Validating Smartphone- and Computer-based Technologies with GPS for Activity Tracking

Anna E Bradshaw

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2018

Committee:

Adam Drewnowski

Anju Aggarwal

Program Authorized to Offer Degree:

Nutritional Sciences Program

University of Washington

Abstract

Validating Smartphone- and Computer-based Technologies with GPS for Activity Tracking

Anna E Bradshaw

Chair of the Supervisory Committee:

Adam Drewnowski

Nutritional Sciences Program

The space in which we live and complete our daily activities such as shopping, eating, and working is known as the activity space. Measuring activity space can provide insights into the relation between built environment and health outcomes. Historically, geospatial research in public health was conducted using paper-and-pencil travel logs. It has since moved to Global Positioning System (GPS)-enabled instruments and to computer-assisted interviews. The purpose of this study was to validate a newly developed computer-assisted instrument, Karma, against a traditionally used GPS instrument and a smartphone-based application, MapMyRun (MMR), to study activity space. 12 participants, recruited in the spring of 2018, were asked to collect data using the three instruments over the same three days. Four primary outcome variables were tested for each participant-day (n=29): dwell point count, active dwell duration (in minutes), travel time (in minutes), and track length (in kilometers). Statistically significant correlations were observed for active dwell duration, travel time and track length from Karma with both GPS (satellite-based instrument) and MMR (smartphone-based instrument). The

only exception was the dwell point count variable that did not show significant correlation between Karma and GPS. Additional analyses suggested slightly different travel patterns for food shopping days vs. non-food shopping days and for weekends vs. weekdays. Limited sample size did not allow further stratified analyses. Despite a small sample size, the present findings suggest potential use of Karma to measure activity space in lieu of GPS instruments. Further studies are needed to test the use of Karma with a larger sample size and in population segments that depend on modes of transportation other than car for their primary travel.

Table of Contents

Introduction

I. Background

Measuring activity space, or the space where people dwell for specific purposes such as working or shopping, has been a point of interest for public health researchers.[1-3] Looking at the active environment where one shops, sleeps, works, and eats provides context for larger public health issues such as disparities in food access or in health outcomes.[4,5] Policies and public health initiatives that aim at promoting health equity would benefit from a better understanding of the environmental context of health behaviors and other daily activities.[6-8] Measuring activity spaces provides insights into environmental exposure and its likely impact on health.[4,7]

Historically, activity space has been measured using travel recalls, extensive paper diaries that required participants to record when, where, and how they completed their trips.[9] Every block, every turn, and every stop was recorded as they went about their day.[10] These paper-and-pencil interviews (PAPI), were later joined with similar computer-assisted telephone interviews (CATI) and computer-assisted self-interviews (CASI).[9] Due to the heavy participant burden, as Global Positioning System (GPS) technology was being developed, it was quickly adopted as a viable means of measuring activity space.[9] Originally installed into cars, the technology eventually progressed to smaller, handheld devices that could be attached one's person.[11]

Despite its benefits, GPS technology was not without limitations.[12] The instruments were expensive and easy for participants to lose.[12] GPS datasets were often unwieldy, with thousands of data that required extensive cleaning to eliminate "noise" like loss of signal and signal drifting.[13] In large cities, the instruments suffered from the "urban canyon" effect, an increased rate of signal loss due to high-rise buildings blocking the GPS line-of-sight to satellite signal.[14] Finally, interpreting travel mode and trip purpose from the provided data was challenging without extensive use of computer algorithms or participant reports.[13]

Newly developed computer-based software has also been tested as a potential means of recording simplified travel information from participants, with the software filling in potential gaps through services like Google Maps and others.[15] While these instruments offer great promise, they still must be validated against more traditional instruments to ensure that the data they provide remains accurate and useful for researchers.[15]

II. Study Purpose and Specific Aims

The purpose of this study was to validate Karma, a novel computer-based interview instrument designed to capture habitual activity space, against traditionally used instruments in the field such as a GPS tracker instrument and a smartphone-based application, MapMyRun (MMR). This is one of the first studies to test three different methods of gathering geospatial data concurrently. Whereas GPS assesses positions based on satellite signals, MMR relies on a combination of cellphone towers and Wi-Fi.[9,12]

The specific aim of the study was to examine the degree of concordance between a computer-assisted instrument (Karma) with a satellite-based instrument (GPS) and a smartphone-based application (MMR) using four activity space measures (dwell counts, active dwell duration, travel time and track length).

The secondary aim of the study was to examine if the habitual travel patterns, as captured by the three instruments, vary by food shopping days vs. non-food shopping days, and by weekdays vs. weekends.

Smartphone applications have been brought to the forefront of conversation as a potential solution to the participant burden of GPS instruments.[16] The ubiquity of smartphone instruments coupled with the ease of downloading applications with built-in GPS capabilities shows incredible promise for measuring geospatial movement.[16,17]

MMR is a free smartphone application (https://www.mapmyrun.com/) owned by Under Armour® to track distance and speed for athletes as they run.[18] After being developed in 2005, the application rose quickly as a potential fitness tracking application and has gained over 20 million registered users since launching onto the smartphone market.[18]

Karma is a computer-based travel recall software developed by the University of Washington specifically to measure activity space. As an instrument to measure activity space, Karma is akin to other CATI instruments, but is streamlined to minimize participant burden. By using the search engines Google Maps (https://www.google.com/maps/) and Bing Maps (https://www.bing.com/maps), Karma is able to track an estimated trip pathway from destinations provided by participants. This is in contrast to MMR and GPS, which use tracking to estimate destinations. In addition to reporting their destinations, participants report dwell times and trip times, creating a comprehensive outline of their activity patterns for the day in conjunction with the maps generated through the instrument. Through the testing of the validity of these instruments in comparison to the GPS, new avenues can be explored for activity space data collection.

Methods

I. Background Research of Study Instruments

To determine an appropriate smartphone application to compare to Karma, a literature review was conducted over the summer of 2017 to narrow down potential candidates. For the smartphone application to be useful in the study, it needed to be free for participants who would be installing it on their phones, be intuitive to use, be accessible on both Android and iPhone instruments, and provide a means for researchers to download individuals' data without accessing their personal information such as email addresses and passwords. In addition to all of these conditions, the application also needed to provide data collected in discrete time intervals. The potential instrument also needed to provide values for distance, longitude, and latitude. These variables would provide information that could later be compared against validated instruments.

Two free smartphone fitness applications, MMR and Moves (http://moves-app.com/), were tested in June 2017 for their potential as data collection instruments. Similarly to MMR, Moves was developed by a smaller company and later purchased by a larger corporation, Facebook, in 2014.[19] After collecting data on the instruments over a period of five days, the respective datasets were downloaded and compared. While Moves provided distance and estimated mode of transportation, it provided an aggregated distance for the day, making it impossible to compare against the individual time intervals found in the GPS instrument dataset. In contrast, MMR provided one-second time intervals, which could be compared against the GPS dataset. Because of the limitations of Moves, MMR was chosen as the smartphone application to be used in this study. It should be noted that in addition to its limitations, the Moves application has been cancelled by Facebook as of July 2018.[19]

II. Pilot Study

Prior to data collection, a pilot study was conducted in January 2018 to test the three instruments (the QSTARZ Travel Recorder XT GPS instrument, MMR, and Karma) in the field.[20] Three days of data were collected during the pilot study and then analyzed using preliminary data analysis algorithms and techniques. The analysis showed a potential for Karma as a means of measuring activity space and provided further background to support moving forward with the study. In addition, the pilot data narrowed potential variables of interest to the data values that directly measured time and distance for participants' recorded days.

III. Data Collection Phase

A. Preparing the GPS Instrument and MMR Application

To prepare the GPS tracker, the Travel Recorder XT, the instrument was configured using QTravel software.[21] The instrument was set to record data in one-minute intervals that provided information on the longitude and latitude coordinates, satellite information, and distance travelled for each time interval. Following and individual's data collection period, the information was pulled from the GPS instrument, and the instrument was cleared and reconfigured for future data collection. For MMR, individual accounts with unique passwords were generated for participants to use. These accounts were attached to dummy email addresses that would be cleared following the study.

B. Recruiting Participants

Participants were recruited through April 2018. 18 participants were initially recruited by word-of-mouth for the study. After a preliminary screen was conducted over the phone, 12 potential participants completed in-person meetings with the graduate researcher. During these meetings participants were briefed on study protocol and provided consent forms approved by the University of Washington Institutional Review Board. A demographic survey was conducted after consent was obtained. Finally, participants were provided individual GPS instruments, MMR accounts, Karma accounts, and were provided both written and verbal explanations on how to use the instruments. As a final step to the in-person meeting, participants scheduled times during which a graduate researcher would contact them over the phone to complete the Karma survey portion of data collection.

C. Primary Data Collection

For three days, each participant was instructed to turn on their GPS instrument and MMR smartphone application before leaving the house each day. Participants were encouraged to keep both instruments running throughout the day to provide a consistent dataset but were instructed to pause the MMR data collection while in activity spaces if their phone was low on battery. This ensured that data would accurately reflect travel patterns and allow for dwell point to be estimated from the missing time. During the data collection period, participants were encouraged to record at least one food shopping day using either their primary or secondary food shopping locations. Participants completed their data collection for each day once they had returned home for the final time. At this point, participants were instructed to turn off the instruments and charge the instruments. If a participant completed an impromptu trip from home after turning off the instruments, they were instructed to turn them back on to collect the new data.

Karma data collection was staggered to the GPS and MMR instruments. To complete the data collection for Karma, participants were contacted over the phone at designated times over three days to complete a travel recall for the day prior. Participants were asked to provide information on their common activity spaces such as where they shopped and worked, and then they were asked to recall the trips and destinations that they took during the day in a two-pass method. In the first pass, participants were asked to provide a general timeline of their activities for the day. During the second pass, the researcher conducting the interview would review the listed trips and destinations and ask for specific details on when trips began and ended and what mode of transportation was used. In addition to recording the trip information on Karma, the researcher made note on paper of any discrepancies in trip path between the Karma estimated pathway and participants' recollections.

While data collection followed seemingly disparate pathways for each instrument, ultimately the variables collected were similar. Time, distance, and location were provided by all three instruments that could later be analyzed during the data analysis stage.

IV. Data Analyses

A. Downloading Data from Each Instrument

Data were downloaded from the GPS units using QTravel software and stored as comma-separated value (CSV) and project (ITM) files.[21] A series of GPS variables were obtained including time stamp, latitude, longitude, height, speed, heading direction, satellite information, and distance travelled between data points. Quality control of data collected from the GPS was conducted using a custom script in the R statistical software.[22] The custom R script compared data entry counts to ensure that all data points were collected from the instrument.

MMR data was downloaded from the MMR server as TCX files, a text format that was later converted into an Excel file using an in-house algorithm. The MMR files contained UTC time stamp, latitude, longitude, maximum speed for the day, estimated calories burned, altitude, and daily cumulative distance traveled.

Raw Karma data were provided as CSV files containing names and addresses of six key destinations reported by the participant (home address, place of work or school, primary food store, secondary food store, primary fast food restaurant, and primary activity) with up to ten additional destinations. The dataset also contained detailed information on trips taken on each of three days, following the order in which places were visited. Trip-level data included information on start and end times, origin and destination locations, travel mode, duration, and distance traveled using Google Maps. Trip data were also provided as pathways rendered in Google Maps and Bing Maps.

B. Data Cleaning

Several steps were followed for data cleaning. First, a Python algorithm was developed to condense GPS and MMR data to five-minute intervals. Variables were populated using the first data point at or after five minutes from the previous data point. The distance variable was summed over each five-minute interval (to reflect the total distance traveled during that interval). The Karma dataset, being available in five-minute intervals, required no further processing.

The condensed five-minute interval data from GPS and MMR were then collected manually for all participants into a datafile, one for each instrument. Each respondent for each day was assigned a unique participant-day ID, which allowed researchers to study both intra-level, such as different days for the same participant, and inter-level variation in travel pattern. Each participant day was categorized as either a food shopping or non-food shopping day and either a weekday or weekend-day based off of participant recall in Karma.

After the data was condensed, extraneous variables from the dataset were removed. The primary variables of interest that were retained from raw GPS data were: UTC and local time stamps, latitude, longitude, and distance. Extraneous GPS information variables, such as heading, speed, height, and satellite information were removed. A speed variable was created using the distance between data points divided by the time interval. This variable was later used to identify a "dwell" location, discussed in detail below.

Similar procedures were followed to clean data for MMR. Variables of interest were UTC time stamp, latitude, longitude, and distance. The altitude variable was removed. A speed variable was also created from MMR similar to that computed for GPS data. No cleaning was required for the Karma dataset.

Once the data were compiled into 36 participant-day samples, the data were further cleaned to eliminate any missing data collection samples due to participant non-compliance. In doing this, seven samples were eliminated, leaving the final sample size (n=29). The final step in data cleaning eliminated extraneous dwell time due to participants accidentally keeping either the GPS or MMR application "on" after they had finished their data collection period and returned home. For secondary data analysis, data was also stratified into either food shopping or non-food shopping days and weekends or weekdays based on trip information gathered through Karma.

