Survey Research Methods (2017)
Vol. 11, No. 3, pp. 329-344
doi:10.18148/srm/2017.v11i3.6741

ISSN 1864-3361
http://www.surveymethods.org

Comparing Paper and Tablet Modes of Retrospective Activity Space Data Collection

Scott T. Yabiku
The Pennsylvania State University

Jennifer E. Glick
The Pennsylvania State University

Elizabeth A. Wentz
Arizona State University

Dirgha J. Ghimire
University of Michigan

Qunshan Zhao
Arizona State University

Individual actions are both constrained and facilitated by the social context in which individuals are embedded. But research to test specific hypotheses about the role of space on human behaviors and well-being is limited by the difficulty of collecting accurate and personally relevant social context data. We report on a project in Chitwan, Nepal, that directly addresses challenges to collect accurate activity space data. We test if a computer assisted interviewing (CAI) tablet-based approach to collecting activity space data was more accurate than a paper map-based approach; we also examine which subgroups of respondents provided more accurate data with the tablet mode compared to paper. Results show that the tablet approach yielded more accurate data when comparing respondent-indicated locations to the known locations as verified by on-the-ground staff. In addition, the accuracy of the data provided by older and less healthy respondents benefited more from the tablet mode.

Keywords: tablet; data collection; human-computer interaction; spatial; activity spaces

1 Introduction

As researchers gain accessibility to increasingly dynamic, accurate, and comprehensive spatial data from respondents, researchers face new data collection challenges. In this paper, we report on a project in Chitwan, Nepal, that directly addresses the opportunities in collecting accurate activity space data and the challenges that accompany these new modes. We developed a computer assisted interviewing (CAI) tablet-based approach to collecting activity space data. We subsequently evaluated the accuracy of this mode through a direct comparison with a paper map-based method for collecting the same activity space information.

Accurate spatial data provide important information on the role of geographic space on human behavior and well-being (Entwisle, 2007; Kuai & Zhao, 2017; R. Sampson, Morenoff, & Gannon-Rowley, 2002). Until recently, research has primarily focused on the neighborhood where individuals live as the relevant social context. Now, researchers are increasingly looking beyond the residential neighborhood to other spaces where individuals spend time. Studies from developed settings find that individuals spend large portions of their days outside their residential neighborhood (Browning & Soller, 2014). This may explain why some studies have found weak or no influence of neighborhood context on individual health and well-being, even when these associations have been strongly predicted by theory (Crowder & South, 2011; Inagami, Cohen, & Finch, 2007; Sastry & Pebley, 2010; Wodtke, Harding, & Elwert, 2011).

Activity space methods offer conceptual and empirical advantages over prior research on neighborhood context. First, measuring the social context at individuals' various activity spaces, which are not bounded by the residential neighborhood, may more realistically capture the exposures individuals receive (Kwan et al., 2008).For example, Lipperman-Kreda, Morrison, Grube, and Gaidus (2015) compared youths' exposures to tobacco outlets measured through activity spaces assessed via GPS tracking with exposures measured through typical neighborhood methods (800 meter buffers around home and school). Compared to activity space measures, the neighborhood measures greatly underestimated the potential exposure to opportunities for purchasing tobacco (Lipperman-Kreda et al., 2015). A second advantage of activity space methods is that they allow researchers to more finely conceptualize the idea of place.

Contact infromation: Scott T. Yabiku, 306 Oswald Tower, Department of Sociology and Criminology, The Pennsylvania State University, University Park, PA, 16802 (E-Mail: sty105@psu.edu)

Using LAFANS data, Sharp, Denney, and Kimbro (2015) showed that disadvantage in activity spaces and disadvantage in neighborhood context were independently associated with self-rated health. This prior activity space research on neighborhood contexts suggests that there may be different mechanisms through which place affects behaviors and health. The way neighborhoods affect individuals can be very different from the ways other places outside neighborhoods exert their influences. In sum, activity space approaches may offer better empirical measurement of exposure and may further the development of theory on the relationships between place, behaviors, and outcomes.

The collection of neighborhood context data through survey methods has a long history in the social sciences, and a large literature describes different approaches for measuring neighborhood context. For example, the neighborhood history calendar method is a retrospective approach for collecting data on how a neighborhood changes in relation to accessibility to important services and organizations, such as schools, health clinics, and transportation infrastructure (Axinn, Barber, & Ghimire, 1997). Systematic social observation is a technique for rigorously coding the presence of neighborhood features, such as physical or social disorder (R. J. Sampson & Raudenbush, 1999). This approach has now been adapted to use Google Street View data: comparisons between primary data field audits and Google Street View data confirm that this secondary digital imagery provides comparable measures of neighborhood context and interviewer observations at lower cost (Rundle, Bader, Richards, Neckerman, & Teitler, 2011).

Although substantial methodological work has examined the collection of neighborhood context data, we lack studies that compare different approaches for collecting activity space data. The studies on neighborhood context offer some guidance, but the data collection methods are different. Measuring activity spaces requires detailed information on where each individual participates in various domains of activities, such as consumption, education, production, and recreation. Collecting highly accurate activity space data from respondents can be difficult. For example, the LAFANS project in Los Angeles is a leading study for examining links between place and health, socioeconomics, and family well-being. LAFANS asked respondents for precise location data (numeric addresses or cross-streets) on the physical location of places respondents frequented, such as grocery stores, churches, and prior residences. Even though activity space data were requested of places just within Southern California, only about three-fourths of activity locations were successfully geocoded (Pebley & Sastry, 2004, p. 39). Collection of activity space data is even more challenging in areas that do not use numeric address systems, or do not use consistently named streets and roads as is often the case in economically developing settings.

2 Modes of spatial data collection: paper and electronic

The recent increase in research projects using activity space data aligns with a long history in the geographical sciences of asking respondents themselves to describe the spatial features of their environments (Hägerstraand, 1970). More recently, paper and electronic data collection methods have been used extensively in research to understand where and how people live and interact in space.

Paper-based approaches to spatial data collection have several strengths. One benefit of paper is that it presents a blank slate that can capture the respondents' spatial representations, unprompted by researchers. For questions that seek to understand the mental maps in people's thinking, to understand how people conceptualize places, and to learn how people assign priority to different places, asking respondents to draw their own maps on paper is an effective strategy (Boschmann & Cubbon, 2014; Poole, 1995). It is also comparatively easy to have respondents annotate paper maps (Mather, de Boer, Gurung, & Roche, 1998). Another benefit of paper collection of spatial data is that it is simple, has few upfront costs (no equipment or programming), requires minimal training for interviewers or respondents, and needs little maintenance or supplies (e.g., batteries) (Van Wart, Tsai, & Parikh, 2010). Finally, paper is an established method for collecting spatial data and has been used in many studies. Paper is inherently collaborative (Van Wart et al., 2010) and familiar. Some survey methodology research suggests that, compared to paper, electronic modes are more likely to raise issues of mistrust: in electronic modes, respondents may feel more uncomfortable and less likely to volunteer information (Wright, Aquilino, & Supple, 1998), although this is not consistent, and other studies have found no differences in reporting by mode (Bates & Cox, 2008).

Paper modes of data collection come with weaknesses as well. Cost and effort for processing paper data are higher. It is common for paper maps, either annotated or drawn by respondents, to be digitized for subsequent spatial analysis (Ramirez-Gomez et al., 2015). Data entry costs are typically higher for paper than electronic modes (Fricker & Schonlau, 2002). The cost differential is even higher for spatial data collected through paper because the data entry requires scanning, calibration, and high physical accuracy. Another weakness is the limited and static quantity of information that paper maps can display (Reilly, Rodgers, Argue, Nunes, & Inkpen, 2006). If a large geographic area is displayed on a paper map, high resolution detail must be sacrificed as a tradeoff, and the spatial accuracy of the map is decreased. Spatial accuracy can be preserved only with increasing physical map sizes, which can become unwieldy.

Electronic modes of spatial data collection address these main weaknesses of paper modes. With electronic methods, data processing costs are minimal because spatial information is collected from the respondents on the same geographic projection and datum presented in the software; the data require fewer processing steps to be ready for further analysis. In contrast to the fixed resolution of paper, the dynamic nature of electronic maps allows researchers to eliminate the tradeoffs between visual complexity, geographic scale, and spatial accuracy (Li & Ho, 2004; Reilly et al., 2006). Because it is zoomable, a tablet can show both a regional overview as well as a detailed, fine resolution focus on a specific area.

