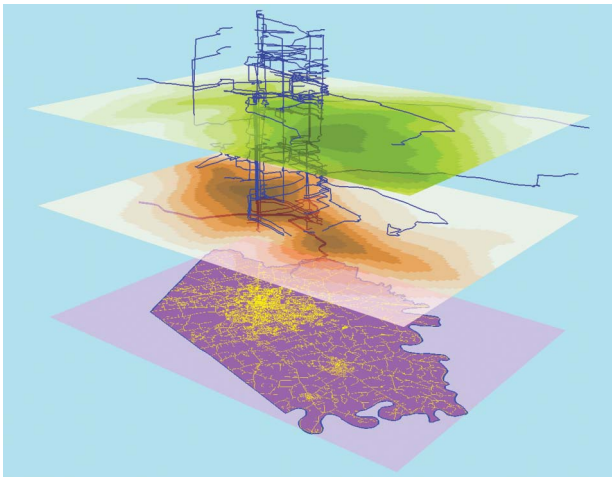

Figure 1. Certain contextual influences may vary over space and time in a highly complex manner (e.g., traffic-related air pollution). As people move through the changing pollution field over time during the day, their exposures to these environmental influences also change. GIS and detailed GPS data can be used to more accurately assess people's exposures to these environmental risk factors (after Gulliver and Briggs 2005).

different neighborhood contexts besides their residential neighborhoods (Sampson *et al*. 2002, Matthews 2011). This means that people's activities (and thus exposures to environmental influences) do not take place at one time point and wholly within any conventionally defined neighborhood. Their use of different physical resources and their social interactions with friends, peers, and others may take place at different times of the day and in disparate geographic areas outside of their home neighborhoods (Kwan 2009). The neighborhood of residence is thus only one of the places people spend their time, and it may not adequately capture people's exposure to relevant contextual influences. Further, besides moving around to undertake their daily activities, people also move around over time. They may change their residence in the same city (residential mobility) or move to another (migration). As a result of moving to different neighborhoods, people's exposure to environmental influences may also change over time.

3. Recent evidence on the UGCoP

It is important to emphasize that the UGCoP arises because of spatial and temporal uncertainties in the contextual influences that affect where, when, and for how long people experienced these influences. The fundamental source of the problem is that in most, if not all, studies the true geographic context for a particular person is unknown because of its complex spatial and temporal configuration. Researchers can use sound conceptual frameworks and methods to better approximate the true context and to mitigate the UGCoP, but it is not clear when one can validly claim that the problem has been solved.

Unlike addressing the MAUP, where the primary analytical task is to identify the best areal division, neighborhood size, or geographic scale as well as to compare the effects of difference zoning schemes and scales on analytical results, addressing the UGCoP involves more accurately measuring and estimating the 'true causally relevant' geographic context (Diez-Roux and Mair 2010). Further, it may be difficult to fully appreciate the significance of the problem if one adheres to the conventional spatial frameworks of area-based data (even using very small enumeration units) or when these data are the only available data for studying a particular social issue.

The UGCoP is a significant methodological problem because it means that analytical results can be different for different delineations of contextual units even if everything else is the same (Kwan 2012). It is perhaps an important reason why research findings concerning the effects of social and physical environments on human behaviors and experiences are often inconsistent. Recent research provides evidence that illuminates various aspects of the UGCoP. Concerning the issue that people move around in their daily lives and that the residential neighborhood may not capture a significant portion of their exposure to various contextual influences, recent studies that collected detailed data about people's out-of-home activities and travel routes using GPS or other location-aware devices provide some important clues about the UGCoP. Basta *et al*. (2010), for instance, observed that half of the sampled participants (15–19 years old) spent 92% of their time outside of their residential neighborhoods. Elgethun *et al*. (2003), on the other hand, found that participants (children 2–8 years old) on average spent most of their time inside schools on weekdays, while spending most of their time in establishments like restaurants and cinemas on weekend days.

In a study of the risk behavior of female adolescents, Wiehe *et al*. (2008) found that "participants spent one-third of their time in locations more than 1 km from home, which is the distance used in many previous studies for defining neighborhood. This means that the participants spent a considerable amount of time in their daily life outside of what has conventionally been defined as geographic context or neighborhood. The study also found considerable day-to-day variability in participants' activity locations besides their variability by time of day" (Wiehe *et al*. 2008, Kwan 2012, p. 961). These studies reveal the daily and day-to-day variability in human activity locations and raise serious concerns about using conventional static contextual units in health and social science research.