C. Calculating Interim Analytical Variables

In order to calculate the final analytical variables of interest (see **Table 1**), several interim analytical variables needed for the GPS and MMR datasets. The difference in time between data intervals in minutes, calculated distance in meters (for MMR), and calculated speed in kilometers per hour were calculated for each data point. These values were then used to estimate the number of dwell points for both GPS and MMR. Once dwell points were established, the final variables such as track length, dwell duration, and time spent travelling could be calculated. These variables are explained in more detail in **Table 1.**

D. Development of Primary Outcome Variables

Following generation of the interim analytical variables, the primary outcome variables could be calculated. The first primary outcome variable to be calculated for both GPS and MMR was dwell count. After reviewing the literature and pilot data, a rule-based system rather than a probability-based system was used to measure the data, with a threshold at which a person was considered "dwelling" to be less than 4.5 kilometers per hour (km/h) for greater than five minutes.[23-26] Using the conditional formatting feature on Microsoft Excel, all time intervals during which participants were travelling at less than 4.5 km/h for five minutes or greater were established as dwell times.[27] Dwell points were defined as the time interval during which a participant remained within the 4.5 km/h threshold. Once a participant increased their speed past the threshold for five minutes or greater, they were considered "travelling" and the dwell point ended. In addition to the calculated dwell points, a final dwell point was added to samples that did not end with a dwell point due to participants turning off the MMR and GPS instruments before five minutes had passed once they had returned home. The number of discrete dwell points were then summed for the final dwell count value for each participant-day.

After establishing the time intervals during which participants were dwelling, the final three variables could be calculated. Active dwell duration (in minutes) was derived from the summed time intervals spent during each dwell point. Time spent at home once the data collection period ended (once the participant had returned home for the final time) was not included as an active dwell period and so was not calculated into the final active dwell duration variable. Travel time (in minutes) was calculated by summing all time established as non-dwell time. Track length (in kilometers) was calculated by summing the distance data collected during the non-dwell time.

No interim analytical variables were needed for Karma. Instead, each trip destination was counted and summed to calculate dwell count. Track length in miles was provided for each trip, allowing a simple sum of the values to provide the track length variable for each sample. This value was then converted into meters for consistency between instruments. Travel time was summed from Google Map-generated values provided through Karma. An additional subjective travel time was computed based on the time it took the respondent to travel from one destination to the other as self-reported using the Karma instrument. This analytical variable was meant for additional sensitivity analyses on whether self-reported travel times matched with the Google Maps-based objective variable. Dwell duration was calculated as the participant-reported time between the end of one trip and the beginning of another.

Two indicator variables were created to reflect food shopping vs. non-food shopping day, and a weekday vs. weekend. The indicator variables were used to stratify the data into the various sub-categories for secondary data analysis.

Table 1: Primary outcome variable development by instrument

Primary Outcome Variable Name	Method of Variable Development		
	GPS	MapMyRun	Karma
Dwell Count	Dwell points were calculated as time intervals during which a participant was moving < 4.5 km/hour for > five minutes. An additional dwell point was established at the end of each participant-day as needed to account for the dwell point of "home. These points were then summed to generate a final dwell count.	Dwell points were calculated as time intervals during which a participant was moving < 4.5 km/hour for > five minutes. An additional dwell point was established at the end of each participant-day as needed to account for the dwell point of "home. These points were then summed to generate a final dwell count.	Dwell points in Karma were identified as the end points or destinations in a self-reported trip. These points were then summed to generate a final dwell count.
Active Dwell Duration (min)	Dwell duration was established by summing the estimated dwell point time intervals	Dwell duration was established by summing the estimated dwell point time intervals	Dwell duration was calculated as the difference between the end time and beginning time of participant-reported trips
Travel Time (min)	Travel time was established by summing the estimated non-dwell point time intervals	Travel time was established by summing the estimated non-dwell point time intervals	Time spent commuting (in minutes) for each trip on a given day was summed across all the trips for that day to create a total travel time indicator. This indicator was generated using Google Maps programmed in Karma. A corresponding subjective indicator of total travel time was also computed based on the self-reported time that the respondent took to travel from one destination to the other for each trip for a given day.
Track Length (km)	Total distance travelled was established by summing the distanced in the estimated non-dwell point time intervals	Total distance travelled was established by summing the distanced in the estimated non-dwell point time intervals	Total distance traveled was calculated by summing the shortest distance for each trip as provided by Google Maps programmed in Karma.

E. Statistical Data Analyses

Primary Data Analyses

Statistical analyses were conducted using SAS software.[28] Several different analysis methods were used to validate the primary outcome variables between the GPS tracker instrument, MMR, and Karma.

First, basic distributions of the primary outcome variables were studied using histograms, means, and medians. Second, due to continuous nature of the data, a Pearson's correlation coefficient was used in lieu of the kappa statistic to test the correlation between individual data points between instruments.[29-30] Significance was tested using a t-distribution test.[31] Third, in order to study the difference between instrument values for each outcome variable, the Bland-Altman method was used with 95% limits of agreement.[32] This was defined as the difference in instrument outcome variables between ± 1.96 SD of the bias. [32] Using Bland-Altman plots allowed researchers to visualize the agreement between instruments.[32] Fourth, all these analyses were replicated after stratifying by food shopping day versus not; and weekday versus weekend. The final analytical sample for the overall sample was 36 person-days. Among 36-person days, seven were food shopping days and 17 were weekdays.

Additional Sensitivity Analyses

A series of additional analyses were conducted to test the robustness of the findings. First, the travel time variable (time spent commuting per day) as obtained objectively from Google Maps was compared to subjective travel time variable obtained from self-reported data from Karma using a Pearson correlation coefficient.[30]. The purpose of this analysis was to test the degree of concordance across objective vs. self-reported data. Second, dwell count indicator was compared with three other activity space indicators (dwell duration, travel time and track length) to study the degree of correlation across activity space measures.

Results

I. Descriptive Statistics

First, basic distributions of each of the outcome variables were studied using histograms, means (SD) and median (IQR) for the entire sample (n=29) **(Table 2)**. Visual distribution was studied using box plots **(Figure 1)**. Histograms revealed fairly normal distribution for all four activity space measures across three instruments **(Appendix 1)**. Average number of dwell counts per participant per day were comparable across three instruments (Mean(SD): 6.0(2.3) for Karma, 5.9(2.3) for GPS and 6.0(2.9) for MMR) **(Table 2).** Corresponding median values were 6.0, 6.0, and 5.0 respectively **(Fig. 1A)**. For the active dwell duration variable, defined as total time spent dwelling per participant-day, the means were comparable across three instruments (517.1(228.6) minutes for Karma, 527.9(246.2) minutes for GPS and 540.0(239.8) minutes for MMR). Active dwell duration means and medians remained fairly consistent across the three instruments, deviating by less than 40 minutes across instruments for both values. GPS provided the largest IQR **(Fig. 1B).** In terms of the travel time variable, which was defined as non-dwell time spent for each person-day, the mean values were 95.5(59.8) minutes from Karma, 95.2(60.4) for GPS and 96.5(51.5) for MMR. The mean and median differed by less than two and 20 minutes respectively across the three instruments, despite the thousands of minutes recorded **(Fig. 1C)**. As seen in **Fig. 1C**, the total range of data and the IQR between instruments was also similar.

The track length variable, defined as the distance travelled during non-dwell time, had a comparable distribution across all three instruments. The mean(SD) was 36.2(26.0) km for Karma, 34.0(25.1) km for GPS and 36.9(25.4) km for MMR **(Fig. 1D)**.

Additional descriptive analyses were conducted to examine if activity space measures differ across food shopping days vs. non-food shopping days (**Appendix 2**), and weekdays vs. weekends (**Appendix 3**). Among 29 person-days, seven were food shopping days and 22 were non-food shopping days. As shown in **Appendix 2**, the mean number of dwell counts were slightly higher for food shopping days (Mean(SD): 6.3(2.7) from Karma) as compared to non-food shopping days (Mean(SD): 5.9(2.2) from Karma). However, the reverse was observed with GPS and MMR with mean dwell counts higher for non-food shopping days vs. food shopping days (Mean(SD): 6.3(2.4) and 4.7(1.4) respectively from GPS and 6.5(3.1) and 4.7(1.4) from MMR). Corresponding median values from these instruments did not show much difference in dwell counts by food shopping and non-food days across instruments. In terms of other activity space measures (active dwell duration, travel time and track length), the means and medians were consistently higher for food shopping days vs. non-food shopping days across all three instruments **(Appendix 2**). Limited sample size did not allow tests for statistical significance.

Additional analyses stratified by weekdays vs. weekend (analytical n= 17 and 12 respectively) revealed similar travel patterns (**Appendix 3**). The mean number of dwell counts from Karma were the same for weekday vs. weekend (Mean(SD): 6.0(2.4) and 6.0(2.3) respectively). Similar pattern was observed with MMR (5.0(2.3) and 5.0(3.6) respectively). However, slightly higher dwell counts were obtained from GPS for weekdays (with mean(SD): 6.0(2.5)) vs. weekends (4.8(1.4)). No further analyses were conducted due to small sample size.

In terms of other activity space measures, the mean and median active dwell durations, travel time and track length seem to be much higher for weekdays vs. weekends across all three instruments (**Appendix 3)**. Formal tests for statistical significance could not be conducted due to limited power.

Table 2 Distributions of activity space indicators as computed from three instruments: GPS, MMR and Karma (n=29)

	Dwell Count		**Active Dwell Duration (min)**		**Travel Time (min)**		**Track Length (km)**	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
GPS	5.9 (2.3)	6.0 (3.0)	527.9 (246.2)	580.7 (405.7)	95.2 (60.4)	70.4 (65.9)	34.0 (25.1)	30.7 (40.0)
MMR	6.0 (2.9)	5.0 (1.0)	540.0 (239.8)	610.7 (331.1)	96.5 (51.5)	87.8 (61.8)	36.9 (25.4)	33.8 (41.4)
Karma	6.0 (2.3)	6.0 (4.0)	517.1 (228.6)	565.0 (340.0)	95.5 (59.8)	81.0 (51.0)	36.2 (26.0)	32.0 (41.5)