Electronic approaches, similarly, have strengths and weaknesses. The positive is that results are captured digitally, minimizing post-collection intervention. Furthermore, respondents and interviewers are increasingly familiar with digital devices such as phones and tablets. Nevertheless, electronic modes require more upfront costs for hardware and programming. For some research projects with relatively few respondents, these costs may greatly outweigh any benefits: the greater efficiency of electronic modes becomes appreciable at larger sample sizes. Another disadvantage is that some respondents may still view electronic devices as novel and may have trouble using them. This concern is likely to be population-specific. In industrialized settings, many respondents are very familiar with tablets and other handheld electronic devices, and researchers frequently instruct respondents to interact with devices directly. For example, Gourmelon, Le Guyader, and Fontenelle (2014) used tablets that let respondents draw their activities and the locations on a tablet. In research by Schoepfer and Rogers (2014), respondents used a zoomable tablet interface to indicate neighborhood features, such as boundaries, crime areas, and parks. In other settings, however, electronic modes may impede data collection if respondents or interviewers have less familiarity with these devices (van Heerden, Norris, Tollman, & Richter, 2014). Electronic hardware may also be more challenging to use in settings where electricity is unreliable and devices are exposed to the elements (precipitation, dust, or heat) and rougher transportation (bicycle, motorbike, or crowded public transportation) that can break electronics (Caviglia-Harris et al., 2012). Electronic devices may be more difficult to view in sunlight, which is a concern if interviewing will be conducted in the natural environment (Caviglia-Harris et al., 2012).

As the research community increasingly uses tablet and screen-based collection of spatial data, questions remain about whether electronic modes of data collection are accurate and if electronic modes significantly improve over previous paper-based methods. Specifically, are spatial data collected with tablets more accurate than spatial data collected with paper? Mode comparisons are common in the field of survey research methods, but these comparisons do not consider spatial accuracy. Typical metrics for mode comparisons include response rate (Kaplowitz, Hadlock, & Levine, 2004), cost (Kaplowitz et al., 2004), item non-response (Woo, Kim, & Couper, 2014), duration of interview (Gummer & Roßmann, 2015; Watson & Wilkins, 2015), and reporting on sensitive behaviors (Burkill et al., 2016). To our best knowledge, there are no experimental studies with random assignment that compare tablet and paper modes for the collection of spatial data in face-to-face interviews. To design effective data collections, it is important to know if tablets provide higher accuracy than paper. It is also important to assess if tablets provide higher accuracy to all or only to a subgroup of users.

Furthermore, there will be increasing need for more methodological work comparing spatial accuracy of different approaches, such as active versus passive data collection, personal computer versus mobile device, and individual versus crowd-sourced. Our work here suggests initial methods to compare spatial accuracy across different modes. In the future, these comparisons will become more important in the survey research literature as electronic spatial data collection proliferates.

3 The design of effective computer assisted interviewing methods for spatial data collection

In any CAI application, usability is a main concern (Couper, 2000, 2008), and web survey usability from the respondent perspective is well studied (Geisen & Bergstrom, 2017). In CAPI applications, however, usability from the respondent's perspective is often not considered. For some information, it may not be necessary that respondents see data as it is entered or see any visual aids from the interviewer. The recall of spatial data, however, is especially primed and eased by visual aids – most notably maps (Brewer, 1986, 1988; Williams, Healy, & Ellis, 1999). For finding locations in settings that do not use numeric address systems, it is important that respondents can view maps or satellite imagery and point to locations. A diversity of respondents has demonstrated facility with maps including young children who can effectively use maps to navigate and understand their environment (Sandberg & Huttenlocher, 2001; Wiegand, 2006). The quality of maps also matters: the more closely a map represents the real world, the better respondents can effectively use it to recall pertinent information (Downs, 1981; Guelke, 1979; Peterson, Kulhavy, Stock, & Pridemore, 1991; Schwartz & Kulhavy, 1981). Satellite imagery, with street overlays, provides all sorts of graphic detail such as buildings, roads, and bodies of water.

Usability is enhanced when the interface is carefully designed for the needs of the users, and map design is no exception to this principle (Konečný, Kubíček, Stachoň, & Šašinka, 2011). A Zoomable User Interface (ZUI) in a CAI approach prevents the user from being overwhelmed with data, yet it allows the user to seek out details when appropriate. In the field of Human Computer Interaction (HCI), ZUIs have been widely accepted as an effective way of allowing access to more information than can fit on one screen (Bederson, 2011). Broadly speaking, ZUIs show or hide information to present the user with the ideal Information-to-Interface Ratio (Harrower & Sheesley, 2005). When ZUIs are applied to satellite imagery, at a low zoom level, the user can see a larger geographic area, but few details; for example, only major roads might be visible. At a high zoom level, the user sees a smaller geographic area, but many more details: residential roads as well as individual homes, trees, and natural features. The ZUI is fluid and the visual flow allows enhanced comprehension of the spatial data (Bederson, 2011).

It is important to design a spatial data collection interface so respondents can collaboratively use the instrument with the interviewer. Collaborative, semi-structured data collection approaches can lead to higher data quality when respondents are tasked with activities of high cognitive difficulty (Schober & Conrad, 1997). The widely used life history calendar (Freedman, Thornton, Camburn, Alwin, & Young-DeMarco, 1988) method is an example of a semistructured approach that usually leads to higher data quality than if questions were asked in a purely structured manner. In the design employed in the current study, the aim is to extend this semistructured approach to the design of a human-computer interface based on a multi-touch display, which has been shown to be superior to mouse and keyboard for collaborative work (Cooperstock, Fels, Buxton, & Smith, 1997; Forlines, Wigdor, Shen, & Balakrishnan, 2007). Although our study did not collect information from multiple respondents simultaneously, the interface was collaborative because the interviewer needed to observe the respondent using the map and intervene to show how to use the interface if the respondent had difficulty. Scrollbars, boxes and buttons, and small-sized controls on the screen require precision movements with a mouse or keyboard, which interrupt collaboration because they are single user input methods. One person has to wait until the other finishes; furthermore, the second user is not always aware what the first user has clicked or typed, which impedes information flow and cognition between users (Schneiderman & Plaisant, 1998; Scott, Shoemaker, & Inkpen, 2000). Experiments show that navigating a map interface by touch is more intuitive, faster, and more preferred than keyboard or mouse (MacKay, Dearman, Inkpen, & Watters, 2005; Sears & Shneiderman, 1991), and thus may improve spatial recall. In our CAI approach, respondents can drag a finger on the screen to scroll the map. Natural gestures in human computer interfaces reduce errors (Burigat, Chittaro, & Gabrielli, 2008) and increase the engagement of the user even on a small screen (Brandl, Forlines, Wigdor, Haller, & Shen, 2008; Hinckley et al., 2010; Yee, 2004). Effective use of touch technology allows a computer display to maintain the engagement of a respondent much like a large sheet of paper in a traditional collaborative approach.

4 Hypotheses

We test the change in positional accuracy when respondents participate in the collection of spatial activity space data via computer assisted interfaces versus a paper map. This design was guided by several hypotheses regarding data quality issues between tablet and paper mode. Our data collection plan randomly assigned CAI tablet or paper map mode to respondents, and then later independently confirmed the locations of the spaces respondents indicated through a "ground truth" process. Ground truth is a term used to describe the process in which a location provided by one source (e.g., a respondent's answer) is verified by a second method with higher accuracy (e.g., field staff with a GPS) that involves physically visiting the location "on the ground" (Foody, 2002).

Accuracy of data. First, we hypothesized that the CAI tablet approach would yield more accurate data. Specifically, the difference between the ground truth location and the respondent-provided location will be smaller for tablet than paper mode. The tablet interface is zoomable, pannable, and dynamic, which allows respondents to better orient themselves to the satellite imagery. The paper mode, in contrast, is static satellite imagery in hard copy. Furthermore, due to the enhanced spatial comprehension offered to respondents when using a zoomable tablet interface, we hypothesize that the tablet mode will come closer to its lowest possible error, compared to the paper mode. In other words, we predict that, with regards to accuracy, tablet data collection will outperform paper, even after acknowledging the higher spatial resolution of the tablet device.

Respondent differences in performance. Second, we hypothesized that the CAI approach would not be equally accurate for all respondents in our diverse sample. Various background factors affect how well respondents perform in surveys. In general, respondent characteristics that capture cognitive abilities have been found to be correlated with response accuracy. Age and education frequently have been found to affect response time and consistency (Eggs & Jaeckle, 2015; Sauer, Auspurg, Hinz, & Liebig, 2011); lower education individuals may be less likely to understand the tasks being asked of them (Krosnick, 1992). Health, as another indicator of cognitive ability, is also likely to be a predictor of response errors. Although gender is not an indicator of cognitive ability, Nepal is a very gender segregated setting and women are often not given the same educational experiences as men; furthermore, women's experiences are also more limited with regards to activities outside the home. In the Nepali context, therefore, female gender is also predicted to be associated with higher survey errors. Unfortunately, we did not have measures known to predict the ease of map use, such as cognitive measures and visual-spatial ability (Allen, 1999; Sholl & Egeth, 1982).