Other studies have shed light on how the misspecification of contextual units or inappropriate temporal characterization of the contextual influence may confound research results. For instance, Kwan *et al*. (2009) "found significant differences in the size and shape of three different delineations of geographic context: two delineations of activity space (the standard deviational ellipse and the

kernel density surface) and the home census tract. The study observed that for certain gender and racial groups, neighborhood effects based on people's home census tracts tend to overestimate their actual exposure to social disadvantage (because characteristics of the non-residential neighborhoods people visit may mitigate the disadvantage they experience in their residential neighborhood)" (Kwan 2012, p. 963).

In a study on the association between physical features within 1-km road network buffers of participants' home and workplace and the amount of physical activity, Troped *et al*. (2010) found that physical features that are associated with participants' physical activity around their home and their workplace are different, and none of the physical features showed associations with participants' total physical activity. The study not only shows that people's physical activity may vary according to where they are, but it also suggests that a study that uses only participants' home neighborhood as the contextual unit may not find any association between its physical features and participants' body mass index,[1] because body mass index depends on total physical activity, not just activity around one's home or workplace.

Other studies have provided more direct evidence to show that both contextual variables and research results are sensitive to the choice of contextual units (e.g., Zenk *et al*. 2011). Kwan *et al*. (2011), for example, observed significant difference between the composite deprivation index[2] (as a contextual variable) derived from circular buffers around participants' home addresses and those derived from half-mile road network buffers around participants' GPS tracks. With respect to research results, Oliver *et al*. (2007) found that "the use of different kinds of buffers around participants' home (based on centroids of their home postal codes) as contextual units has a considerable influence on the results: land-use characteristics tend to show greater associations with walking using line-based road network buffers than circular buffers; circular and polygon buffers tend to underestimate the effects of land-use characteristics on walking because they may include large areas that are irrelevant to walking (e.g., industrial land) or inaccessible. These studies indicate that both contextual variables and study results are sensitive to the choice of contextual units" (Kwan 2012, p. 964).

4. Relevance and limitations of past studies

As argued in the last section, the UGCoP arises because of spatial and temporal uncertainties about where, when, and for how long people experienced relevant contextual influences. The fundamental source of the problem is that in most studies researchers do not have perfect knowledge about the true geographic context because of its complex spatial and temporal configuration (cf Couclelis 2003, Shi 2010, Lam 2012). A helpful first step toward mitigating the influence of the UGCoP on contextual variables and research results is the construction of a conceptual model that clearly specifies the causal pathways among the contextual and outcome variables. Contextual units can then be constructed based on such conceptual model. For instance, in a study of the factors that influence people's smoking cessation behavior, Sorensen *et al*. (2004) provide an elaborate socio-contextual model that specifies the important neighborhood or community factors that influence people's smoking behavior (e.g., social deprivation and social cohesion).

Based upon a sound conceptual model that clearly specifies the causal pathways among contextual and outcome variables, researchers can identify appropriate methods that can be used to better approximate the true geographic context for the individuals being studied. The *spatial framework* in which data are organized or provided and the level of analysis (individuals or area-based groups) have significant implications for what kind of uncertainty contributes to the UGCoP and how it may be addressed. It is important to note that the arguments about the UGCoP in this article (as well as those in Kwan 2012) pertain mainly to studies in which area-based contextual variables (e.g., neighborhood poverty and racial segregation) are used to explain or predict individual behaviors or outcomes (e.g., individual decision about using particular health services). In this kind of studies, the contextual variables and the outcome variables are in different spatial frameworks, the former are area-based while the latter are individual-based.

Much of social science research, however, is concerned with the relationships between area-based contextual variables and area-based outcome variables (e.g., cancer or crime rates of census tracts). This kind of research is useful in that it helps us identify the possible relationships between contextual influences and particular social or health phenomena (e.g., low birthweight rates) in different areas or for different social groups using aggregate area-based data, which are often the only data available. Some of these studies also explored the effect of different delineations of contextual units on research findings and developed different means for addressing the MAUP and the small numbers problem, largely through identifying and using the best aggregation scheme, neighborhood size, or geographic scale for the study area (e.g., Hipp 2007, Root 2012, Wang *et al*. 2012). For instance, Guo (2008) and Guo and Wang (2011) developed and refined a method (called the recent group of regionalization method) for generating larger homogeneous areas from smaller ones that mitigate the small numbers problem (the problem that statistical estimates tend to be unstable for small enumeration units with small populations). Mu and Wang (2008) developed a modified scale-space clustering method that merged smaller areal units into larger ones to mitigate the effect of scale and spatial autocorrelation on regression estimates. Wang *et al*. (2012) developed an automated

method for constructing larger areas that are spatially and socioeconomically closer than arbitrary areal units such as census tracks or post code areas. The method was used to overcome the small numbers problems and was applied to analyze late-stage breast cancer risks in Illinois in 2000.