Figure 1 *Boxplots of primary outcome variables by instruments*

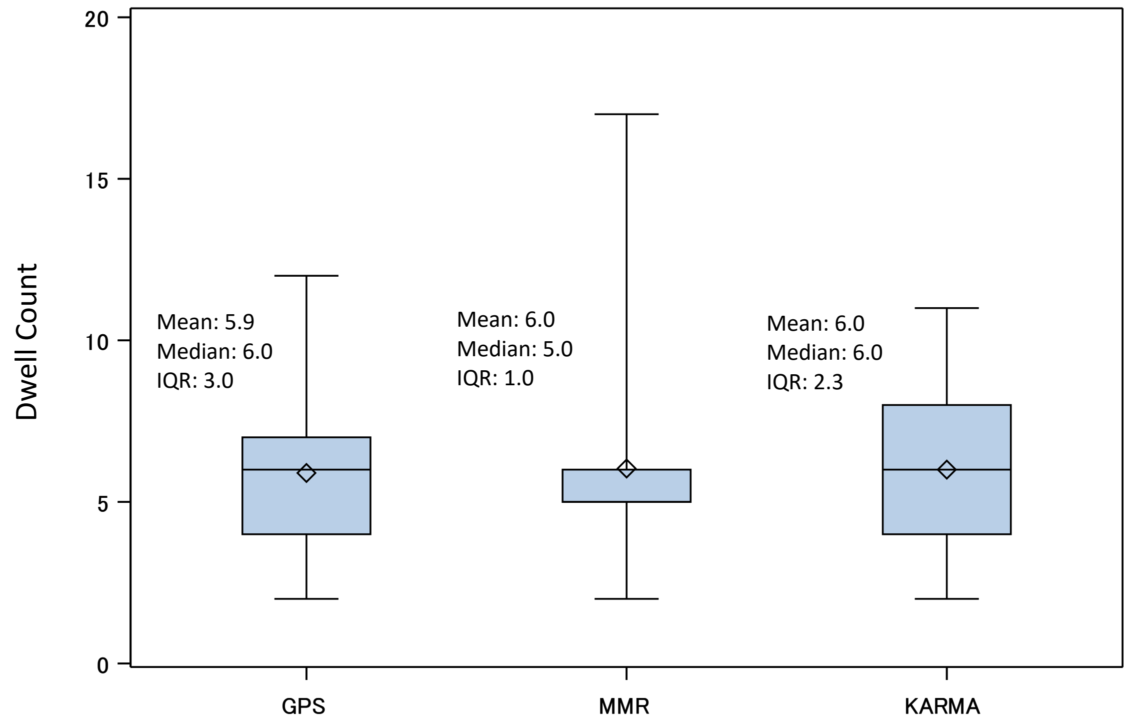


Fig 1(A): Box plots for dwell count by each instrument

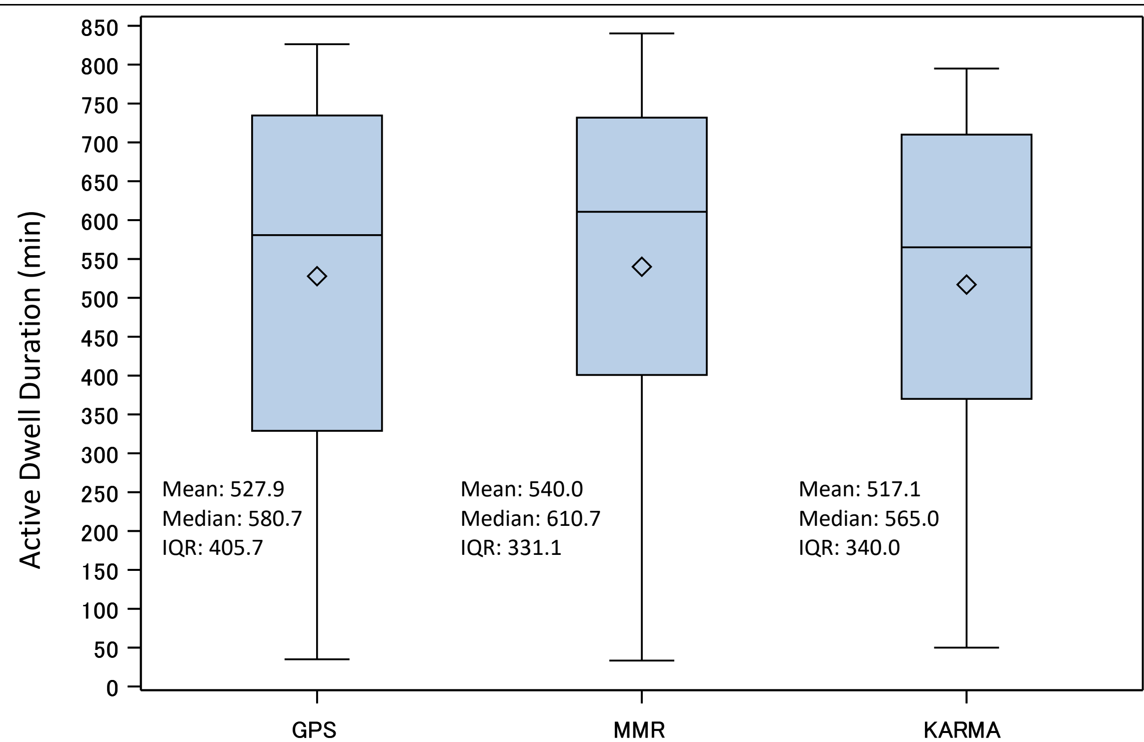


Fig 1(B): Box plots for active dwell duration by each instrument

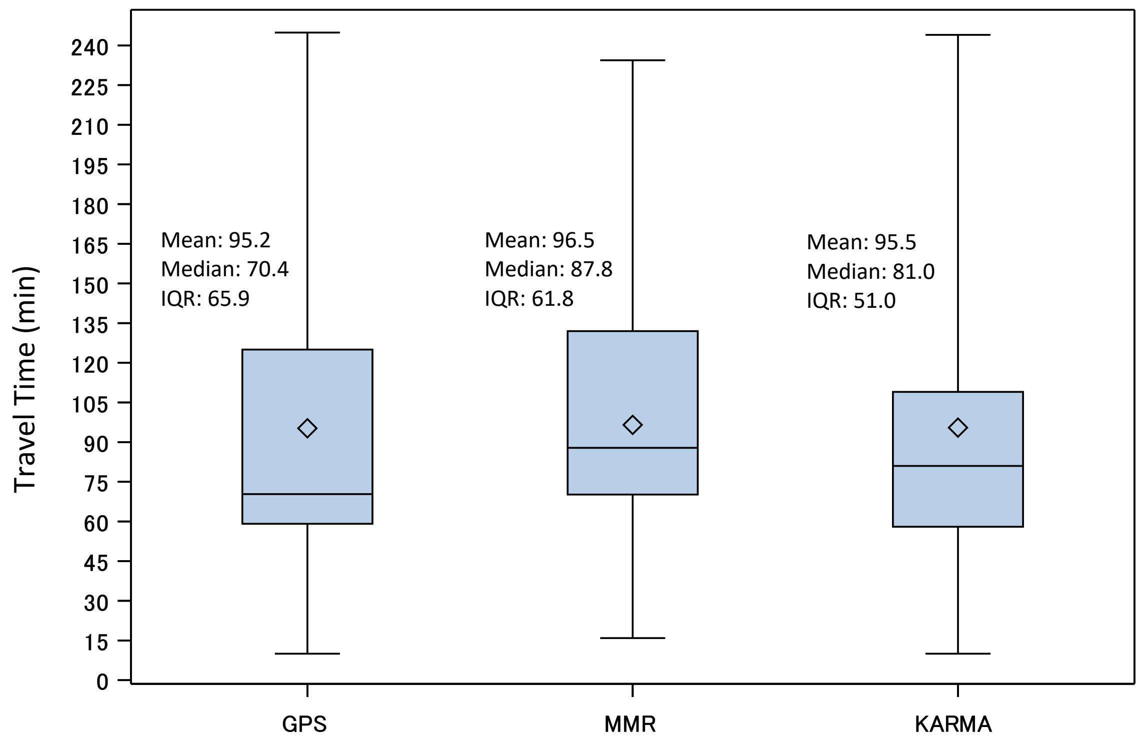


Fig 1(C): Box plots for travel time by each instrument

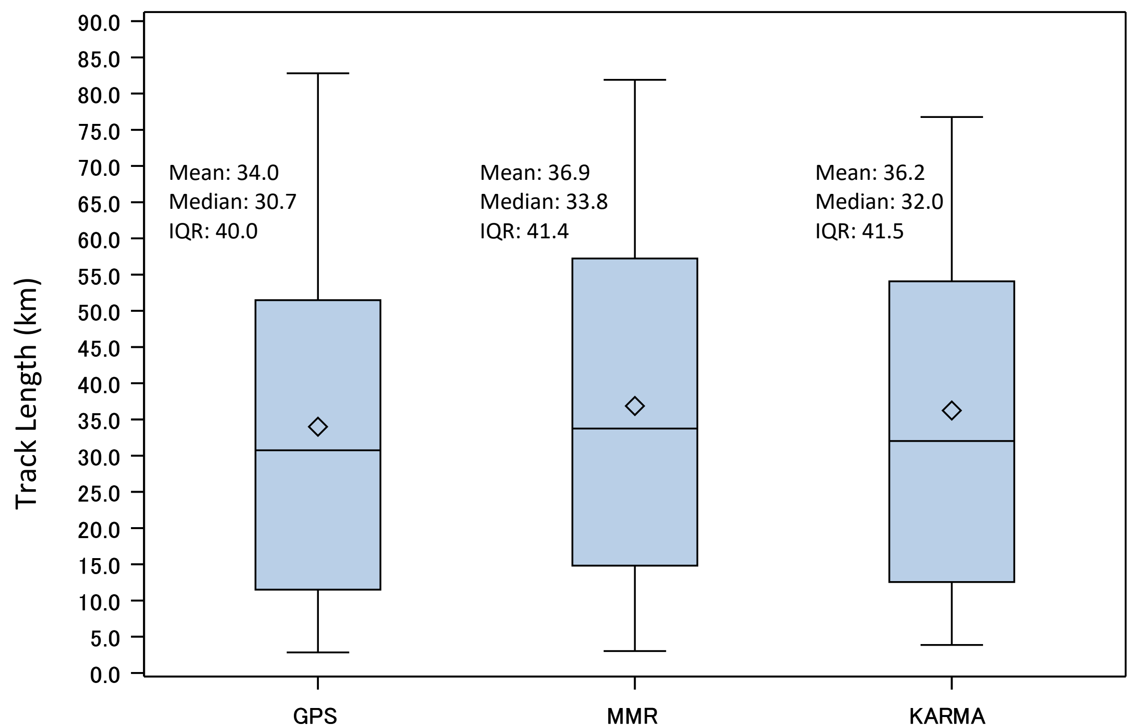


Fig 1(D): Box plots for track length by each instrument

II. Studying Correlations of Activity Space Indicators: Comparing Karma against GPS and MMR

Pearson's correlation coefficients for each outcome variable across three instruments are presented in **Table 3** and **Figure 2**.

In terms of dwell count, Karma demonstrated a stronger correlation with MMR (0.63, p= 0.0003) as compared to GPS (0.43, p=0.02). Active dwell duration showed a correlation of 0.97 (p <0.0001) between GPS and Karma and a correlation of 0.98 (p <0.0001) between Karma and MMR. While not as highly correlated as active dwell duration, travel time (in minutes) had a significant correlation of 0.77 (p <0.0001) between GPS and Karma and 0.83 (p <0.0001) between MMR and Karma. Finally, track length showed significant correlation with GPS and Karma with a correlation of 0.95 (p <0.0001) and Karma and MMR having a correlation value of 0.97 (p <0.0001). All activity space measures except for dwell count across all three instrument comparison groups were significantly correlated with p-values often less than <0.0001. Karma and MMR consistently had stronger correlation in comparison to GPS across all variables.

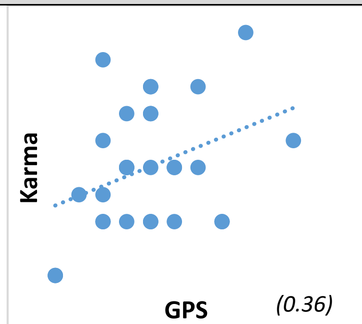
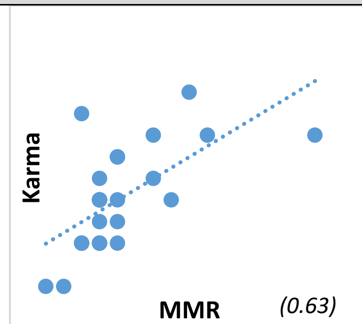
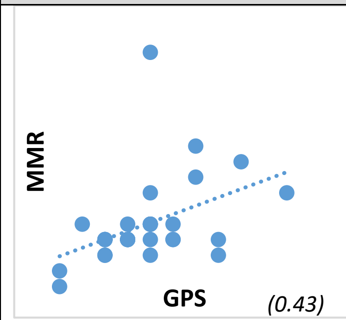
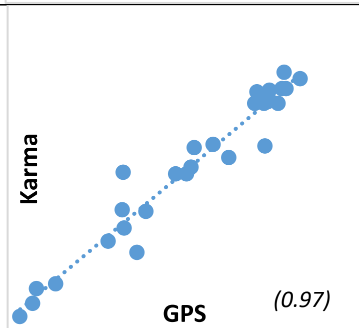
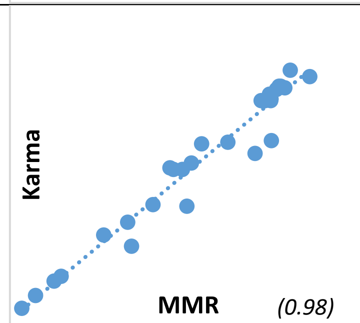
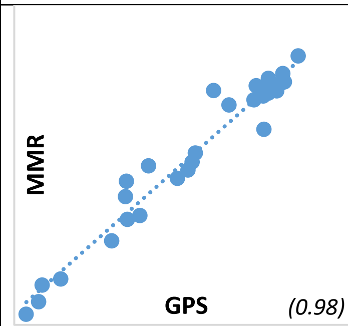
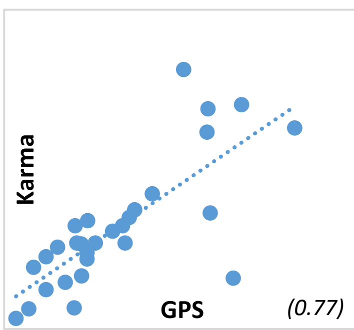
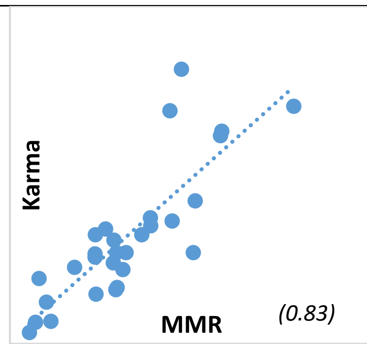
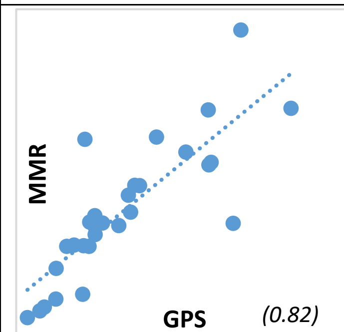
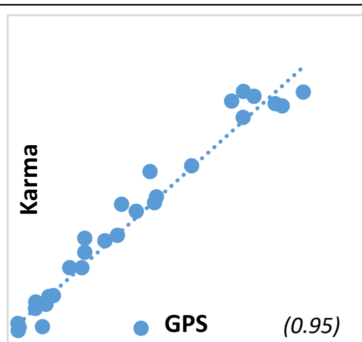
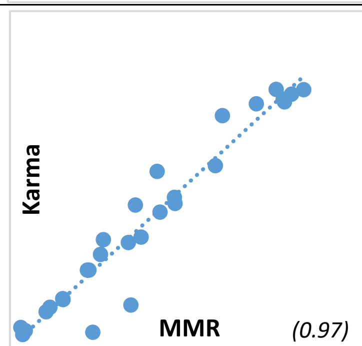
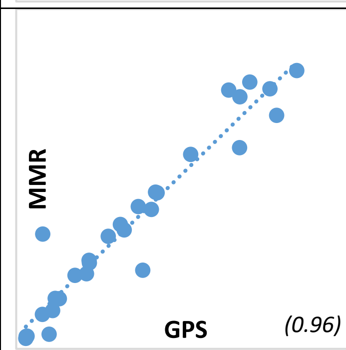
After observing the correlation between GPS and Karma and noting the non-parametric nature of the dwell count variable, a second analysis using a chi-square goodness-of-fit test was conducted across instruments. The values for the various comparisons were 0.18 for GPS and MMR, 0.32 for GPS and Karma, and 0.91 for Karma and MMR. These values indicate that the measures between dwell counts were non-significantly different (**Appendix 4**).