The purpose of our hypotheses, however, is not to test the associations between respondent characteristics and survey response errors. Instead, our hypotheses test which types of respondents benefit more from the tablet mode versus the paper mode, with regards to spatial accuracy. We expect that individuals with more education are expected to be more comfortable using tablet technology and better able to make use of the technology and have more accurate markings. In other words, we hypothesize an interaction between tablet mode and education, in which the errors are decreased significantly more for those with more education when using the tablet. We expect similar interactions for individuals who are younger, male, and healthier.

5 Setting, Data, and Methods

5.1 Setting

Since 1996, the Chitwan Valley Family Study (CVFS) has extensively measured social change and family behaviors in the Chitwan Valley of Nepal. The Chitwan Valley is about 100 kilometers west of Kathmandu, the capital city of Nepal. Chitwan is located in the Terai, a region of low-lying plains along the southern borders of the country. This is an ideal location for our proposed project. First, the study site is typical of many developing areas and does not use any numeric address system, which makes it an excellent location for testing our instrument. Second, there is significant variability across Chitwan, ranging from urban to very rural areas, which allows us to test the instrument in areas of dense and sparse contextual features and activity spaces.

5.2 Data Collection

In the late summer and fall of 2015, our data collection partner – the Institute for Social and Environmental Research-Nepal (ISER-N) – fielded a household, face-to-face survey to collect activity space data. The sampling frame for this survey came from a registry of households in the CVFS that was established in 1996 and continues to the present day. From the most recent registry in February 2015, a sampling frame was constructed of all individuals between the ages of 15 and 49 currently living in the household; individuals who were listed as living away from the household due to migration were excluded. Approximately 1560 individuals between the ages of 15 and 49 were randomly sampled. Of these, interviews were completed with 1433 individuals, a response rate of 92%. No survey incentives were used. The high response rate is due to several reasons: the cultural expectation in this setting is that visitors, such as survey interviewers, are welcomed and receive attention, ISER-N is a well-established scientific research organization that has excellent relationships with the community, and the respondents were existing household registry participants who were familiar with survey routines.

The age range of 15-49 was chosen because individuals within this range are more likely to be engaged in a diverse set of activities. In addition, in Nepal, the declines in health are steeper than in industrialized settings, and even at relatively early ages, physical mobility can be impacted. Furthermore, vision care (e.g., glasses) is not consistently available for all people. In our pilot work, we found that limitations in eyesight prevented some older individuals from using maps.In sum, we wanted to maximize respondents with activities outside the home and who were able to successfully use the paper or tablet spatial data collection modes.

The survey consisted of both a standard, interviewer-led structured questionnaire and a spatial activity space component. The standard survey component included questions about basic demographic characteristics, employment, and overall self-rated health. The activity space component asked respondents about the following activities of everyday life: going to school, shopping, visiting a health provider, restaurants, place of employment, worship, recreation, visiting friends/relatives, banking, clubs/groups, and visiting local government offices. This list of activities was chosen based on our assessment and our local research partners' cultural knowledge and experience of the typical activities in which individuals in this setting regularly engage.

For each activity, respondents were asked if they did the activity in the past week, how many times in the past week they did that activity, and where the activity was located. If a respondent did not engage in a particular activity in the past week (for example a respondent may have a usual doctor, but did not visit the doctor recently), the respondent estimated when he/or she last did the activity, and indicated the location of the activity.

The procedures for generating the list of activities from respondents and collecting information on frequency were identical for tablet and paper mode. The mode of data collection for the activity space measures was randomly assigned to respondents: CAI tablet or paper map. In our sample of 1,433 respondents, approximately 2/3 of respondents recorded the activity locations using a CAI tablet approach, and 1/3 with paper map. Rather than an equal 50-50 split, this distribution was chosen to balance two project goals. The first goal of the project was to compare accuracy of CAI versus paper modes, and this is the work we report in this paper. The second goal was to answer substantive hypotheses about activity spaces and respondent social characteristics, such as gender and health, and these analyses are not reported in this paper. We suspected that the tablet would have better spatial accuracy, so we assigned 2/3 of the sample to the tablet mode in order to collect data with the expected higher spatial accuracy, which would increase statistical power for our substantive hypotheses. Yet to test accuracy between tablet and paper, we still needed a sizeable number of respondents answering with the paper maps for the mode comparisons to have sufficient statistical power, and thus we assigned 1/3 of the sample to paper map mode. Across the CAI and paper mode, approximately 15,500 unique activities were marked by the 1,433 respondents.

We considered the spatial data collection to be a semi-structured interview process because the interaction between respondent and interviewer was variable depending on the respondent's spatial ability. Some respondents used the tablet without any help from the interviewer, some needed continual guidance, while others needed less assistance as they gained familiarity with the task as the interview progressed. Throughout the interview, respondents were encouraged to familiarize themselves with the tablet or paper map, and exploration varied across respondents.

5.3 Computer assisted interviewing (CAI) tablet

The tablet mode was administered by trained interviewers using tablets with Windows 8 operating system (ASUS Transformer Book T-100). The Windows 8 tablets were low-cost (approximately $350 each), had 10.1-inch touch capacitive screens, and were equipped with solid state disk drives, instead of mechanical hard drives, for greater durability in the field. Before use, tablets were loaded with appropriate software and data. The software was custom designed and programmed by our team in C#. The data were derived from satellite images of the area stored on the machine locally for two reasons: 1) it eliminated the need for costly wireless data connections to a remote database, and 2) the loading of imagery was instantaneous from the solid state disk. Satellite images are large files, and processing these images needs substantial computational performance beyond the capacity of field laptops and tablets. In our approach, the processing is done asynchronously: in advance of the field period, raster satellite imagery at multiple zoom levels is subdivided into image tiles, compressed, indexed on a high performance workstation, and stored on the tablets. During the interview, the tiles are "served" locally. This tile-based approach follows the standard architecture in modern web cartography (Barclay, Gray, & Slutz, 2000), i.e., the same as used by Google Maps, Bing Maps, among many others.

A design principle of our CAI interface was to dedicate the majority of the display to visualization in order to increase respondent engagement and promote a collaborative interview environment. The interface was fluid, pannable, and zoomed in and out easily, much like any modern tablet app. Although the interviewer was always present, the CAI interface was designed to be approachable and easy to use, even for a non-expert. Figure 1 shows the CAI interface.

5.4 Paper map

In pilot testing our instruments, we tried multiple approaches to printing paper satellite imagery maps that could encompass the entire Chitwan study area. Smaller paper maps of varying sizes, while easier to transport and handle, did not offer enough detail to observe individual roads and landmarks or a large enough geographic area to capture the range of activities. After piloting different solutions, the most effective approach was a large map approximately 1.2 meters wide by 2.4 meters long, large enough to allow for details and a large area simultaneously. Maps were printed on heavy-duty paper so that they would be durable for repeated interviews. The cost for printing each map was approximately $100 USD each. The large paper size allowed respondents to see the general area of Chitwan and identify major intersections of roads, but not enough to identify individual buildings. The paper map showed an area of about 200 sq. km. and then broke it into squares, each of which was 1600 meters across. To identify locations, respondents pointed to the area on the map and interviewers noted the grid cell and recorded that grid coordinate (e.g., "A17" or "E12"). The centroids of these grid cells were later converted into geographically referenced locations. In both the tablet and paper modes we added landmarks to help respondents orient themselves to the map and increase their ability to recall spatial information. For example, we labeled well-known intersections or "chowks." Chowks are a common point of reference and navigation in this setting. In the tablet mode, the chowk labels were in bright yellow Nepali script and visible at all levels of map zoom because their text size remained constant, even as the respondent zoomed in or out of an area. In the paper map, the chowks were labeled in blue Nepali script because that color provided the highest contrast. Figure 2 shows the paper map (in reduced size).