These studies have contributed in important ways to the development of new methods for identifying better aggregation schemes or spatial scales in studies involving area-based outcome variables and thus are helpful for mitigating the MAUP. While the analytical focus of this kind of studies is group behaviors, decisions, or outcomes (based on area-based data), it is not clear how group-based geographic context can be conceptualized given the spatial and temporal uncertainties of the true geographic context for each individual in a group. Since individuals who live in the same areal unit (e.g., census tract or block) may experience contextual influences from many different areal units besides their home neighborhoods (since they may visit many different places in their daily lives), it is not clear how a true causally relevant geographic context may be meaningfully conceptualized or delineated for an area-based group (e.g., for all the individuals who live in the same census tract). In this light, it remains unclear how attempts to mitigate the MAUP are relevant to addressing the UGCoP, although they are helpful in reducing statistical biases due to the confounding effects of geographic scale and zonal aggregation.

More relevant to addressing the UGCoP are past studies that dealt with particular aspects of the problem by using multilevel models or delineating person-specific neighborhood or contextual units using small-area data (e.g., MacAllister *et al*. 2001, Johnston *et al*. 2005, 2007). Given that contextual influences on individual behaviors or outcomes often operate in a multilevel and multiscale fashion, multilevel models may help researchers capture these influences at multiple levels if these levels are hierarchically organized or nested (e.g., region, county, census tracts). For instance, Mobley *et al*. (2008) examined how contextual variables at different levels (geographic scales) affect women's mammography use decision. The study built contextual variables from fine-grained data based on four levels of contextual units (counties, medical service study areas, primary care service areas, and post code areas). It identified significant contextual influences that operate at the local level (e.g., residential segregation) and concluded that studies using only county-level factors will miss important relationships at the local level. MacAllister *et al*. (2001) constructed person-specific neighborhoods using small-area census data to approximate the true geographic context for each participant of the 1997 British Election Study (BES). Based on each participant's postcode and enumeration district and using enumeration districts as building blocks, bespoke neighborhoods for every participant containing their nearest 500, 1000, 2500, 5000, and 10,000 neighbors were constructed with a customized algorithm. The results, as the authors argued, provided strong evidence on the existence of classic neighborhood effects at small spatial scales.

While these efforts are relevant to addressing the UGCoP, they have limitations that call for the development of new conceptualizations of geographic context and better methods for delineating them in future research. First, all of these studies were still based on data that are tied to arbitrary administrative areas. No matter how small the enumeration units are or how these units are combined to create 'better' contextual units, it is not clear why and how each person's true geographic context can be approximated based on *a priori* assumptions about the appropriate size of neighborhood units. Since people have different spatial mobility, and their activity–travel patterns and social interactions may unfold in space and time in a highly complex manner, there may be considerable differences in the size and shape of their true geographic contexts (e.g., Kwan *et al*. 2008, Shoval *et al*. 2011). Second, when individual outcomes are heavily affected by people's travel routes or movements in space–time (e.g., exposure to traffic-related pollution), and when the temporal characteristics of contextual influences such as duration or cumulative effects are important (e.g., exposure to carcinogenic substances), it will be difficult to create good approximations of people's true geographic contexts without knowing where they went, how much time they spent there, and their travel routes during the relevant time period. Third, multilevel models are helpful only when the multiple levels of contextual influences in question can be hierarchically organized or have a nested structure. But research has observed that a person's neighborhood may consist of several discrete areas that cannot be organized hierarchically (e.g., Wiehe *et al*. 2008). Lastly, while individuals who live in the same household may be exposed to contextual influences from many different geographic areas and thus should have their own person-specific contextual areas, the methods used in these studies do not allow individuals who live in the same household to have different contextual units.

5. Using GIS to address the UGCoP

To overcome these limitations and to address the UGCoP more effectively, we need new conceptualizations of geographic context that can take into account the complex spatial and temporal configuration of individual context, as well as new analytical methods for delineating these contextual units. Drawing upon insights from time geography, Kwan *et al*. (2008) and Kwan (2009, 2012) proposed a *dynamic conceptualization of geographic context* that is individual-based, person-specific, and delineated based on where people go, how much time they spend there, and their travel routes. Hipp and Boessen (2012) also adopted a person-specific approach and developed the notion of *egohood* as the basis for delineating the contextual area

for each individual. The idea of egohood describes the fact that each individual perceives themselves as the center of their neighborhoods and social networks. New conceptualizations of geographic context or neighborhood like these seem useful for informing future research that examines the effects of area-based variables on individual behaviors or outcomes. Implementing these new conceptualizations, however, calls for the development and use of new methods (e.g., Lee *et al*. (2008) developed a method for delineating circular egocentric local environments for each research participant that are not tied to administrative units). Two promising areas involving the use of GIS and geospatial technologies to address the UGCoP are described as follows: (1) GIS, GPS, and location-aware mobile devices, and (2) qualitative GIS and web-based GIS.