Due to limited power, Pearson's correlation could not be studied across food shopping vs. non-food shopping days and weekdays vs. weekends.

Table 3: Pearson's Correlation Coefficients for primary outcome variables by each instrument comparison per participant-day

	GPS	MMR	Karma
Dwell Count			
GPS	1.0	0.43 (0.02)	0.36 (0.06)
MMR	0.43 (0.02)	1.0	0.63 (0.0003)
Karma	0.36 (0.06)	0.63 (0.0003)	1.0
Active Dwell Duration (min)			
GPS	1.0	0.98 (<0.0001)	0.97 (<0.0001)
MMR	0.98 (<0.0001)	1.0	0.98 (<0.0001)
Karma	0.97 (<0.0001)	0.98 (<0.0001)	1.0
Total Travel Time (min)			
GPS	1.0	0.82 (<0.0001)	0.77 (<0.0001)
MMR	0.82 (<0.0001)	1.0	0.83 (<0.0001)
Karma	0.77 (<0.0001)	0.83 (<0.0001)	1.0
Track Length (km)			
GPS	1.0	0.96 (<0.0001)	0.95 (<0.0001)
MMR	0.96 (<0.0001)	1.0	0.97 (<0.0001)
Karma	0.95 (<0.0001)	0.97 (<0.0001)	1.0

Figure 2: Correlation plots comparing Karma with GPS and MMR. Line-of-best fit (Pearson's Coefficient) are present to illustrate the trend of the data.

	GPS vs Karma	MMR vs Karma	GPS vs MMR
Dwell Count			
Active Dwell Duration (min)			
Total Travel Time (min)			
Track Length (km)			

III. Measuring Agreement: Bland-Altman Plots

To provide visualization for the degree of agreement between Karma and GPS and MMR, Bland-Altman plots were created for each primary outcome variables (**Figures 3-6**). Dwell counts reflected similar trends to the correlation values, with GPS comparisons Karma showing greater, more inconsistent spread from the line of zero bias in comparison to the other variables (**Fig. 3A**). In contrast to the GPS vs. Karma plot, the Bland-Altman plot for dwell count between MMR vs. Karma (**Fig. 3B**) showed clustering closer to the line of zero bias, but also displayed a value that was outside of the 95% upper limit and one outside of the 95% lower limit.

Active dwell duration showed a bias towards underestimating as the amount of time had that been dwelled increased for GPS vs. Karma (**Fig. 4A**). The data also showed several instances of outliers that fell outside of the 95% limits between GPS vs. Karma, indicating that extremely dissimilar values existed in the dataset for active dwell count. Between MMR vs. Karma, several values were above the upper 95% limit (**Fig. 4B**).

Travel time disagreement increased as the time spent travelling increased, indicating that the instruments demonstrated greater potential incongruence as participants travelled longer throughout the day (**Figs. 5A-C**). In addition, several data points fell out of the 95% bounds on both sides for all comparisons.

A similar trend to travel time and its increasing disagreement with higher values can be seen for track length for GPS vs. Karma (**Fig. 6A**). In addition, for GPS vs. Karma a few extreme values fell below the 95% lower threshold only. The comparison of MMR vs. Karma (**Fig. 6B**) showed the opposite, with the extreme values falling above the upper 95% limit. As track length increased, the agreement between instruments decreased. In addition to this, the plots show a bias towards Karma overestimating track length in comparison to GPS. In contrast, MMR tended to underestimate track length in comparison to Karma, especially as the length increased. Overall, MMR provided plots that demonstrated similar trends to Karma when compared to GPS data, reflecting the overall correlation values (**Figs. 3-6C**)

Due to limited power, stratified analyses by food shopping days and weekends could not be studied using Bland-Altman plots.

Figure 3: Bland-Altman plot for dwell count across instrument comparisons

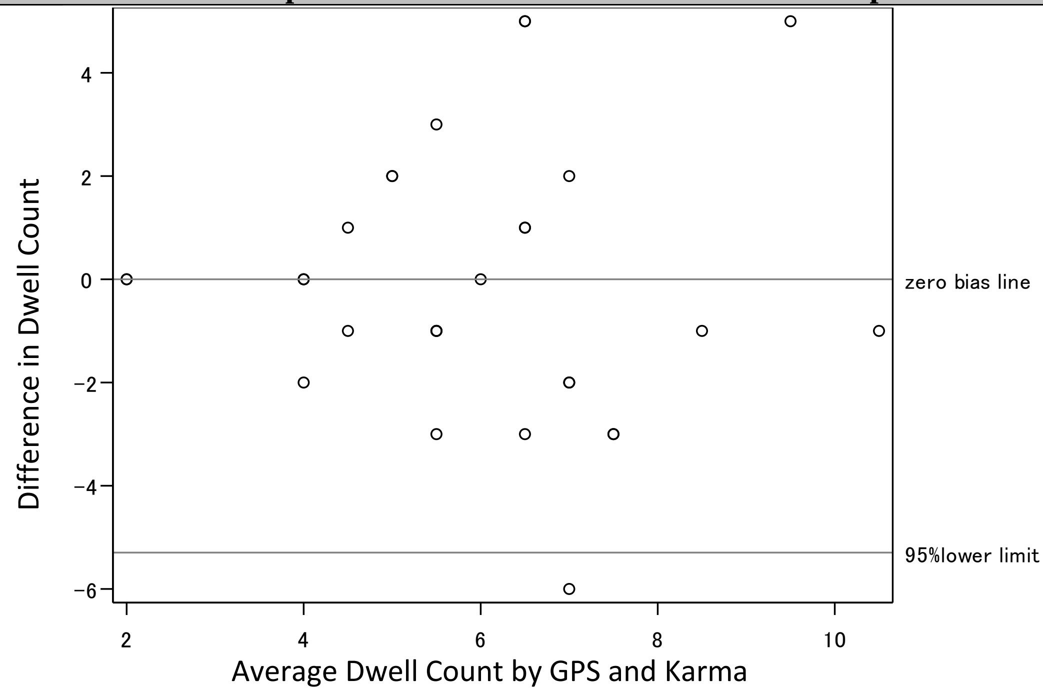


Fig 3(A): Bland-Altman plot between GPS and Karma for dwell count

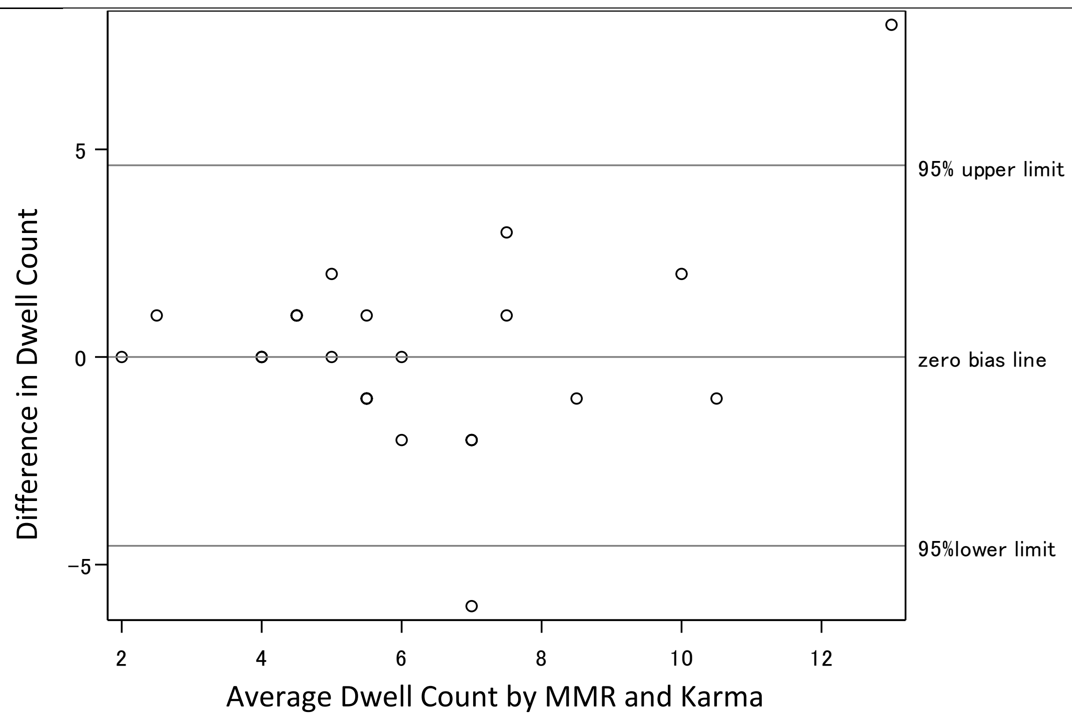


Fig 3(B): Bland-Altman plot between MMR and Karma for dwell count

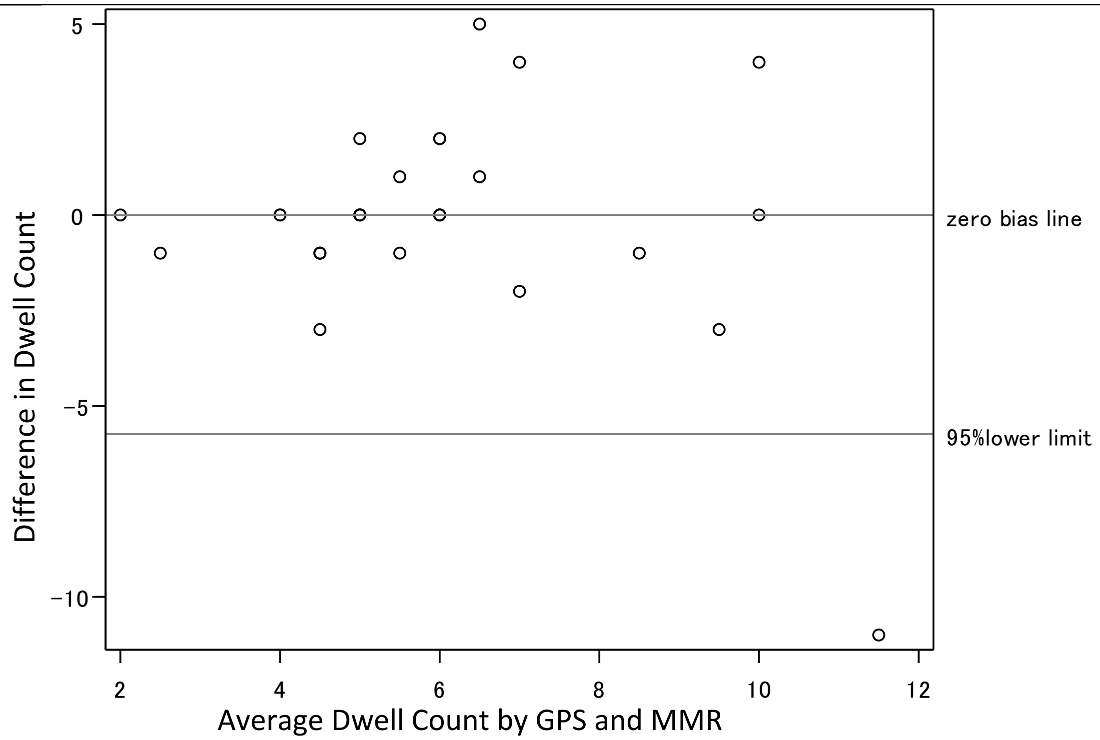


Fig 3(C): Bland-Altman plot between GPS and MMR for dwell count

Figure 4: Bland-Altman plot for active dwell duration across instrument comparisons

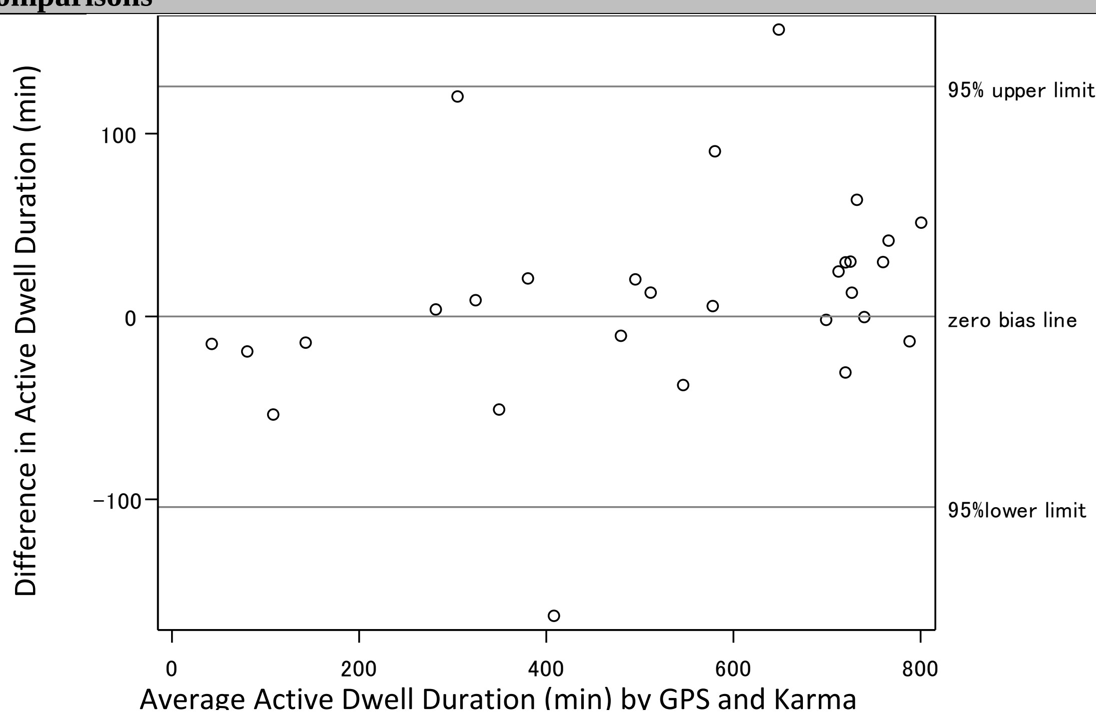


Fig 4(A): Bland-Altman plot between GPS and Karma for active dwell duration (min)