5.5 Analytical Methods

To assess the accuracy of the location data collected with the tablet versus paper modes, we required a benchmark or true reference point of the actual locations respondents intended to mark. When we collected the locations, in both tablet and paper mode, we also asked for the complete name and address information for a subset of activities (in Chitwan, the "address" is typically a town name, intersection of major roads, or an administrative division known as a ward). In Fall 2015, we "ground truthed" the activity space locations for a subset of respondents (400 randomly selected tablet respondents and 400 randomly selected paper respondents). We limited our ground truth activity to a total of 800 respondents, rather than the full sample of 1433, due to budget constraints. Based only on the name and limited address information provided by the respondent (not the geographic data), interviewers visited the activity mentioned in the data (e.g., a school or restaurant or bank) and used GPS to mark the location of the activity. This location marked by the interview staff served as the "gold standard" reference point.Distance in meters between this reference point and the respondent's location (marked via tablet or paper) serves as the depen-

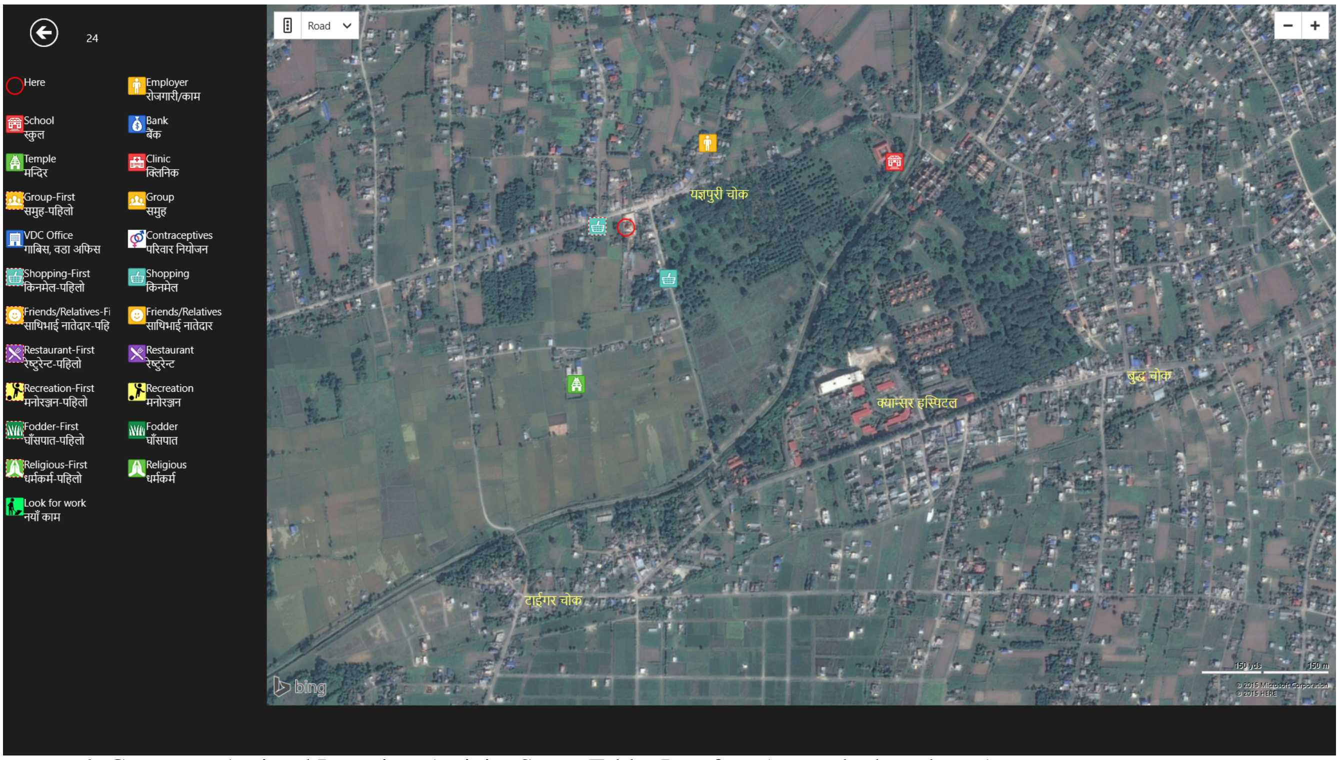

Figure 1. Computer Assisted Interview Activity Space Tablet Interface (example data shown)

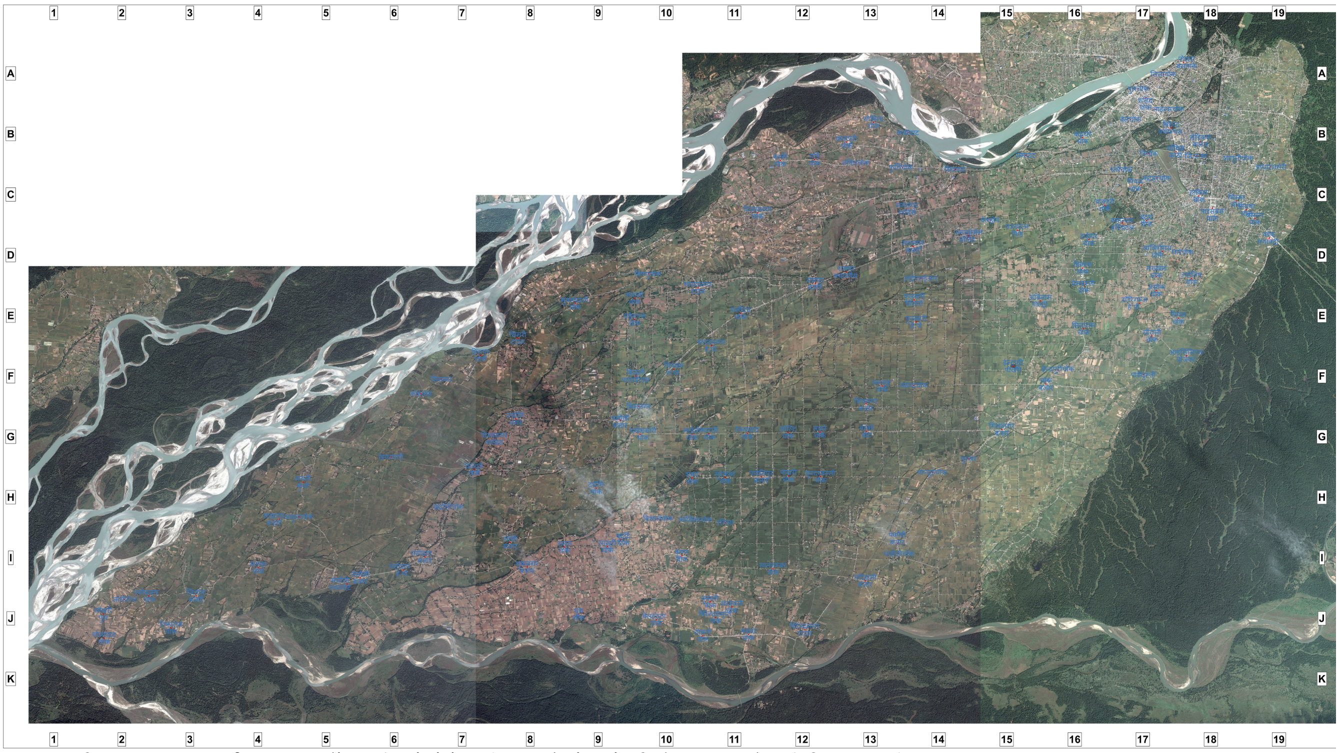

Figure 2. Paper Map for recording Activities (actual size is 2.4 meters by 1.2 meters)

Table 1
Ground truth verification status by collection mode for 3,527 activities across 800 respondents

Ground truth verification status	Paper %	Tablet %	Total %
Activity location verified and found	87	89	88
Unclear if activity found	11	10	11
Activity location not found	2	1	1
N	1783	1744	3527

χ^2-test of independence = 3.27, df = 2, p-value = 0.19

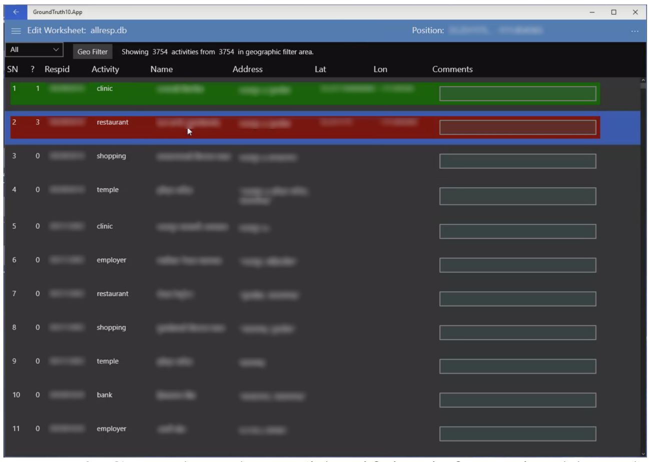

Figure 3. Ground Truth app (identifying information blurred)

dent variable for assessing accuracy of paper or tablet mode. Note that the inherent variability in GPS, which was used in ground truthing, depends mostly on a clear path to the sky, and is typically within 10 meters; this variability is equally distributed across the ground truthing of paper and tablet mode.