5.1. GIS, GPS, and location-aware mobile devices

GIS and GPS technologies are particularly helpful for addressing the UGCoP because they can help researcher to better capture the complex spatial and temporal configuration of people's true geographic context, largely through collecting and analyzing data of people's movements and the spatiotemporal dynamics of environmental influences. For instance, using GIS to delineate people's activity space seems promising for addressing the UGCoP (e.g., Rainham *et al*. 2008, 2010). A person's activity space is the area in which routine daily activities and trips are undertaken (Arcury *et al*. 2005). This space can be used to capture where and how much time people spend in a study area. Sherman *et al*. (2005), for instance, implemented the standard deviational ellipse and three road network-based delineations of activity space using GIS: the road network buffer, the 30-min standard travel time polygon, and the relative travel time polygon. In a study using the Los Angeles Family and Neighborhood Survey (LA FANS) data set, Kwan *et al*. (2009) implemented two delineations of activity space (the standard deviational ellipse and the kernel density surface) using GIS in addition to the home census tract to examine the effect of neighborhood socioeconomic deprivation in Los Angeles on participants' body weight. Activity spaces delineated in these studies are actual activity spaces. They were constructed with data about people's actual activity locations and GIS procedures such as the standard deviational ellipse. However, when data about people's space–time constraints are available (e.g., collected with activity–travel diary surveys), these data can be used to construct potential activity spaces using dedicated geocomputational algorithms (e.g., Kwan 1999). This latter notion of activity space, also known as the daily potential path area in the parlance of time geography, is less commonly applied in social science research when compared to actual activity space because of its data and geocomputational requirements.

While activity surveys like the LA FANS provide useful data for delineating people's activity spaces, information about the timing of people's activities and the routes they used to travel between activity locations is very limited. It is difficult to accurately estimate how much time people spent at each location using these data, and as a result, environmental exposures evaluated based on people's activity spaces may be inaccurate. To overcome this limitation, detailed space–time data of people's activities can be collected using GPS or other location-aware devices such as mobile phones (e.g., Rainham *et al*. 2008, Wiehe *et al*. 2008, Shoval *et al*. 2011). Zenk *et al*. (2011), for instance, use GPS to track participants' movement over a 7-day study period. The GPS recorded participants space–time coordinates every 30 s. The data were used to construct two delineations of participants' activity spaces: one was constructed as a standard deviational ellipse, and the other was delimited as a potential daily path area.

Because of their high spatial and temporal resolutions, GPS data allow researchers to perform 3D geovisualizations of people's space–time paths and construct more relevant contextual areas using various notions of activity space (Kwan 2000, 2004). Based upon a time-geographic framework, Rainhaim *et al*. (2008), for example, collected and used GPS data to delineate participants' activity space using the standard deviational ellipse, which captures the geographic distribution or directional trend of a series of points (Yuill 1971, Wong and Lee 2005). Kwan *et al*. (2011) explored how GPS data can be used to construct more relevant contextual units in a study of smokeless tobacco usage in the Appalachian regions of Ohio. Using GPS data collected from participants, different delineations of geographic context (e.g., potential path area and time-weighted activity space) were implemented to derive values for the contextual variables (e.g., pro-tobacco advertisement and socioeconomic deprivation). These delineations also included road network buffers with different width around participants' GPS tracks.

Another area where GIS can be particularly helpful for addressing the UGCoP is the accurate assessment of people's exposure to environmental influences. Important tasks for such assessment include measuring the spatiotemporal variation of environmental influences (e.g., airborne pollutants) and identifying when individuals are affected by them given their movement in space–time. While these are highly complex tasks, GIS can help capture the spatiotemporal dynamics of contextual influences and detailed space–time trajectories of individuals and integrate all the data into a suitable analytical framework. Gulliver and Briggs (2005), for instance, collected 24-h activity diary data from participants and constructed a space–time exposure modeling (STEM) method to evaluate their journey-time exposure to traffic-related pollution. The model actually integrated four different smaller submodels within a GIS. Each of these submodels deals with a specific component of exposure assessment (e.g., an air pollution dispersion model).