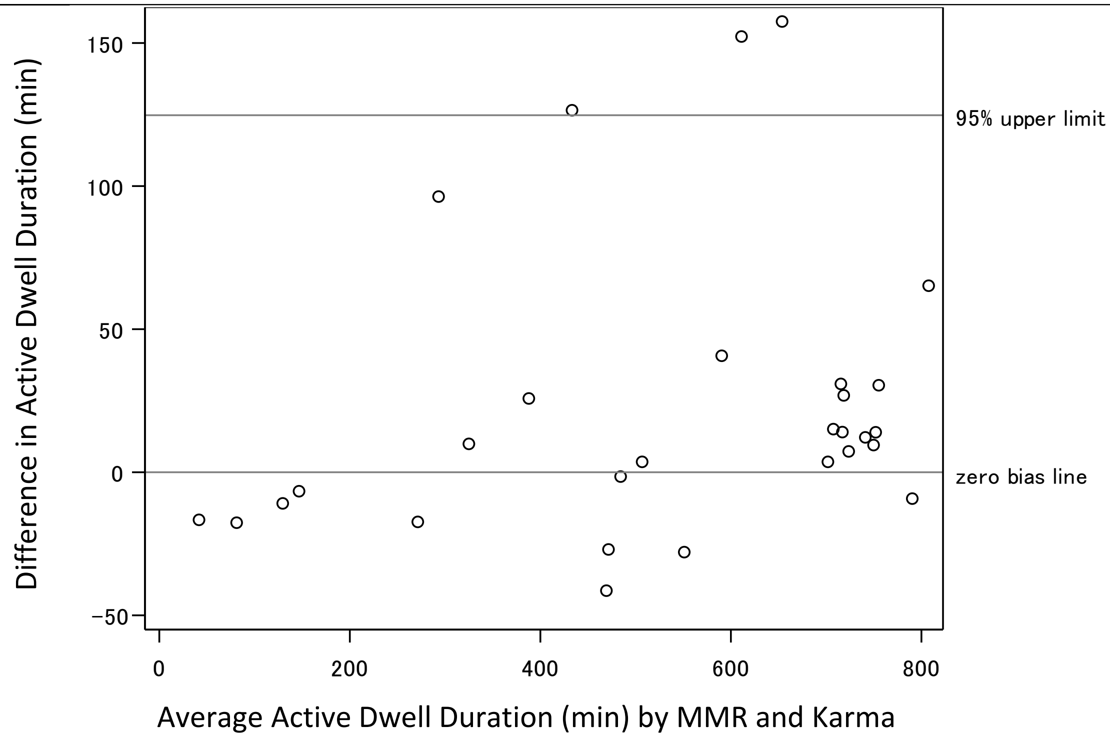


Fig 4(B): Bland-Altman plot between MMR and Karma for active dwell duration (min)

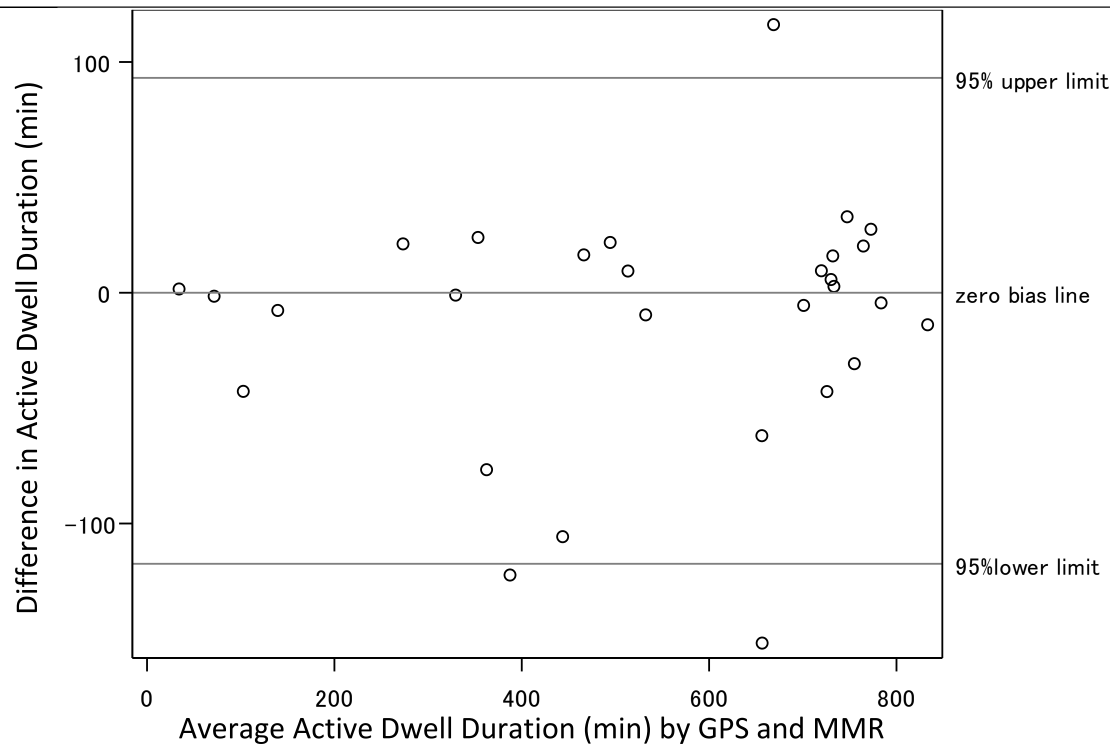


Fig 4(C): Bland-Altman plot between GPS and MMR for active dwell duration (min)

Figure 5: Bland-Altman plot for travel time across instrument comparisons

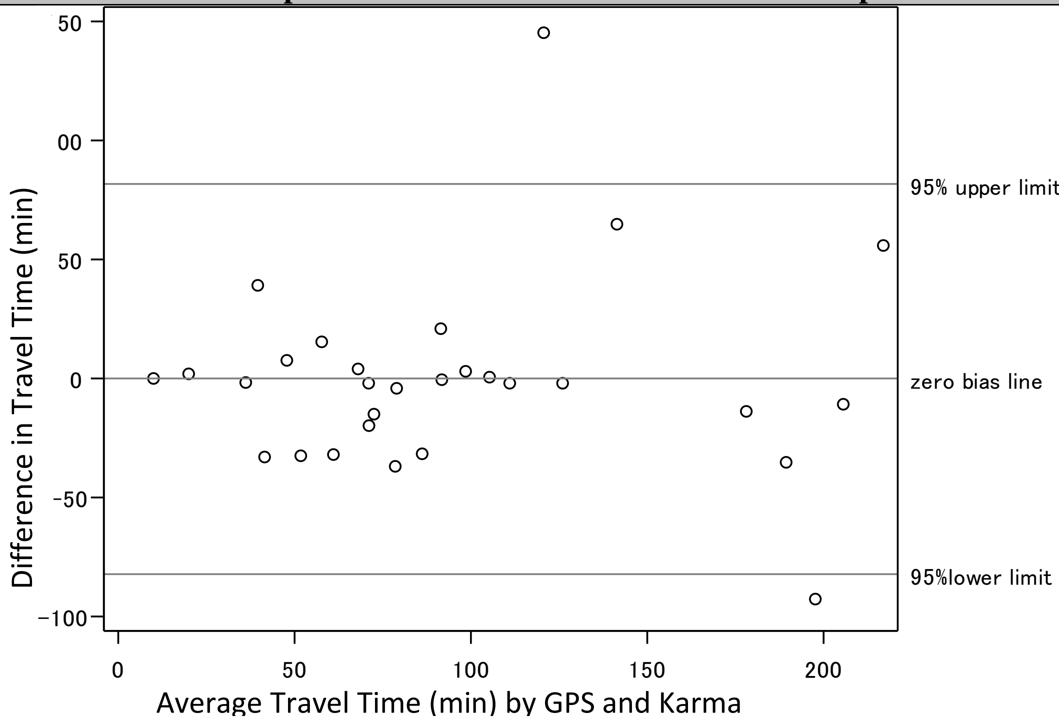


Fig 5(A): Bland-Altman plot between GPS and Karma for travel time (min)

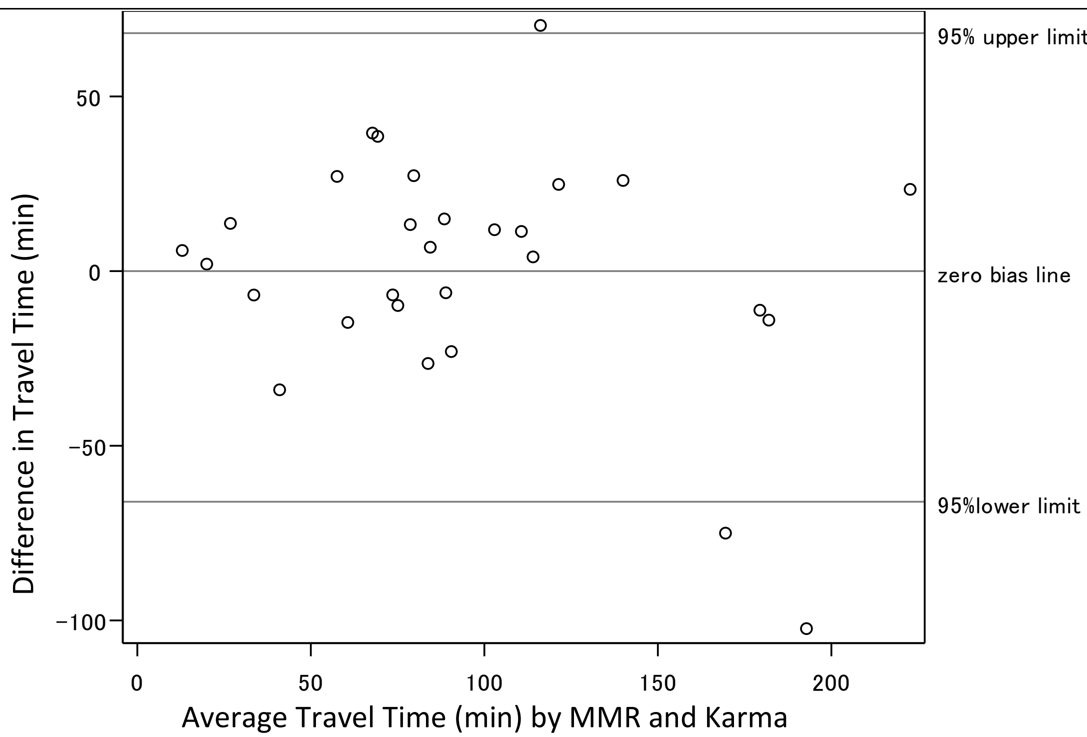


Fig 5(B): Bland-Altman plot between MMR and Karma for travel time (min)

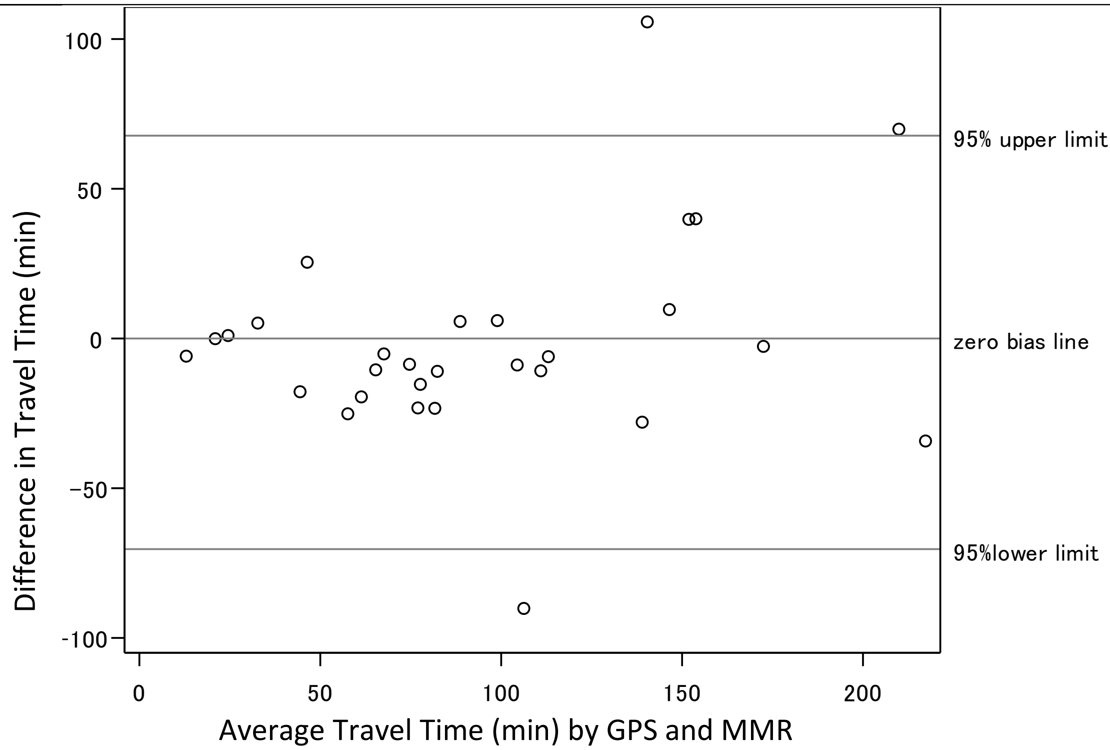


Fig 5(C): Bland-Altman plot between GPS and MMR for active travel time (min)