Due to time and budget constraints, we did not ground truth all activities for the 800 randomly selected respondents. Instead we chose activities that represent common yet diverse activities in this setting: banks, schools, health clinics, employers, restaurants, shops, and temples. These types were chosen in consultation with our partners in Nepal and yielded a set of 3,527 activities that were ground truthed for the 800 respondents. The number of activities to be verified via ground truth differed slightly by mode: 1,783 activities from the 400 paper respondents, and 1,744 from the 400 tablet respondents. This is expected because respondents provided a variable number of activities.Some respondents, for example, did not have a bank, so they did not provide the location of this activity.

Our research team programmed a separate tablet app in C# to facilitate the ground truth process.The ground truth app contained an internal database with all 3,527 activities, listing their name and location information. The tablets were connected to a GPS device, which automatically recorded the location when the activity was verified. After visiting the location of the activity as provided by the respondent, field staff recorded each activity as found, not found, or unclear. Rows for activities in the app would turn to green, red, or yellow to help staff quickly visualize the disposition of activities. The app was designed for touch interaction to allow efficient marking of rows and scrolling through activities. Figure 3 shows the ground truth app.

A verification result of "unclear" could happen if the name or address of the activity provided by the respondent was not detailed enough to conclusively determine if the staff had located the activity. For example, a respondent might have indicated he or she shopped at a "vegetable stand" in a given neighborhood intersection. If, during ground truthing, interviewers visited the location and found two vegetable stands in the area that matched the respondent description, it would be unclear which one the respondent meant.

Overall, most activities were successfully ground truthed by research staff. As shown in Table 1, 3,106 activities (88%) were found and verified as such by our interview staff. For the analyses evaluating accuracy, we use only these 3,106 verified activities. Note that there was no association between mode of activity collection (tablet/paper) and verification status (p=.19). This suggested there would be no sample selection bias by mode when we used the found activities to evaluate accuracy across tablet and paper mode.

To determine the relative accuracy and characteristics of respondents who report their activity spaces, we include several demographic background factors including age, self-rated health (1 to 5 scale), gender (female coded as 1, male 0), and years of education. We also include a measure of how much difficulty the respondent encountered using the tablet or paper map: when the interview finished, the interviewers made an assessment of how well the respondent understood the protocol for indicating locations: very well/excellent, okay/average, or with difficulty. We coded this difficulty scale from 1-3, with 3 representing high difficulty with the task.

5.6 Conceptualizing Spatial Accuracy across Modes

There were important differences in the spatial resolution of paper maps when compared to the tablets. Therefore, we must consider how we might conceptualize accuracy with each mode. Recall that the paper map divided the study area into squares, each of which was 1600 meters across. For a square of side s, the average expected distance $E(d)$ between a set of randomly chosen points in the cell to the center point can be calculated as (Rogerson, 2014, p. 157),

$$E(d) = \frac{s}{6} \cdot \left(\sqrt{2} + \ln\left(1 + \sqrt{2}\right)\right) \approx 0.383s \tag{1}$$

For our paper maps gridded with 1600 meter cells, the average distance is about 616 meters. This is the average error one would expect when using 1600 meter cells to collect exact point data. Thus the highest possible accuracy, on average, for the paper mode would be an average error of 616 meters. In contrast, the satellite imagery used for the tablet interface has sub-meter resolution, and thus the highest possible accuracy for tablet mode would be close to 1 meter.

When comparing the accuracy of the paper versus the tablet models, the most useful question is not simply whether the tablet mode has lower spatial error than paper mode; it most likely will. Instead, we compare each mode to its "best case" scenario (less than 1 meter error for tablet, 616 meters for paper map).

5.7 Models

Our outcome variable was the average difference, in meters, between all activity locations as marked by a respondent and the locations of all these activities as marked by the ground truth process. For the CAI mode, the difference for a single activity was the distance between the point indicated with the respondent's finger and the ground truth GPS coordinate. For the paper mode, the difference for a single activity was the distance between the centroid of the square indicated by the respondent and the ground truth GPS coordinate. For both CAI and paper, these differences were then averaged for all of the activities a respondent indicated. Because our outcome is continuous, we use regular ordinary least squares regression to predict the dependent variable. We estimate three models to assess spatial accuracy and test our hypotheses:

$$Y_i = \beta_0 + \beta_1 \text{Mode}_i + \epsilon_i \quad , \tag{2}$$

$$\begin{aligned} Y_i = \beta_0 &+ \beta_1 \text{Mode}_i + \beta_2 \text{Age}_i + \beta_3 \text{Education}_i + \\ &\beta_4 \text{Female}_i + \beta_5 \text{Health}_i + \beta_6 \text{Difficulty}_i + \epsilon_i \quad , \end{aligned} \tag{3}$$

and

$$\begin{aligned} Y_i = \beta_0 &+ \beta_1 \text{Mode}_i + \beta_2 \text{Age}_i + \beta_3 \text{Education}_i + \\ &\beta_4 \text{Female}_i + \beta_5 \text{Health}_i + \beta_6 \text{Difficulty}_i + \\ &\beta_7 \text{Mode}_i \cdot \text{Age}_i + \beta_8 \text{Mode}_i \cdot \text{Education}_i + \\ &\beta_9 \text{Mode}_i \cdot \text{Female} + \beta_{10} \text{Mode}_i \cdot \text{Health} + \\ &\beta_{11} \text{Mode}_i \cdot \text{Difficulty}_i + \epsilon_i \quad , \end{aligned} \tag{4}$$

where Y_i is the average absolute difference in meters between all activity locations and the corresponding ground truth locations for respondent i, Mode is a binary indicator of tablet (Mode = 1) or paper (Mode = 0) data collection, and the remaining main effect coefficients are demographic characteristics (age, education, female gender, self-rated health) and the interviewer assessment of respondent i's difficulty with the location protocol. The interaction terms are multiplicative interactions between each of the predictors and survey mode, and ϵ_i is the error term. The first model tests the main effects of mode, the second model tests the main effects of demographic characteristics, and the third model tests how usability varies across these demographic characteristics.

6 Results

Table 2 presents the descriptive statistics for the sample. Overall, the respondents averaged 31 years old (note that the maximum age was 50, although the sampling frame was age 15-49: by the time of the survey, a handful of 49-year-old respondents had had another birthday). About 60% of respondents were female. Many respondents had some education although years of formal schooling averaged slightly fewer than 8 years of education, which is to be expected for this age group in this setting. We also note that respondents rated themselves as in fairly good health, averaging 2.78 on the self-rated health assessment (on a 1-5 scale). We tested for differences on these background characteristics across mode. T-tests showed no significant differences by mode in average age, education, sex, or health, which suggests the randomization between tablet and paper mode was effective.

One of the first ways to compare the modes of data collection is to consider how the interviewer rated the respondent's ability to locate places on the map. On a scale of 1 to 3 with a 3 indicating the most difficulty, difficulty with the protocol averaged 1.34. There was little difference by mode in how the interviewer assessed the respondent's difficulty of locating places on the map: interviewers rated tablet respondents' difficulty with an average of 1.36 and paper respondents' difficulty with an average of 1.30. This difference was not significant ($p > 0.05$).