5.2. Qualitative and web-based GIS

Using GIS to capture people's social interactions and experiences in different contexts may also help address the UGCoP. A promising area in this regard is qualitative GIS and mixed methods that integrate both qualitative and quantitative information. Matthews *et al*. (2005), for instance, developed a qualitative GIS method called geo-ethnography for recording and capturing the complexity in people's daily activities and the places they visit. In the study, family and neighborhood ethnographies are integrated in a GIS, which allowed researchers to visualize and better understand the complexity of participants' lives.

In an ongoing project, I and my collaborators use dynamic context and exposure measures to study the health risk of female sex workers at the US–Mexico border through capturing and analyzing their space–time trajectories and fears through a mixed-method approach. The project uses the geo-narrative approach for illuminating the complex sociogeographic context of the participants (Kwan and Ding 2008). Based upon the general principles of narrative inquiry, geo-narrative is particularly useful for discovering and preserving the experiences of ordinary people and for studying the hidden histories, lives, and memories of disadvantaged people. The project seeks to integrate the spatial and temporal dimensions as well as participants' feelings and emotions using qualitative data analysis capabilities originally built within a GIS. Other qualitative GIS approaches such as the grounded visualization approach will also be useful for uncovering how people's social contexts may affect their behaviors and experiences (e.g., Knigge and Cope 2006).

Researchers have also explored the use of web-based mapping tools and GIS in delineating more relevant contextual areas or neighborhoods. Chaix *et al*. (2012), for instance, have developed a web-based tool called VERITAS that integrates various Google Maps interactive mapping functionalities. The tool was used to record the frequency of participants' visits to their destinations and the extent to which they feel attached to their residential neighborhood and other activity places (e.g., workplace, transport mode, and recreational activities). It also allows researchers to geocode and visualize participants' activity locations, and to delineate their perceived or experienced neighborhoods. The method is particularly useful for capturing non-residential activity locations people visit in their daily lives and thus can help research identify the true geographic context.

6. Conclusion

This article argued that the UGCoP is a significant methodological problem because it means that analytical results can differ for different delineations of contextual units even if everything else is the same. Drawing upon Kwan (2012) and other relevant literature, it further elaborated on the nature of the UGCoP and explored how GIS and geospatial technologies can help researchers address the problem. It also suggested that new conceptualizations of geographic context that take into account the actual spatial and temporal configuration of contextual influences would enable us to assess the effects of these influences (or neighborhood effects) more accurately for each individual subject. Because where and when people spend their time differ from individual to individual, these new notions of context need to be operationalized through individualized measures that allow the contextual unit or exposure level to vary even for individuals within the same neighborhood or household. As discussed in this article, data collection with GPS- and GIS-based analytical methods can be very useful for mitigating the impact of the UGCoP.

But there are many challenges that future research that seeks to address the UGCoP will face. First, the collection and integration of vast amounts of complex spatial and temporal data in GIS are challenging. This is the big data challenge that concerns the development and analysis of massive spatiotemporal databases, an area that has attracted much attention lately (Wang 2010). Second, addressing the UGCoP would require the development of sophisticated dynamic methods for measuring contextual exposures, characterizing human movement, and characterizing environmental or contextual influences (e.g., geocomputation and geovisualization methods). Significant advances in these three areas will be necessary for developing methods for the appropriate delineations of contextual units or neighborhoods. Further, as detailed geographic and temporal data may bear considerable risks of revealing the identity of individuals, protecting personal privacy and preserving data confidentiality will also be a significant challenge for future research on using GIS and geospatial technologies to address the UGCoP.

Acknowledgments

Earlier versions of this article were presented as a keynote address at the 20th International Conference on Geoinformatics (Hong Kong, June 2012) and at the NSF-sponsored Symposium on Enabling a National Geospatial Cyberinfrastructure for Health Research (San Diego, July 2012). I thank Mike Goodchild, Sara McLafferty, and the anonymous reviewers for their helpful comments. The papers sent to me by Basile Chaix, Ron Johnston, and Stephen Matthews have also helped improve this article considerably. I thank them and many others for their interest in the UGCoP and their thoughtful suggestions. This article was written while I was supported by the following grants: NIDA R01CA129771 and NIDA R01DA028692.

Notes

1. Body mass index (BMI) is a measure based on a person's weight and height. It provides a good indicator of body fatness for most people and is often used to classify people into different weight categories for screening for health problems.

2. A composite deprivation index is a single measure that was used to capture the multiple dimensions of social deprivation, which is the condition of social disadvantage experienced by the individuals who live in particular neighborhoods. A composite deprivation index is usually constructed as a standardized weighted sum of several individual deprivation scores that measure conditions like neighborhood housing, employment, socioeconomic status, and availability of social services.