Figure 6: Bland-Altman plot for track length across instrument comparisons

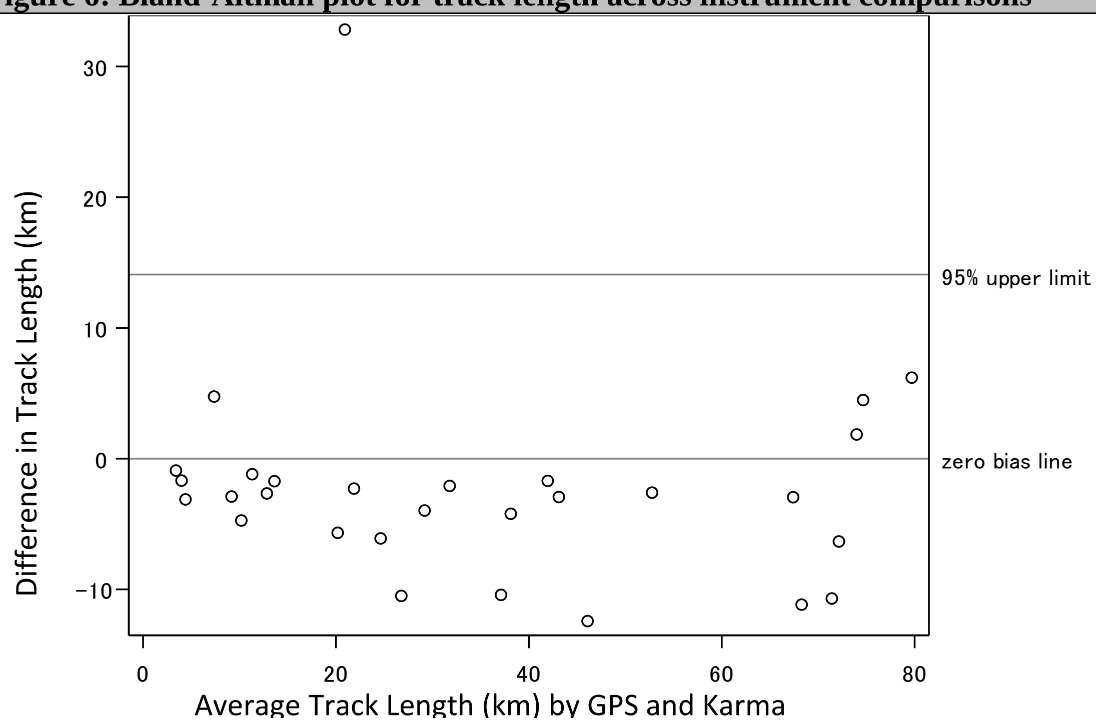


Fig 6(A): Bland-Altman plot between GPS and Karma for track length (km)

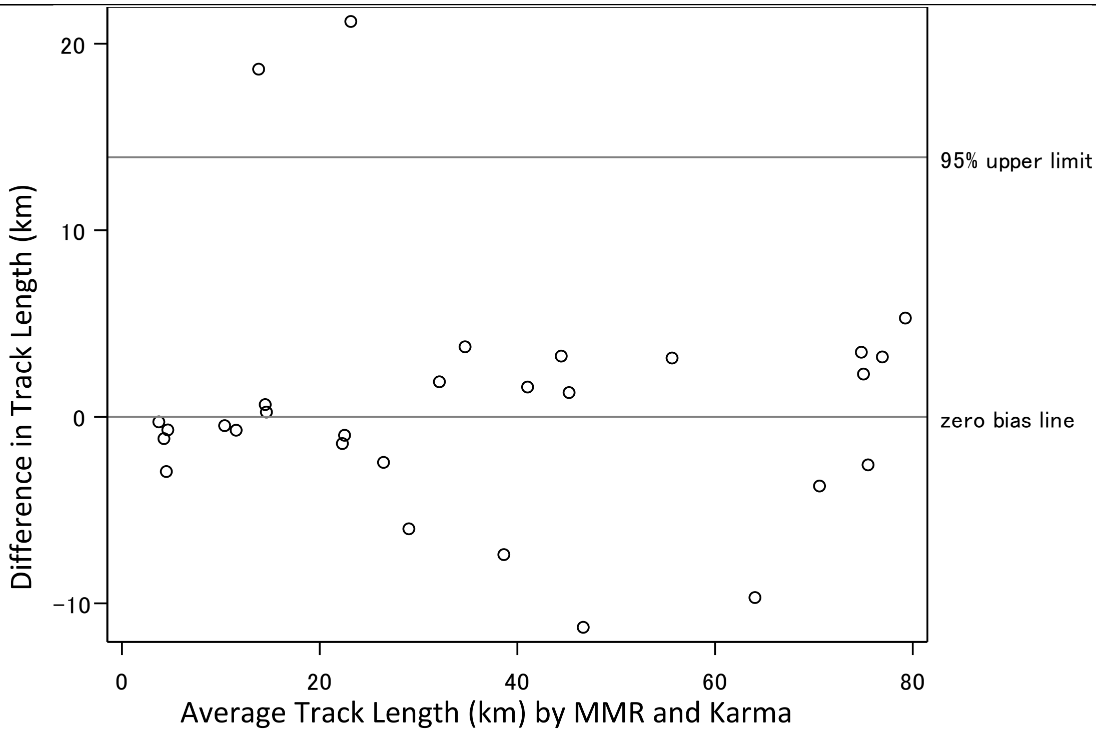


Fig 6(B): Bland-Altman plot between MMR and Karma for track length (km)

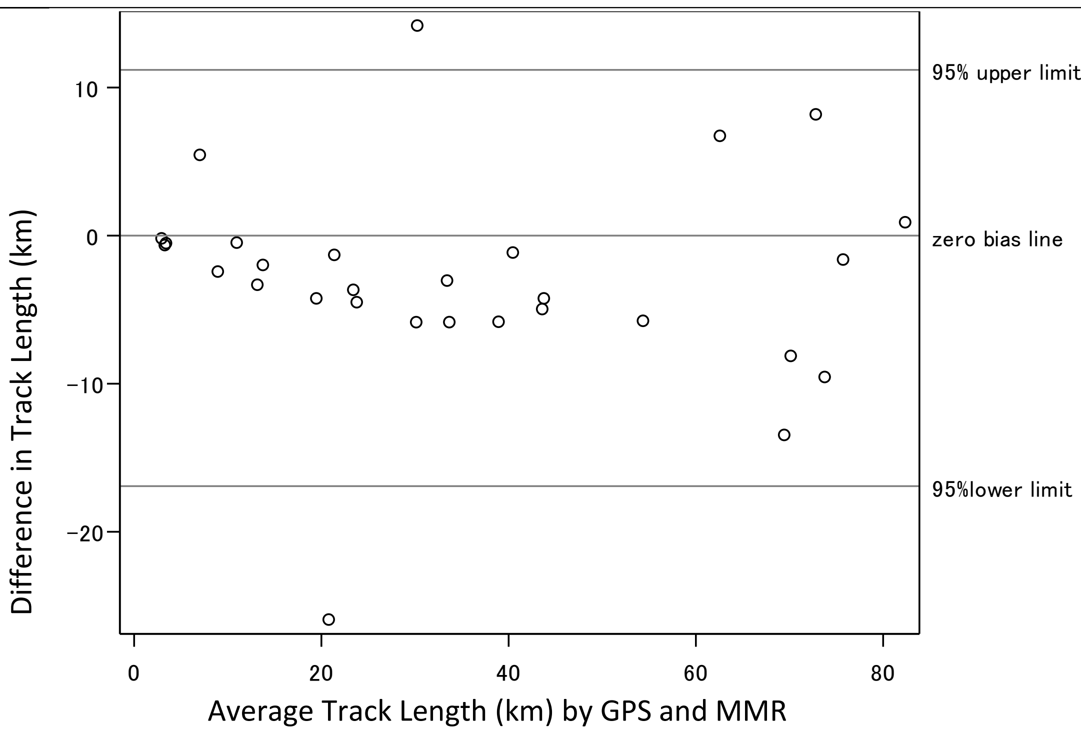


Fig 6(C): Bland-Altman plot between GPS and MMR for track length (km)

IV. Additional Analyses

A series of additional analyses were conducted to test the robustness of the findings.

A. Travel Mode Indicators as Obtained from Karma

Travel mode, defined as the method of transportation used by participants to reach destinations, could only be directly measured by Karma. This indicator can be only be derived and estimated from GPS and MMR using post-processing techniques.

For the present study, we focused on the travel mode variable available from Karma to understand its distribution in the sample **(Appendix 5)**.Of the 173 trips collected by Karma, 58 were walked (33%), 8 were biked (4.6%), 31 utilized public transportation (18%), and 76 involved driving to the respective destinations (44%). On average, participants walked for 9.3 (6.6) minutes with a median of 7.5 (8.8) minutes. Track length was on average 0.8 (0.5) km and 0.6 (0.8) km for walking. In contrast, the average time spent biking was 5.6 (1.3) minutes with a median of 5.0 (1.3) minutes. The average track length biked per participant-day was 1.2 (0.1) km with a median of 1.2 (0.2) km. Public transportation had the largest averages for both time and distance travelled. The average time travelled per participant day for public transportation was 42.9 (22.2) minutes with a median of 37.0 (20.5) minutes. Average track length per participant day was 14.5 (10.1) km with a median of 9.8 (16.7) km. The average time spent travelling via a car was 11.2 (6.5) minutes per participant-day with a median of 11.0 (10.0) minutes. Distance spent driving was on average 7.2 (6.7) km and 6.2 (8.9) km

B. Comparing Objective vs. Subjective Measure of Travel Time as Obtained from Karma

In addition to the objective travel time variable provided by Karma provided by Google Maps, Karma also provided self-reported (by participant) data on time spent commuting per trip per day. These data were compiled to generate a subjective total travel time variable that reflects total time spent commuting by a participant on a given day as self-reported by the participant. The Pearson's correlation coefficient between objective and subjective travel time variables was 0.85 (p <.0001), demonstrating significant correlation between the two estimations. A visual plot of the correlation between the two methods of estimation is presented in **Appendix 6.**

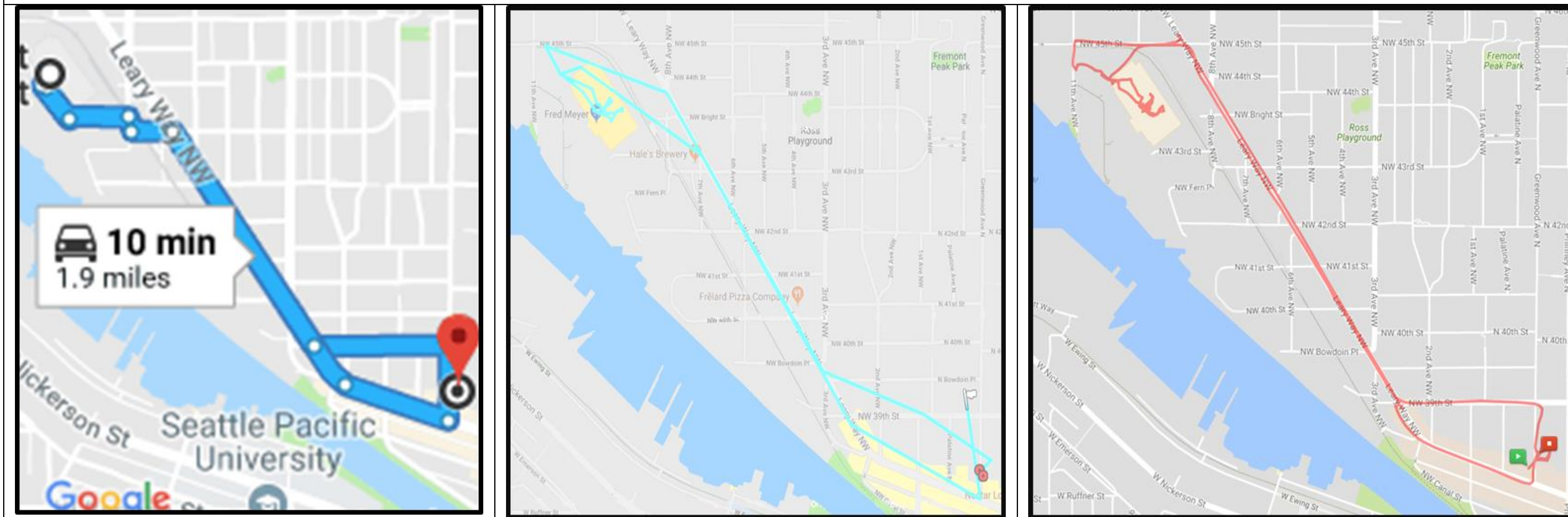
C. Correlation between Dwell Count and Other Primary Outcome Variables (dwell duration, travel time and track length)

Using the Pearson's correlation statistic, dwell count was compared to active dwell duration (in minutes), travel time (in minutes), and track length (in km) overall for each instrument. Dwell count and active dwell duration had a correlation coefficient of 0.29 for the entire dataset across instruments (0.49 for GPS, 0.26 for MMR, and 0.13 for Karma) **(Appendix 7A)**. Dwell count and travel time had an overall correlation of 0.55 (0.67 for GPS, 0.66 for MMR, and 0.34 for Karma) **(Appendix 7B)**. Dwell count and track length had an overall correlation of 0.39 (0.62 for GPS, 0.36 for MMR, and 0.19 for Karma) **(Appendix 7C)**. The overall correlation of the data shows minimal correlation between dwell counts and the other primary outcome variables.