Despite the similarity in interviewer reported ease of use, there was a large difference in the accuracy of location of activities by mode. Recall that 3,106 activities from 400 tablet and 400 paper respondents were ground truthed successfully (activity locations were verified as found). For each respondent, an average error was calculated, which is the mean error in the location of their activities as reported by respondents compared to the locations as verified by interview staff. Across both modes, average error was 405 meters. For the 400 tablet respondents, the average error was 75 meters. For the 400 paper respondents, the average error was 735 meters. This difference in average error across modes was significant ($p < 0.001$). Because of the spatial resolution of the data source, the tablet was clearly more precise, but the key comparison here is not across the two modes, but each mode against its hypothetical minimum error. For tablet mode, the average error was 75 meters more than the hypothetical minimum ($75 - 0 = 75$). For paper mode, the average error was 119 meters more than the hypothetical minimum ($735 - 616 = 119$). In other words, the tablet mode outperformed the paper mode by about 44 meters of accuracy ($119 - 75 = 44$), even after accounting for the better spa-

Table 2
Descriptive Statistics for 800 respondents

	Mean	SD	Min	Max
Average Error between Respondent Locations and Ground Truth Locations (meters)	404.98	381.44	5	3737
Tablet mode (versus paper)	0.50	0.50	0	1
Age	31.47	10.57	16	50
Categories of age				
15-19 Years	0.18	0.38	0	1
20-29 Years	0.27	0.44	0	1
30-39 Years	0.26	0.44	0	1
40-50 Years	0.29	0.45	0	1
Years of education	7.56	4.32	0	16
Categories of education				
0-6 Years	0.34	0.47	0	1
7-9 Years	0.23	0.42	0	1
10-12 Years	0.39	0.49	0	1
13+ Years	0.04	0.20	0	1
Female (versus male)	0.61	0.49	0	1
Self-Rated Health	2.78	1.10	1	5
Categories of self-rated health				
Poor	0.03	0.17	0	1
Fair	0.52	0.50	0	1
Good	0.23	0.42	0	1
Very Good	0.09	0.29	0	1
Excellent	0.13	0.37	0	1
Difficulty with protocol	1.34	0.50	1	3

tial resolution of the tablet (less than 1 meter) compared to the gridded paper map with 1600 meter cells. This suggests strong support for our first hypothesis that the tablet mode increases spatial accuracy compared to a paper map mode.

Our second hypothesis focused on the characteristics of respondents that were most sensitive to mode of data collection and whether respondent accuracy was better for some types of respondents when using a tablet compared to the paper mode of activity space data collection. These hypotheses were tested with a series of interaction coefficients. Before estimating these interaction models, we present the main effects of predictors. Model 1 of Table 3 predicted average error with only mode (tablet or paper) as a predictor. This essentially replicated the mean differences reported in the prior paragraph: the intercept of 735.1 was the average error for paper, and the tablet coefficient (−660.2) indicates that tablet mode error was 660.2 meters less than paper, or 74.9 meters (735.1 – 660.2 = 74.9).

Model 2 predicted average error with demographic characteristics and the indicator of respondent difficulty with the location protocol. None of the predictors, such as age, education, gender, or health, were significantly associated with spatial accuracy. Although this may appear unexpected, most of the literature on demographic characteristics and response error comes from research on standard survey questions, not novel cognitive spatial tasks. The setting (rural Nepal) also differs from the location of most prior survey research literature, which may further help explain the lack of strong associations between demographic characteristics and respondent performance.

Model 3 tested our primary hypothesis that respondent accuracy may be different across respondents in the two different data collection modes.We hypothesized that younger respondents, more educated respondents, men, and healthier respondents would show less error in tablet mode when compared to similar respondents using paper. In this case, the coefficients of interest are the interactions, which test which types of respondents are most accurate (or have less error) on each mode of data collection.

The results of these interaction models were contrary to

Table 3
Predicting Error (meters) between Respondent Location and Ground Truth Location

	Model 1		Model 2		Model 3	
	Pred. err	SE	Pred. err	SE	Pred. err	SE
Tablet mode (reference is paper)	-660.158^{**}	13.501	-661.522^{**}	13.556	94.240	-693.837^{**}
Age	-	-	-0.142	0.738	1.443	1.067
Education	-	-	0.916	1.945	0.829	2.756
Female (reference is male)	-	-	10.283	14.454	17.027	20.800
Health	-	-	1.909	6.342	-11.299	9.015
Difficulty with location protocol	-	-	22.289	15.169	-3.760	22.896
Age × Tablet Mode	-	-	-	-	-2.943^{*}	1.473
Education × Tablet Mode	-	-	-	-	0.558	3.878
Female × Tablet Mode	-	-	-	-	-18.275	28.885
Health × Tablet Mode	-	-	-	-	25.924^{*}	12.657
Difficulty with protocol × Tablet Mode	-	-	-	-	45.104	30.483
Intercept	735.063^{**}	9.547	691.959^{**}	47.629	708.969^{**}	67.955
N	800		800		800	

* $p<0.05$ ** $p<0.01$

initial expectations. Model 3 demonstrated that older respondents' accuracy was significantly higher when the mode was tablet compared to paper map (significant negative interaction between age and tablet mode) when compared to younger respondents. The negative coefficient for tablet (which represented lower error for tablet versus paper) grew in magnitude as age increased.

Similar findings appeared for the self-rated health of respondents.The accuracy of respondents in poorer health was again significantly higher when they used tablet compared to paper. The significant positive interaction between health and tablet mode means that the effect of tablet on reducing error was not as strong for healthier respondents compared to poorer health respondents. Health was coded such that higher values represented better health; as health improved, the coefficient for tablet became less negative. There were no significant interactions between mode and education, gender, or the respondent's difficulty with the location protocol. Overall, the results demonstrated lower errors (higher accuracy) for the tablet data collection mode. Furthermore, the results suggested that the accuracy benefits of tablet technology were especially observed among older respondents and respondents reporting poorer health.

7 Discussion and Conclusions

In this paper, we examined a technological approach to measuring retrospective activity spaces using tablets in a collaborative interview and compared it to a traditional paper approach in the same interview format. Our goal was to assess how using an interactive tablet interface affected the accuracy of data when compared to a more traditional paper-based approach. Our questions have high relevance to current research: many projects are increasingly turning to electronic modes of collecting spatial data, yet it is not known how this shift from paper to electronic modes affects data accuracy on activity space studies.

Overall, we conclude that using this technology is quite feasible and yields higher accuracy than paper modes. The ground truthing exercise demonstrates that a CAI approach enables respondents to mark these locations with better accuracy than paper mode. The tablet interface is zoomable, pannable, and dynamic, which allows respondents to better orient themselves to the satellite imagery. The paper mode, in contrast, is static satellite imagery in hardcopy. Consistent with our hypothesis, the tablet approach yielded data that was significantly more accurate, even when accounting for the fact that the paper map, which used 1600 meter cell grids, had inherently lower spatial resolution than the zoomable satellite imagery.

Our second hypothesis was that accuracy would increase most significantly for younger male respondents and for those with more education and who report being in better health because these respondents were expected to be better positioned to take advantage of the increased potential in accuracy afforded by the tablet mode. A common assumption is that younger respondents are more adept at using new technologies and interfaces, and some studies find that the young perform better with novel interfaces (Bobeth et al., 2014). There are sound biophysical reasons why older individuals and individuals with poorer health would have more difficulty with technological input devices: worsening eyesight, less motor control, and decreased strength (Caprani, O'Connor, & Gurrin, 2012; Taveira & Choi, 2009). Touch interfaces sometimes use specific gestures for zoom, selection, and panning, and the precision required may challenge older individuals (Harada, Sato, Takagi, & Asakawa, 2013).

These results, however, are more consistent with expectations that increased technology enhances the ability of older and less healthy respondents to provide accurate spatial data. In this case, the technology's precision may help compensate for limitations of age or poor health that hamper respondents' ability to successfully navigate around the paper map. Rather, with the ability to zoom in on areas and orient themselves with a touch screen interface, these respondents could be even more precise when indicating their activities. This is somewhat contrary to expectations that those with the least experience with technology, such as older respondents in this developing setting, would have the most difficulty indicating their activities with the tablets.

The specifics of our data collection point to potential reasons why the CAI approach worked well for older adults: while the paper map mode was a very familiar medium, its large size (1.2 meters by 2.4 meters) was probably a limitation for older, more frail adults. During interviews, maps were typically unrolled on the floor, a porch, or outside courtyard. Although sitting on the floor or mat is common in this setting, it seems likely that younger respondents are less likely to have difficulty moving around on the floor to observe the paper map closely whereas this was likely to be a greater challenge for some of the older adults and those in poor health. In this case, a tablet Zoomable User Interface – even though novel and perhaps unfamiliar – more effectively served the needs of older adults tasked with marking spatial locations.

Our findings are encouraging because they suggest that new technologies for data collection – in our study, touch screen tablets – should not immediately be assumed to be a barrier to data collection for older adults or those in poorer health. On the contrary, the accuracy of these individuals' reports appeared to have improved the most when they used tablets, relative to paper map mode. Research suggests that well-designed touch interfaces can be highly intuitive, even for older adults (Culén & Bratteteig, 2013). Rather than a technological barrier to be overcome, a well-designed interface is an assistive technology that enhances respondents' performance. Our findings are consistent with other studies that have found that older adults can use tablet modes equally well as or even more effectively than paper modes in tasks such as standard survey questionnaires (Fanning & McAuley, 2014). Our research contributes to this literature with evidence that tablet modes can successfully be used to collect spatial data among older, less healthy individuals.