References

Arcury, T.A., *et al*., 2005. The effects of geography and spatial behavior on health care utilization among the residents of rural region. *Health Services Research*, 40 (1), 135–155.

Basta, L.A., Richmond, T.S., and Wiebe, D.J., 2010. Neighborhoods, daily activities, and measuring health risks experienced in urban environments. *Social Science & Medicine*, 71 (11), 1943–1950.

Berke, E.M., *et al*., 2007. Association of the built environment with physical activity and obesity in older persons. *American Journal of Public Health*, 97 (3), 486–492.

Browning, C.R., *et al*., 2010. Commercial density, residential concentration, and crime: land use patterns and violence in neighborhood context. *Journal of Research in Crime and Delinquency*, 47, 329–357.

Brownson, R.C., *et al*., 2004. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. *American Journal of Public Health*, 94 (3), 473–483.

Burnett, P. and Hanson, S., 1982. The analysis of travel as an example of complex human behavior in spatially-constrained situations: definition and measurement issues. *Transportation Research A*, 16 (2), 87–102.

Chaix, B., 2009. Geographic life environments and coronary heart disease: a literature review. Theoretical contributions, methodological updates, and a research agenda. *Annual Review of Public Health*, 30, 81–105.

Chaix, B., *et al*., 2012. An interactive mapping tool to assess individual mobility patterns in neighborhood studies. *American Journal of Preventive Medicine*, 43 (4), 440–450.

Cook, T.D., 2003. The case for studying multiple contexts simultaneously. *Addiction*, 98 (Suppl. 1), 151–155.

Couclelis, H., 2003. The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS*, 7 (2), 165–175.

Cressie, N., 1996. Change of support and the modifiable areal unit problem. *Geographical Systems*, 3, 159–180.

Cummins, S., 2007. Investigating neighborhood effects on health – avoiding the 'local trap'. *International Journal of Epidemiology*, 36, 355–357.

Curtis, S. and Jones, I.R., 1998. Is there a place for geography in the analysis of health inequality? *Sociology of Health & Illness*, 20 (5), 645–672.

Dietz, R.D., 2002. The estimation of neighborhood effects in the social sciences: an interdisciplinary approach. *Social Science Research*, 31 (4), 539–575.

Diez Roux, A.V., 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health*, 88 (2), 216–222.

Diez Roux, A.V., 2001. Investigating neighborhood and area effects on health. *American Journal of Public Health*, 91 (11), 1783–1789.

Diez Roux, A.V. and Mair, C., 2010. Neighborhood and health. *Annals of the New York Academy of Sciences*, 1186, 125–145.

Duncan, M.J., Badland, H.M., and Mummery, W.K., 2009. Applying GPS to enhance understanding of transport-related physical activity. *Journal of Science and Medicine in Sport*, 12 (5), 549–556.

Elgethun, K., *et al*., 2003. Time–location analysis for exposure assessment studies of children using a novel global positioning system instrument. *Environmental Health Perspectives*, 111 (1), 115–122.

Entwisle, B., 2007. Putting people into place. *Demography*, 44 (4), 687–703.

Flowerdew, R., Manley, D.J., and Sabel, C.E., 2008. Neighbourhood effects on health: does it matter where you draw the boundaries? *Social Science & Medicine*, 66, 1241–1555.

Fotheringham, A.S. and Wong, D.W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23 (7), 1025–1044.

Frank, L.D., *et al*., 2005. Linking objectively measured physical activity with objectively measured urban form – findings from SMARTRAQ. *American Journal of Preventive Medicine*, 28, 117–125.

Galster, G., 2001. On the nature of neighbourhood. *Urban Studies*, 38 (12), 2111–2124.

Gatrell, A.C., 2011. *Mobilities and health*. Aldershot: Ashgate.

Gulliver, J. and Briggs, D.J., 2005. Time–space modeling of journey-time exposure to traffic-related air pollution using GIS. *Environmental research*, 97, 10–25.

Guo, D., 2008. Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science*, 22, 801–823.

Guo, D. and Wang, H., 2011. Automatic region building for spatial analysis. *Transactions in GIS*, 15 (s1), 29–45.

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

Hipp, J., 2007. Block, tract, and levels of aggregation: neighborhood structure and crime and disorder as a case in point. *American Sociological Review*, 72, 659–680.

Hipp, J. and Boessen, A., 2012. *Living in your own private Idaho: egohoods as a new measure of neighborhood* [online]. Available from: http://www.irle.berkeley.edu/culture/conf2012/hipp_boessen12.pdf [Accessed 21 September 2012].