V. Visual Plots from Karma, GPS and MMR

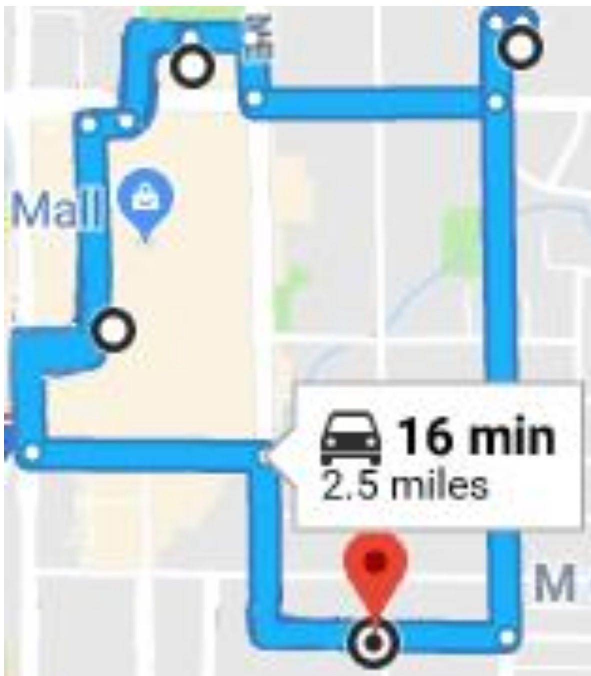
One of the unique features of Karma is that it automatically provides visual plots of the travel pattern for a given participant by each trip or compiled for a day, with minimal post processing time. Similar visuals can be obtained from GPS and MMR with post-processing**.** As seen in **Figure 7A,** which depicts a trip to a food store, both Karma and MMR more closely followed the path of the road when compared to GPS. While Karma stopped generating a path once the destination was reached, MMR and GPS continued trip coverage within the food store.

Figure 7A: Visual snapshot of a trip for a given participant as obtained from three devices: Karma, GPS and MMR

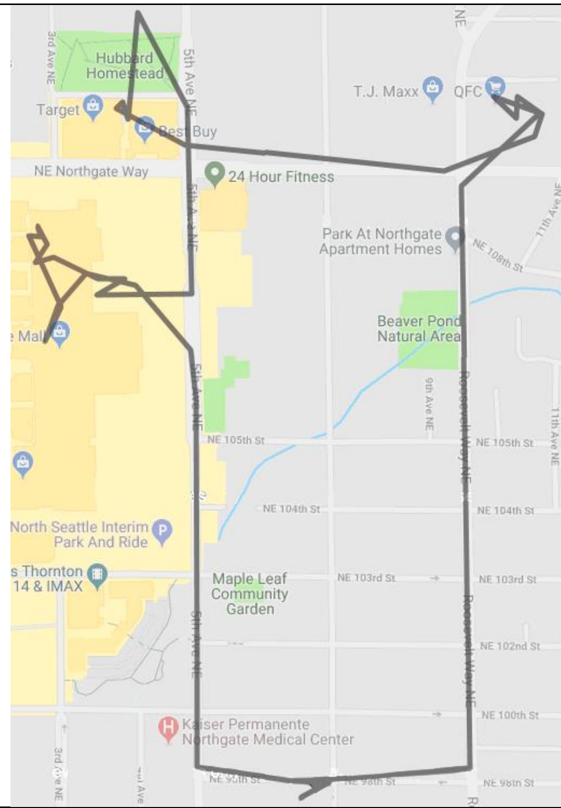


Karma	GPS	MMR

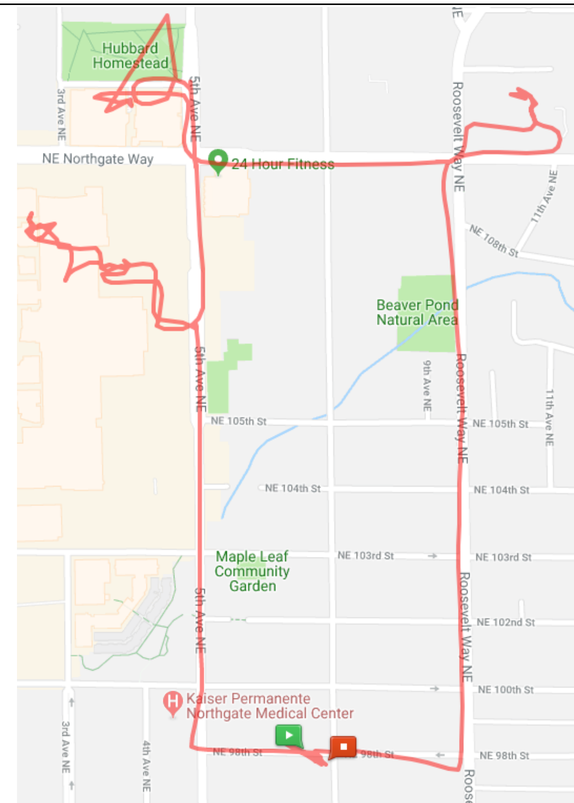
Figure 7B: A Visual snapshot of activity space for a given participant for an entire day as obtained from three devices: Karma, GPS and MMR



Karma



GPS



MMR

Discussion

I. Discussion of Results

The present study is one of the very few studies to study the use of Karma, a newly developed computer-assisted instrument to study activity space, in relation to traditionally used instruments in the field – the satellite-based GPS and MMR which uses cellular data and Wi-Fi to generate a track using a smartphone's internal assisted-GPS device.[33-34] In addition, this study is one of the first studies to compare the different activity space measures (dwell count, active dwell duration, travel time, and track length) across instruments. Most of the findings support Karma as an alternative valid instrument to both MMR and the traditional GPS instrument to capture habitual travel patterns and study activity space.

The present study computed four indicators of activity space from Karma that have been traditionally obtained from GPS and MMR data - number of dwell counts per participant-day, total dwell duration in minutes per participant-day, total time spent commuting in minutes per participant-day and track length in km per participant-day. The present sample made an average of five trips per day with an average of six dwell counts per day and a total dwell duration 9.5 hours per day (~570 min) during the active data collection period. The total time spent commuting was approximately 1.5 hours per day with a track length (total distance traveled per day) of 33 km. Of all the trips made over 29 person-days, the majority were made by car (44%) followed by walking (33%). Only 18% of the trips utilized public transportation and the remaining 4.6% were biked.

As observed through Pearson's correlation coefficients and Bland-Altman plots, there was a strong agreement between Karma, GPS and MMR for active dwell duration, travel time, and track length. Dwell count showed a relatively higher agreement across Karma and MMR as compared to GPS.

Active dwell duration showed a bias for lower values for Karma in comparison to both GPS and MMR. This is likely due to the nature of the data collection. Since active dwell time was calculated from participants' estimations, they could have consistently underestimated the time spent in their dwell locations, which would have led to the bias towards lower active dwell duration. Agreement for travel time and track length between instruments tended to decrease as both travel time and track length increased. Considering that traffic is a large factor in travel time (and sometimes track length as participants shift routes to other routes), it is possible that the increasing loss of agreement was due to traffic during peak commuting hours, which was unaccounted for in Karma. Similarly, traffic in a large city like Seattle could have led participants to be relatively stationary in "urban canyons" for extended periods of time, leading to further signal loss from either instrument.[14]

Despite the supportive results found in the other three variables, dwell count correlation for Karma vs GPS was less than 0.5, indicating a weak correlation between the datasets. Why this may be the case remains unclear, especially considering the strong correlation between the other variables that are based upon the estimated dwell point. Likely there were many factors that led to the inconsistency in dwell count. A small sample size likely led to minor differences of one or two dwell points between instruments having a significant effect on the final correlation. There were also two samples taken that provided extremely different dwell counts across instruments. One of these was due to the higher sensitivity of Karma in documenting trips that were less than five minutes long. In addition to potential extremes, aggregating the data into five-minute intervals could have hidden other brief dwell points or trips that were less than five minutes. Similarly, the rate of data collection (one minute per data point for GPS and one second per data point for MMR) and difference in signals (satellite and cellular respectively) could have led to data loss.[9,12] The different methods of data collection between devices could potentially explain the higher correlation between MMR and Karma when compared against GPS.

In addition to potential errors in the data collection and preparation, it is possible that the method of establishing dwell point was inadequate at identifying dwell points consistently between instruments. Using 4.5 km/h as a speed threshold for dwell points could have led to a loss (or gain) in dwell count as the data was analyzed. Similarly, the requirement that the low speed be maintained for at least five minutes may have created similar issues of sensitivity or specificity. Since most participants reported travelling for less than ten trips per day as per Karma, even the difference of one dwell point that was lost could lead to a significant difference. Signal loss for both GPS and MMR could have led to the loss of dwell points, and signal drift could have established new ones as the speed threshold was "exceeded" due to the apparent movement of the instrument.

In addition to comparisons for Karma against GPS and MMR, data collected during the study allowed for further analyses of food shopping vs. non-food shopping days and weekdays vs. weekends was conducted. One hypothesis for the secondary aim was whether the number of dwell counts, total time spent commuting and total track length was higher for food shopping days and weekends. Contrary to the hypothesis, non-food shopping days had significantly larger active dwell durations, travel time, and track length values for both mean and median. A similar trend was observed for weekdays vs. weekends, with weekdays having much higher values. Due to the small sample size, no further statistical testing could be conducted on the stratified data.

There could be a myriad of reasons as to why there is such a drastic difference between the respective pairs of days. For most of the participants, weekdays were days during which they attended work or school, which could very likely increase active dwell duration. Travel time and track length could have increased as participants faced heavy traffic to and from work. Similarly, non-food shopping days could have frequently fallen on weekdays during which participants were more likely to be at work, increasing active dwell duration and potentially travel time and track length. Or, alternately, participants were deterred from food shopping *due to* the greater amount of time spent actively dwelling and travelling already present in their day. The small sample size due to stratification may have affected the values across the various days as well. What could have been an anomalously brief shopping day would have played a large role in skewing the small sample to a lower average and median value that may not have necessarily reflected typical behavior patterns.

Comparisons across instruments revealed that for both food shopping and non-food shopping days, Karma tended to generate lower values for active dwell duration and travel time in comparison to GPS and MMR. This is likely to due the nature of Karma's data collection. Since Karma utilizes Google Maps, it generates the "shortest" path based off of the time at which the trip data is gathered (which may not be the same time as the trip occurred). “Shortest” is not based on distance, however, but instead *time*. Because of this, Karma would be expected to provide lower values on average for travel time, while track length may not necessarily follow the same pathway.

In addition to the primary and secondary aims being tested, several other statistical analyses were conducted to review additional outcome variables. Since mode of transportation was uniquely provided by Karma, the data generated could only be tested within itself. After conducting descriptive statistical measurements, it became clear that while walking and driving were the most frequent means of transportation amongst participants, on average the greatest time and distance spent per participant-day was through public transportation. When thinking about city planning and transportation, this information follows. Buses and trains are not intended to take the most expedient route, in fact it is quite the opposite. In order to function in an efficient and effective manner, buses must stop frequently to increase the number of passengers. These stops increase both time and distance as arterioles are taken in lieu of thoroughfares.

Other analyses were used to test the validity of the data itself, rather than provide new information. A correlation test between subjective participant estimation and objective Karma-generated estimation for travel time was used. While the travel time estimated by participants was ultimately not used in the larger study, this secondary analysis allowed researchers to observe the potential shortcomings in participant recall against a more standardized means of observation (Google Maps). In conducting this analysis, it became clear that participants were fairly reliable narrators of their own day.

The final two analyses were intended to test and validate the dwell count values across instruments, and in doing so potentially provide insight into the dramatic swings that were seen for the correlation across variables. A chi-square goodness-of-fit test was used to compare the data in a non-parametric fashion. This test showed non-significant difference for dwell counts between instruments. This supports the theory that the low correlation seen in dwell count may be due to the nature of the data and the inappropriate use of a parametric measurement (Pearson’s correlation) rather than a flaw in the data collection or processing.