Our study is not without limitations. First, our list of activities was predetermined through a series of discussions with collaborators in our setting and through pre-testing the instrument. While we believe our activities are comprehensive and capture the breadth of daily life in this context, it is possible that we have missed important, emergent activities that might have been revealed from an open-ended approach. Second, our project collected only the point locations of where activities occurred. The travel paths respondents take to these activities may also be important components of exposure (van Heeswijck et al., 2015). It is unknown if our results, which suggest the higher accuracy of tablet approaches, would apply when asking respondents to draw travel paths. Third, our measure of self-rated health and the limited age range of our sample (ages 15–50) make it difficult to conclude exactly what aspects of health interact with survey mode to influence accuracy. While we suspect the tablet mode made it easier for older respondents in poorer health to provide spatial information, another study with a wider age range and measures of specific functional limitations will be needed to fully understand how technology may benefit these individuals.

In addition, researchers should carefully consider the setting and research infrastructure before favoring tablet technology over paper for spatial data collection. In our project, the limitations of these devices did not outweigh the advantages. Interviewers had access to reliable power sources, the devices were not burdensome to carry on bicycle or motorbike, and the screens were bright enough to use outside (or tree cover and shade was easily available). There are also cost issues. It is clear that electronic devices cost more than paper, although the lower data entry and processing cost can offset the hardware costs. What may be less obvious is research support needed for fielding a tablet-based data collection: devices need to be maintained and continuously updated with software patches; there must be procedures for security and electronic data management. For spatial data collection, it is likely that custom software will have to be written. Our data collection partner, ISER-N, is a mature organization with professional research staff, a history of using CAPI methods for standard questionnaires, and a dedicated IT manager to work with us on our tablet software as we developed, piloted, and fielded the final version. Projects that do not have access to this infrastructure need to factor in the costs of acquiring this critical support.

We also note that our findings may not completely apply to other innovative ways to collect activity space data, such as prospective methods. Wearable GPS data loggers and mobile phones offer the possibility of continual monitoring of activity spaces and highly accurate spatial measurement. These technologies, however, have a different set of weaknesses compared to retrospective methods. The battery life of these devices is not long enough to be used without recharging by the respondent during data collection (Vazquez-Prokopec et al., 2009), and they may not be feasible in populations without reliable electricity sources. Using the respondents' mobile phones is a promising alternative for research in some settings (Matthews, 2011). Nevertheless, mobile phone penetration is far from universal and respondents may not agree to participate in a project that is perceived as invading their privacy – even when the project passes ethical and IRB review. Finally, unless the researcher can convince the respondent and/or data carrier to release administrative or passively collected data, a respondent's personal mobile device cannot easily assist the collection of retrospective data. In sum, even as the proportion of study populations that carry mobile phones increases, there remains a need for solutions to collect retrospective spatial data.

In the world of both researcher-collected data and volunteered geographic information, respondents – even those in remote locations – are increasingly aware of geospatial technologies that support data sharing, navigation, and location services. These types of digital services potentially enhance both researchers' insights into human behavior and individuals' access to new locations and broader awareness of the geographic region. While this particular study explored the use of interactive digital technology versus static paper maps, we did not test possible interface design variations within the tablet mode that might increase accuracy or other aspects of data quality. Fundamental properties of human-computer interaction (HCI), which is a large and separate research field in itself, shaped how respondents provided information using the tablets. In future research, a more thorough incorporation of HCI principles will likely improve how respondents provide spatial information in electronic formats as this mode becomes common practice.

Acknowledgements

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development under award number R21HD073758 and the National Science Foundation under Grant No. 1211498, Dynamics of Coupled Natural and Human Systems Program. We acknowledge assistance provided by the Population Research Institute at Pennsylvania State University, which is supported by an infrastructure grant by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P2CHD041025). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation

References

Allen, G. L. (1999). Spatial abilities, cognitive maps, and wayfinding: bases for individual differences in spatial cognition and behavior. In R. G. Golledge (Ed.), *Wayfinding behavior: cognitive mapping and other spatial processes* (pp. 46–80). Johns Hopkins University Press: Baltimore.

Axinn, W. G., Barber, J., & Ghimire, D. (1997). The neighborhood history calendar: a data collection method designed for dynamic multilevel modeling. *Sociological Methodology*, *27(1)*, 355–392.

Barclay, T., Gray, J., & Slutz, D. (2000). Microsoft TerraServer: a spatial data warehouse. In *Sigmod '00 proceedings of the 2000 ACM SIGMOD international conference on management of data.* New York: ACM Press.

Bates, S. C. & Cox, J. M. (2008). The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors. *Computers in Human Behavior*, *24(3)*, 903–916.

Bederson, B. B. (2011). The promise of zoomable user interfaces. *Behaviour & Information Technology*, *30(6)*, 853–866.

Bobeth, J., Schrammel, J., Deutsch, S., Klein, M., Drobics, M., Hochleitner, C., & Tscheligi, M. (2014). Tablet, gestures, remote control? Influence of age on performance and user experience with iTV applications. In *Proceedings of the 2014 ACM international conference on interactive experiences for TV and online video.* (pp. 139–146). New York: ACM Press.

Boschmann, E. E. & Cubbon, E. (2014). Sketch maps and qualitative GIS: using cartographies of individual spatial narratives in geographic research. *The Professional Geographer*, *66(2)*, 236–248.

Brandl, P., Forlines, C., Wigdor, D., Haller, M., & Shen, C. (2008). Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proceedings of the working conference on advanced visual interfaces* (pp. 154–161). New York: ACM Press.

Brewer, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 25–49). Cambridge, England: Cambridge University Press.

Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: ecological and traditional approaches to the study of memory* (pp. 21–90). Cambridge, England: Cambridge University Press.

Browning, C. R. & Soller, B. (2014). Moving beyond neighborhood: activity spaces and ecological networks as contexts for youth development. *Cityscape*, *16*, 165–196.

Burigat, S., Chittaro, L., & Gabrielli, S. (2008). Navigation techniques for small-screen devices: an evaluation on maps and web pages. *International Journal of Human-Computer Studies*, *66(2)*, 78–97.

Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., . . . Erens, B. (2016). Using the web to collect data on sensitive behaviours: a study looking at mode effects on the british national survey of sexual attitudes and lifestyles. *PloS one*, *11(2)*, e0147983. doi:https://doi.org/10.1371/journal.pone.0147983

Caprani, N., O'Connor, N., & Gurrin, C. (2012). Touch screens for the older user. In F. A. Auat Cheein (Ed.), *Assistive technologies*. Intech.

Caviglia-Harris, J., Hall, S., Mulllan, K., Macintyre, C., Bauch, S. C., Harris, D., . . . Cha, H. (2012). Improving household surveys through computer-assisted data collection use of touch-screen laptops in challenging environments. *Field Methods*, *24(1)*, 74–94.

Cooperstock, J., Fels, S., Buxton, W., & Smith, K. (1997). Reactive environments: throwing away your keyboard and mouse. *Communications of the ACM*, *40(9)*, 65–73.

Couper, M. (2000). Usability evaluation of computer-assisted survey instruments. *Social Science Computer Review*, *18*, 384–396.

Couper, M. (2008). *Designing effective web surveys*. New York: Cambridge University Press.

Crowder, K. & South, S. (2011). Spatial and temporal dimensions of neighborhood effects on high school graduation. *Social Science Research*, *40*, 87–106.

Culén, A. L. & Bratteteig, T. (2013). Touch-screens and elderly users: a perfect match? *Changes*, *7*, 15.

Downs, R. M. (1981). Maps and metaphors. *The Professional Geographer*, *33(3)*, 287–293.

Eggs, J. & Jaeckle, A. (2015). Dependent interviewing and sub-optimal responding. *Survey Research Methods*, *9(1)*, 15–29.

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

Fanning, J. & McAuley, E. (2014). A comparison of tablet computer and paper-based questionnaires in healthy aging research. *JMIR research protocols*, *3(3)*.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment, 80(1)*, 185–201.

Forlines, C., Wigdor, D., Shen, C., & Balakrishnan, R. (2007). Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 647–656). New York: ACM Press.

Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: a technique for collecting retrospective data. *Sociological Methodology*, *18*, 37–68.

Fricker, R. D. & Schonlau, M. (2002). Advantages and disadvantages of internet research surveys: evidence from the literature. *Field methods*, *14(4)*, 347–367.

Geisen, E. & Bergstrom, J. R. (2017). *Usability testing for survey research*. Morgan Kaufmann.