Holt, D., *et al*., 1996. Aggregation and ecological effects in geographically based data. *Geographical Analysis*, 28, 244–261.

Inagmi, S., Cohen, D.A., and Finch, B.K., 2007. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social Science & Medicine*, 65, 1779–1791.

Johnston, R.J., *et al*., 2005. Spatial scale and the neighborhood effect: multinomial models of voting at two recent British general elections. *British Journal of Political Science*, 35 (4), 487–514.

Johnston, R.J., *et al*., 2007. Region, local context, and voting at the 1997 general election in England. *American Journal of Political Science*, 51 (3), 640–654.

Kawachi, I. and Berkman, L.F., eds., 2003. *Neighborhoods and health*. Oxford: Oxford University Press.

Knigge, L. and Cope, M., 2006. Grounded visualization: integrating the analysis of qualitative data through grounded theory and visualization. *Environment and Planning A*, 38 (11), 2021–2037.

Kwan, M.-P., 1999. Gender and individual access to urban opportunities: a study using space–time measures. *The Professional Geographer*, 51 (2), 210–227.

Kwan, M.-P., 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, 185–203.

Kwan, M.-P., 2004. GIS methods in time-geographic research: geocomputation and geovisualization of human activity patterns. *Geografiska Annaler B*, 86 (4), 267–280.

Kwan, M.-P., 2009. From place-based to people-based exposure measures. *Social Science & Medicine*, 69 (9), 1311–1313.

Kwan, M.-P., 2012. The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102 (5), 958–968.

Kwan, M.-P. and Ding, G., 2008. Geo-narrative: extending geographic information systems for narrative analysis in qualitative and mixed-method research. *The Professional Geographer*, 60 (4), 443–465.

Kwan, M.-P., *et al*., 2008. Reconceptualizing sociogeographic context for the study of drug use, abuse, and addiction. *In*: Y. Thomas, D. Richardson, and I. Cheung, eds. *Geography and drug addiction*. New York: Springer, 437–446.

Kwan, M.-P., *et al*., 2009. Activity-space measures for studying spatial crime and social isolation. Paper presented at the *Annual meeting of the Association of American Geographers (AAG)*, 22–27 March, Las Vegas, NV.

Kwan, M.-P., *et al*., 2011. Sociogeographic context, protobacco advertising, and smokeless tobacco usage in the Appalachian Region of Ohio (USA). Paper presented at the *2011 International Medical Geography Symposium*, 10–15 July, Durham, UK.

Lam, N.S., 2012. Geospatial methods for reducing uncertainties in environmental health risk assessment: challenges and opportunities. *Annals of the Association of American Geographers*, 102 (5), 942–950.

Lee, B.A., *et al*., 2008. Beyond the census tract: patterns and determinants of racial segregation at multiple geographic scales. *American Sociological Review*, 73, 766–791.

MacAllister, I., *et al*., 2001. Class dealignment and the neighborhood effect: Miller revisited. *British Journal of Political Science*, 31 (1), 41–49.

Macintyre, S., 1997. What are spatial effects and how can we measure them? *In*: A. Dale, ed. *Exploring national survey data: the role of locality and spatial effects*. University of Manchester, 1–17.

Macintyre, S., Ellaway, A., and Cummins, S., 2002. Place effects on health: how can we conceptualise, operationalise and measure them? *Social Science & Medicine*, 55, 125–139.

Manley, D., Flowerdew, R., and Steel, D., 2006. Scales, levels and processes: studying spatial patterns of British census variables. *Computers, Environments and Urban Systems*, 30 (2), 143–160.

Matthews, S.A., 2008. The salience of neighborhood: some lessons from sociology. *American Journal of Preventive Medicine*, 34 (3), 257–259.

Matthews, S.A., 2011. Spatial polygamy and the heterogeneity of place: studying people and place via egocentric methods. *In*: L. Burton, *et al*., eds. *Communities, neighborhoods, and health: expanding the boundaries of place*. New York: Springer, 35–55.

Matthews, S.A., Detwiler, A.J., and Burton, L., 2005. Geoethnography: coupling geographic information analysis techniques and ethnographic methods in urban research. *Cartographica*, 40 (4), 75–90.

Mobley, L.R., Kuo, T.M., and Andrews, L., 2008. How sensitive are multilevel regression findings to defined area of context? A case study of mammography use in California. *Medical Care Research and Review*, 65 (3), 315–337.

Mu, L. and Wang, F., 2008. A scale–space clustering method: mitigating the effect of scale in the analysis of zone-based data. *Annals of the Association of American Geographers*, 98, 85–101.

Oliver, L.N. and Hayes, M.V., 2007. Does choice of spatial units matter for estimating small-area disparities in health and place effects n the Vancouver Census Metropolitan Area. *Canadian Journal of Public Health*, 98 (Suppl. 1), 27–34.

Oliver, L.N., Schuurman, N., and Hall, A.W., 2007. Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. *International Journal of Health Geographics*, 6, 41.

Openshaw, S., 1984. *The modifiable areal unit problem*. Norwich: Geo Books.

Openshaw, S., 1996. Developing GIS-relevant zone-based spatial analysis methods. *In*: P. Longley and M. Batty, eds. *Spatial analysis: modelling in a GIS environment*. New York: Wiley, 55–73.

Pickett, K.E. and Pearl, M., 2001. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*, 55, 111–122.

Rainham, D., *et al*., 2008. Development of a wearable global positioning system for place and health research. *International Journal of Health Geographics*, 7, 59.

Rainham, D., *et al*., 2010. Conceptualizing the healthscape: contributions of time geography, location technologies and spatial ecology to place and health research. *Social Science & Medicine*, 70, 668–676.

Riva, M., *et al*., 2009. Disentangling the relative influence of built and socioeconomic environments on walking: the contribution of areas homogenous along exposures of interest. *Social Science & Medicine*, 69 (9), 1296–1305.

Root, E.D., 2012. Moving neighborhoods and health research forward: using geographic methods to examine the role of spatial scale in neighborhood effects on health. *Annals of the Association of American Geographers*, 102 (5), 986–995.

Sampson, R., Raudenbush, S., and Earls, F., 1997. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*, 277, 918–924.

Sampson, R.J., Morenoff, J.D., and Gannon-Rowley, T., 2002. Assessing "neighborhood effects": social processes and new directions in research. *The Annual Review of Sociology*, 28, 443–478.

Setton, E., *et al*., 2010. Gender differences in chronic exposure to traffic-related air pollution – a simulation study of working females and males. *The Professional Geographer*, 62 (1), 66–83.

Sherman, J.E., *et al*., 2005. A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics*, 4, 24.

Shi, W., 2010. *Principles of modeling uncertainties in spatial data and spatial analyses*. Boca Raton, FL: CRC Press.

Shoval, N., *et al*., 2011. Use of the global positioning system to measure the out-of-home mobility of older adults with differing cognitive functioning. *Ageing & Society*, 31, 849–869.

Sorensen, G., *et al*., 2004. Reducing social disparities in tobacco use: a socialcontextual model for reducing tobacco use among blue-collar workers. *American Journal of Public Health*, 94, 230–239.

Troped, P.J., *et al*., 2010. The built environment and location-based physical activity. *American Journal of Preventive Medicine*, 38 (4), 429–438.

Vallée, J., *et al*., 2010. The combined effects of activity space and neighborhood of residence on participation in preventive health-care activities: the case of cervical screening in the Paris metropolitan area (France). *Health & Place*, 16, 838–852.

van Ham, M., *et al*., eds., 2012. *Neighbourhood effects research: new perspectives*. Dordrecht: Springer.

Wang, F., Guo, D., and McLafferty, S., 2012. Constructing geographic areas for cancer data analysis: a case study on late-stage breast cancer risk in Illinois. *Applied Geography*, 35 (1–2), 1–11.

Wang, S., 2010. A cyberGIS framework for the synthesis of cyberinfrastructure, GIS, and spatial analysis. *Annals of the Association of American Geographers*, 100 (3), 535–557.

Weiss, L., *et al*., 2007. Defining neighborhood boundaries for urban health research. *American Journal of Preventive Medicine*, 32 (6S), S154–S159.

Wiehe, S.E., *et al*., 2008. Adolescent travel patterns: pilot data indicating distance from home varies by time of day and day of week. *Journal of Adolescent Health*, 42, 418–420.

Wilks, D., *et al*., 2010. Objectively measured physical activity and obesity prevention in children, adolescents and adults: a systematic review of prospective studies. *Obesity Reviews*, 12 (5), 119–129.

Wong, D.W.S. and Lee, J., 2005. *Statistical analysis of geographic information with ArcView GIS and ArcGIS*. New York: Wiley.

Yuill, R.S., 1971. The standard deviational ellipse: an updated tool for spatial description. *Geografiska Annaler B*, 53, 28–39.

Zenk, S.N., *et al*., 2011. Activity space environment and dietary and physical activity behaviors: a pilot study. *Health & Place*, 17 (5), 1150–1161.