Gourmelon, F., Le Guyader, D., & Fontenelle, G. (2014). A dynamic gis as an efficient tool for integrated coastal zone management. *ISPRS International Journal of Geo-Information*, *3*(2), 391–407.

Guelke, L. (1979). Perception, meaning and cartographic design. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *16(1)*, 61–69.

Gummer, T. & Roßmann, J. (2015). Explaining interview duration in web surveys a multilevel approach. *Social Science Computer Review*, *33(2)*, 217–234.

Hägerstraand, T. (1970). What about people in regional science? *Papers in regional science*, *24(1)*, 7–24.

Harada, S., Sato, D., Takagi, H., & Asakawa, C. (2013). Characteristics of elderly user behavior on mobile multi-touch devices. In *IFIP conference on human-computer interaction* (pp. 323–341). Springer.

Harrower, M. & Sheesley, B. (2005). Designing better map interfaces: a framework for panning and zooming. *Transactions in GIS*, *9(2)*, 77–89.

Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., . . . Buxton, B. (2010). Manual deskterity: an exploration of simultaneous pen + touch direct input. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2793–2802). New York: ACM Press.

Inagami, S., Cohen, D. A., & Finch, B. K. (2007). Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social Science & Medicine*, *65*(8), 1779–1791.

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly*, *68(1)*, 94–101.

Konečný, M., Kubíček, P., Stachoň, Z., & Sašinka, bibinitperiod. (2011). The usability of selected base maps for crises management – users' perspectives. *Applied geomatics*, *3*(4), 189–198.

Krosnick, J. A. (1992). The impact of cognitive sophistication and attitude importance on response-order and question-order effects. In *Context effects in social and psychological research* (pp. 203–218). New York: Springer.

Kuai, X. & Zhao, Q. (2017). Examining healthy food accessibility and disparity in Baton Rouge, Louisiana. *Annals of GIS*, 1–14.

Kwan, M.-P., Peterson, R. D., Browning, C. R., Burrington, L. A., Calder, C. A., & Krivo, L. J. (2008). Reconceptualizing sociogeographic context for the study of drug use, abuse, and addiction. *Geography and Drug Addiction*, 437–446.

Li, Z. & Ho, A. (2004). Design of multi-scale and dynamic maps for land vehicle navigation. *The Cartographic Journal*, *41*, 265–270.

Lipperman-Kreda, S., Morrison, C., Grube, J. W., & Gaidus, A. (2015). Youth activity spaces and daily exposure to tobacco outlets. *Health & Place*, *34*, 30–33.

MacKay, B., Dearman, D., Inkpen, K., & Watters, C. (2005). Walk 'n scroll: a comparison of software-based navigation techniques for different levels of mobility. In *Proceedings of the 7th international conference on human computer interaction with mobile devices & services.* (pp. 183–190). New York: ACM Press.

Mather, R., de Boer, M., Gurung, M., & Roche, N. (1998). Aerial photographs and 'photo-maps' for community forestry. *Rural Development Forestry Network*, *23*.

Matthews, S. A. (2011). Spatial polygamy and the heterogeneity of place: studying people and place via egocentric methods. In *Communities, neighborhoods, and health* (pp. 35–55). Springer.

Pebley, A. & Sastry, N. (2004). Los Angeles Family and Neighborhood Study (L.A.FANS). Original RAND User Guide, Wave 1, Public Use. Retrieved from http://www.icpsr.umich.edu/cgi-bin/file?comp=none%5C&study=22940%5C&ds=0%5C&file_id=991074

Peterson, S. E., Kulhavy, R. W., Stock, W. A., & Pridemore, D. R. (1991). How map features cue associated verbal content. *Bulletin of the Psychonomic Society*, *29*, 158–160.

Poole, P. (1995). Land-based communities, geomatics and biodiversity conservation. *Cultural Survival Quarterly*, *18(4)*, 74–76.

Ramirez-Gomez, S. O., Torres-Vitolas, C. A., Schreckenberg, K., Honzák, M., Cruz-Garcia, G. S., Willcock, S., … Poppy, G. M. (2015). Analysis of ecosystem services provision in the colombian amazon using participatory research and mapping techniques. *Ecosystem Services*, *13*, 93–107.

Reilly, D., Rodgers, M., Argue, R., Nunes, M., & Inkpen, K. (2006). Marked-up maps: combining paper maps and electronic information resources. *Personal and Ubiquitous Computing*, *10(4)*, 215–226.

Rogerson, P. A. (2014). *Statistical methods for geography: a student's guide*. Sage.

Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using google street view to audit neighborhood environments. *American Journal of Preventive Medicine*, *40(1)*, 94–100.

Sampson, R. J. & Raudenbush, S. W. (1999). Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *American Journal of Sociology*, *105(3)*, 603–651.

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

Sandberg, E. H. & Huttenlocher, J. (2001). Advanced spatial skills and advance planning: components of 6-year-olds'navigational map use. *Journal of Cognition and Development*, *2*, 51–70.

Sastry, N. & Pebley, A. (2010). Family and neighborhood sources of socioeconomic inequality in children's achievement. *Demography*, *47(3)*, 777–800.

Sauer, C., Auspurg, K., Hinz, T., & Liebig, S. (2011). The application of factorial surveys in general population samples: the effects of respondent age and education on response times and response consistency. *Survey Research Methods*, *5(3)*, 89–102.

Schneiderman, B. & Plaisant, C. (1998). *Designing the user interface*. Don Mills, Ontario: Addison Wesley.

Schober, M. & Conrad, F. (1997). Does conversational interviewing reduce survey measurement error. *Public Opinion Quarterly*, *61*, 576–602.

Schoepfer, I. & Rogers, S. R. (2014). A new qualitative GIS method for investigating neighbourhood characteristics using a tablet. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *49*(2), 127–143.

Schwartz, N. H. & Kulhavy, R. W. (1981). Map features and recall of discourse. *Contemporary Educational Psychology*, *6*, 151–158.

Scott, S. D., Shoemaker, G. B., & Inkpen, K. (2000). Towards seamless support of natural collaborative interactions. In *Proceedings of Graphics Interface, May* (pp. 103–110).

Sears, A. & Shneiderman, B. (1991). High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, *34*, 593–613.

Sharp, G., Denney, J. T., & Kimbro, R. T. (2015). Multiple contexts of exposure: activity spaces, residential neighborhoods, and self-rated health. *Social Science & Medicine*, *146*, 204–213.

Sholl, M. J. & Egeth, H. E. (1982). Cognitive correlates of map-reading ability. *Intelligence*, *6(2)*, 215–230.

Taveira, A. D. & Choi, S. D. (2009). Review study of computer input devices and older users. *Intl. Journal of Human–Computer Interaction*, *25*(5), 455–474.

Van Wart, S., Tsai, K. J., & Parikh, T. (2010). Local ground: a paper-based toolkit for documenting local geo-spatial knowledge. In *Proceddings of the first ACM symposium on computing for developement*.

van Heerden, A. C., Norris, S. A., Tollman, S. M., & Richter, L. M. (2014). Collecting health research data comparing mobile phone-assisted personal interviewing to paper-and-pen data collection. *Field Methods*, *26(4)*, 307–321.

van Heeswijck, T., Paquet, C., Kestens, Y., Thierry, B., Morency, C., & Daniel, M. (2015). Differences in associations between active transportation and built environmental exposures when expressed using different components of individual activity spaces. *Health & Place*, *33*, 195–202.

Vazquez-Prokopec, G. M., Stoddard, S. T., Paz-Soldan, V., Morrison, A. C., Elder, J. P., Kochel, T. J., … Kitron, U. (2009). Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *International Journal of Health Geographics*, *8*(1), 68.

Watson, N. & Wilkins, R. (2015). Design matters: the impact of CAPI on interview length. *Field Methods*, *27(3)*, 244–264.

Wiegand, P. (2006). *Learning and teaching with maps*. Routledge.

Williams, J., Healy, H., & Ellis, N. (1999). The effect of imageability and predicability of cues in autobiographical memory. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(3), 555–579.

Wodtke, G., Harding, D., & Elwert, F. (2011). Neighborhood effects in temporal perspective: the impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review*, *76*, 713–736.

Woo, Y., Kim, S., & Couper, M. P. (2014). Comparing a cell phone survey and a web survey of university students. *Social Science Computer Review*, *33(3)*, 399–410.

Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly*, 331–353.

Yee, K.-P. (2004). Two-handed interaction on a tablet display. In *CHI'04 extended abstracts on human factors in computing systems* (pp. 1493–1496). ACM.