To further study the relationship between dwell count and the other variables, a series of Pearson’s correlations were calculated across instruments. Theoretically, as dwell count increased so would the value for active dwell duration. Conversely, travel time (and potentially track length) would decrease since they were essentially the inverse values of dwell times. After generating the various correlations, these trends were not observed. In fact, there was relatively low correlation across instruments and variables and no negative correlations at all. It is likely that there is no relationship between dwell count and the other variables. As an example to explain this finding, a person may make 10 stops in a day, but if those stops are only five minutes each, the final value will be lower than the person who stopped three times for three hours each. Dwell count is independent of duration and distance. These findings support further research regarding the relationship between dwell count, activity space measures, and behavior patterns.

II. Strengths and Limitations of Instruments

The present study revealed several strengths and limitations for each instrument to capture, analyze and utilize data on activity space. **Table 4** below compares the traditional used instruments (GPS and MMR) with the newly developed Karma on multiple parameters.

While GPS is treated as a standard for measuring the geospatial environment, it is not without shortcomings.[33] As mentioned previously, it is prone to signal loss and drifting, which can affect the accuracy of measurements.[13, 14] While this can be corrected for in processing, it requires software and algorithms to ensure that it is done successfully.[11] There is also an increased risk for participants losing the instrument, which can be costly for researchers in larger studies.[12] As evidenced by this study, there is also a risk for participants failing to follow protocol, leading to data loss as instruments remain off or at home during data collection periods. In addition to the inherent limitations of the instrument, GPS instruments are also unable to identify trip purpose and mode of transportation.[11] While there have been several studies that have tried to engineer techniques that allow for accurate estimations of these variables, they have had mixed success.[11]

MMR falls into similar pitfalls as the traditional satellite GPS instrument. Travel mode and trip purpose must either be inferred through software or algorithms or explicitly outlined by the participant.[10] Unlike a traditional GPS, however, there is no cost to create MMR accounts and use the application.[18] Participants can uninstall the application after data collection, and the data will still exist within the MMR servers associated with the corresponding account until it is deactivated by researchers. That being said, despite its utility and ease of use, there are still shortcomings in the application that can affect data collection. As it records data in one-second intervals, there is significant strain on both battery life and data usage for the smartphone. Throughout the study, this was a frequent complaint of participants. MMR also suffered from signal loss akin to that of the traditional GPS instrument, likely due to the nature of the assisted-GPS chip used in the smartphone.[33] Failure to follow study protocol for MMR was also the primary cause of data loss of the three instruments, which may contraindicate the use of the software for future studies to measure activity space. MMR, much like GPS, required extensive data cleaning and processing before it could be used to estimate activity space, making it more challenging to use in larger studies.

In contrast to both MMR and GPS, Karma required very little processing and minimal cleaning to provide the information for outcome variables. Inherent in its design, Karma provided mode of transportation as reported by participants. It also automatically provided self-reported and Google Maps-generated travel times and destinations (dwell points) for each trip. Active dwell duration could be easily calculated from participant estimations, and track length was provided by Google Maps.

Shortcomings existed for the Karma instrument as well. As an instrument that is exclusively reliant upon participant memory, it requires that participants be accurate in their reporting. Since the track is automatically generated from Google Maps at the moment that it is being recorded, the provided travel time and suggested track may not necessarily reflect the path a participant took previously. As it stands, the instrument is also limited in the number of trips that can be recorded, which makes trips that utilize multiple methods of transportation (such as driving to a transit center to take a bus) challenging if at all possible.

Table 4: ***Strengths and limitations across study instruments***			
Features	**GPS**	**MMR**	**Karma**
Data Collection	• Objective • Uses satellite data • Validated	• Objective • Uses cellular and Wi-Fi data • Not validated • Can potentially change functionality with updates or be removed from public use • Uses smartphones, which are relatively ubiquitous	• Subjective • No data usage – computer-assisted • Not validated
Type of Data Collection	• Participant wears a GPS instrument that auto-records their travel route • A QStarz GPS instrument is initialized and given to a participant to wear • Instrument needs to be charged by the participant each night	• Participants use a free smartphone application that can auto-record travel route • Participants are provided a username and password that they can use to upload their trips on the MapMyRun servers • Heavy data and battery usage for cellphone	• Web-based, interviewer-assisted program • Conducted during scheduled times over the phone • Program completed in a short amount of time (5-20 minutes), depending on the number of trips taken
Data Input	• Electronic, automated recording of travel patterns collected at pre-set, regular inputs • Potential for data loss due to non-compliance, signal loss, or signal drifting	• Electronic, automated recording of travel patterns collected at pre-set, one second intervals • Participants able to review their own paths during the study through their accounts • Potential for data loss due to non-compliance and signal loss	• Participant-reported address and trip details are entered into the program by interviewer • Participants are held accountable due to scheduled phone call
Real-time Feedback from Participants	• No	• No	• Yes • An interviewer administers the Karma instrument with the participant, who can provide real-time feedback on the

			accuracy of the mapped locations and routes taken.
Participant Burden	• High	• High	• Low
Data Output *Interim Variables*	• Date & time, latitude & longitude, and distance traveled for each record (auto-provided from the GPS data) • Time difference, speed, and estimated dwell points were calculated using the auto-provided values • Interim variables were required for estimation and calculation of primary outcome variables	• Date, time, latitude & longitude, and distance for each record (auto-provided from the MMR data) • Time difference, speed, and estimated dwell points were calculated using the auto-provided values • Interim variables were required for estimation and calculation of primary outcome variables	• Interim variables not required for calculation of primary outcome variables
Quality of the Data Output Generated	• Manually developed using provided data and interim variables • Quality of data variable by participant • Signal loss, drift, and noise affected dwell point estimation • Dwell points only estimated as GPS is "blind" to when dwells begin and end • Participant non-compliance led to extra "inactive" dwell time at end of day	• Manually developed using provided data and interim variables • Quality of data variable by participant • Signal loss, drift, and noise affected dwell point estimation • Dwell points only estimated as MMR is "blind" to when dwells begin and end • Largest loss of data due to non-compliance across instruments • Participant non-compliance led to extra "inactive" dwell time at end of day	• Auto-computed using self-reported data • Quality of data output consistent • Unaffected by signal loss and noise
Post-Processing Time	• Very High • Data need to be downloaded from each instrument every time, and to be manually checked for completeness • A large amount of data on latitude and longitude every minute by GPS was gathered, which was converted	• Very High • Data is manually downloaded from each account for each day • Data provided in proprietary file format that must be converted using an algorithm	• Low • The data are automatically entered when typed into the web-based program • All primary outcome variables were automatically provided with minimal processing

	to five-minutes using an in-house algorithm • Dwell point was manually established, and then primary outcome variables were calculated from dwell points	• An extremely large amount of data on latitude and longitude every second by MMR was gathered, which was converted to five-minutes using an in-house algorithm • Dwell point was manually established, and then primary outcome variables were calculated from dwell points	
Data Quality	• Moderate • Data collected objectively, which minimizes biases • Data prone to drift and signal loss	• High • Data collected objectively, which minimizes biases • Data prone to signal loss	• High • The data are self-reported, so they are subject to recall and social desirability biases. • Strong correlation observed between participant estimations and Karma-generated travel times
Data Completeness	• Partial • Only provides objective data on travel routes or traces. • No information on actual destinations visited, at what time and in what order. • No information provided on method of transportation or trip purpose	• Partial • Only provides objective data on travel routes or traces. • No information on actual destinations visited, at what time and in what order. • No information provided on method of transportation or trip purpose	• More complete • Provides frequency of visits to key locations and potential demographic information • Provides self-reported information on location and trip-level data (start/end time, routes, method of transportation)
Potential Biases	• Subject to social desirability bias if someone decides to switch off the instrument, as well as loss of data due to a lost signal or participant noncompliance (i.e. non-wearing according to protocols).	• Subject to social desirability bias if someone decides to switch off the instrument, as well as loss of data due to a lost signal or participant noncompliance (i.e. not saving or using the application according to protocols)	• Subject to recall and social desirability biases
Economic Cost	• High	• Low	• Low

Total Time cost (Including Data Collection and Data Processing)	• Very high • Can be cumbersome and time consuming for the participant to wear the instrument, with the potential for low data quality due to lack of signal and noncompliance. • Involves huge post processing time to download and process the data.	• Very high • Participants must be willing to use data and battery from their own cellphones for data collection. Participants also must remember and log in to MMR using unique email addresses and passwords generated for the study • Involves huge post processing time to download and process the data.	• Lowest • Involves a quick session with the trained interviewer (<15 minutes) to collect all the data. • Zero post processing time and burden because all the analytical variables are auto-computed in real time.

III. Limitations of the Study

One of the main limitations of the study was the small sample size. Second, dwell counts were determined from GPS and MMR using a newly developed algorithm using <4.5 km/h for five minutes as the cut off to identify a dwell point. This threshold was based on review of the existing literature. Since all other variables were contingent upon the dwell count estimation, this could have led to further errors in the dataset. Additional sensitivity analyses should be conducted with other cut-offs in future studies.

IV. Conclusion

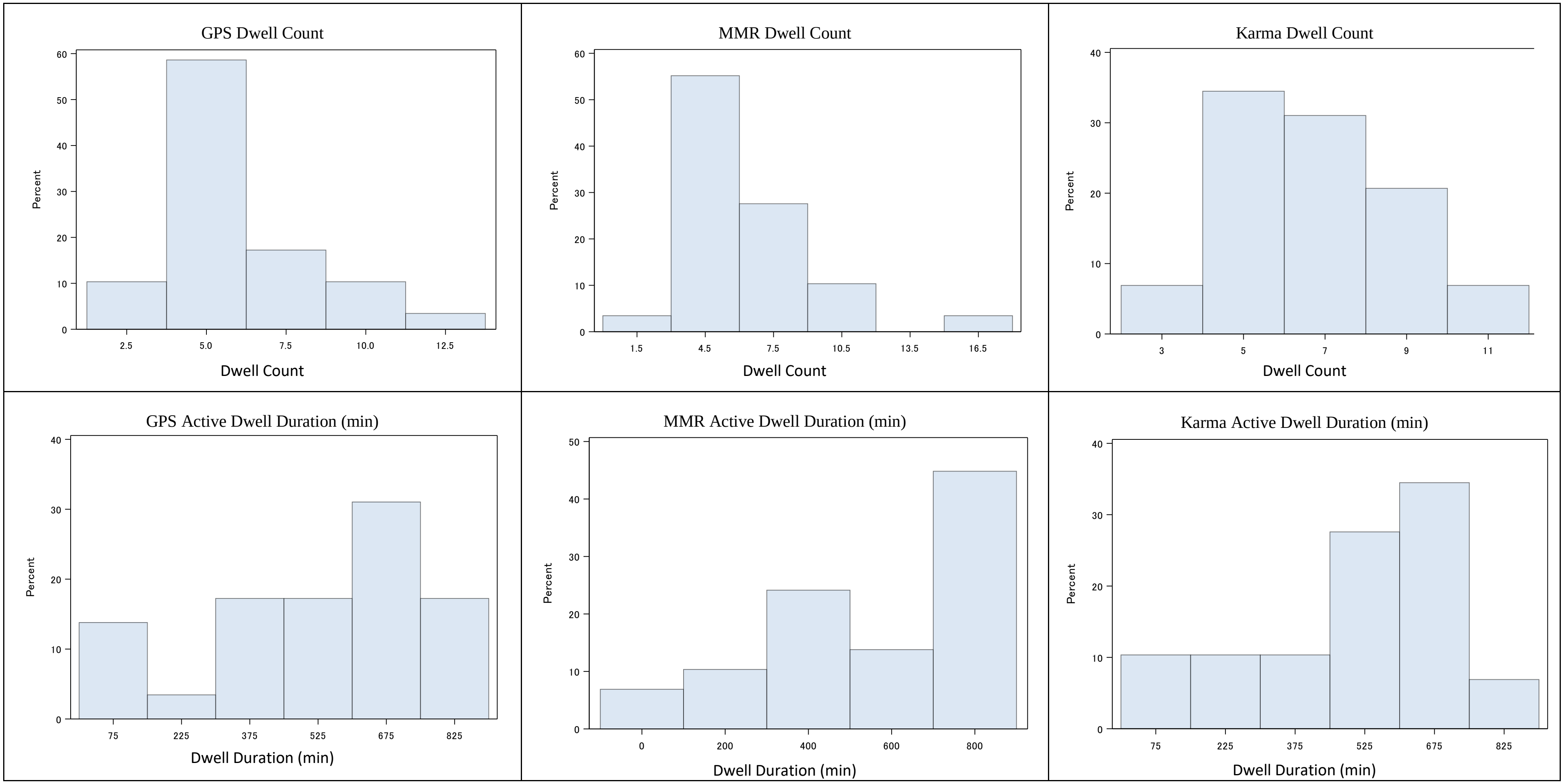
While GPS and travel logs remain as standards for measuring geospatial activity space, instruments akin to Karma will continue to offer researchers other avenues to pursue data collection.[35] As activity space continues to be studied on larger scales, it becomes essential to develop instruments that are free, available on ubiquitous platforms like smartphones and computers, and provide datasets that require minimal processing Karma is one such instrument that can be utilized to collect and study data on activity space with least respondent burden and minimal data processing time.

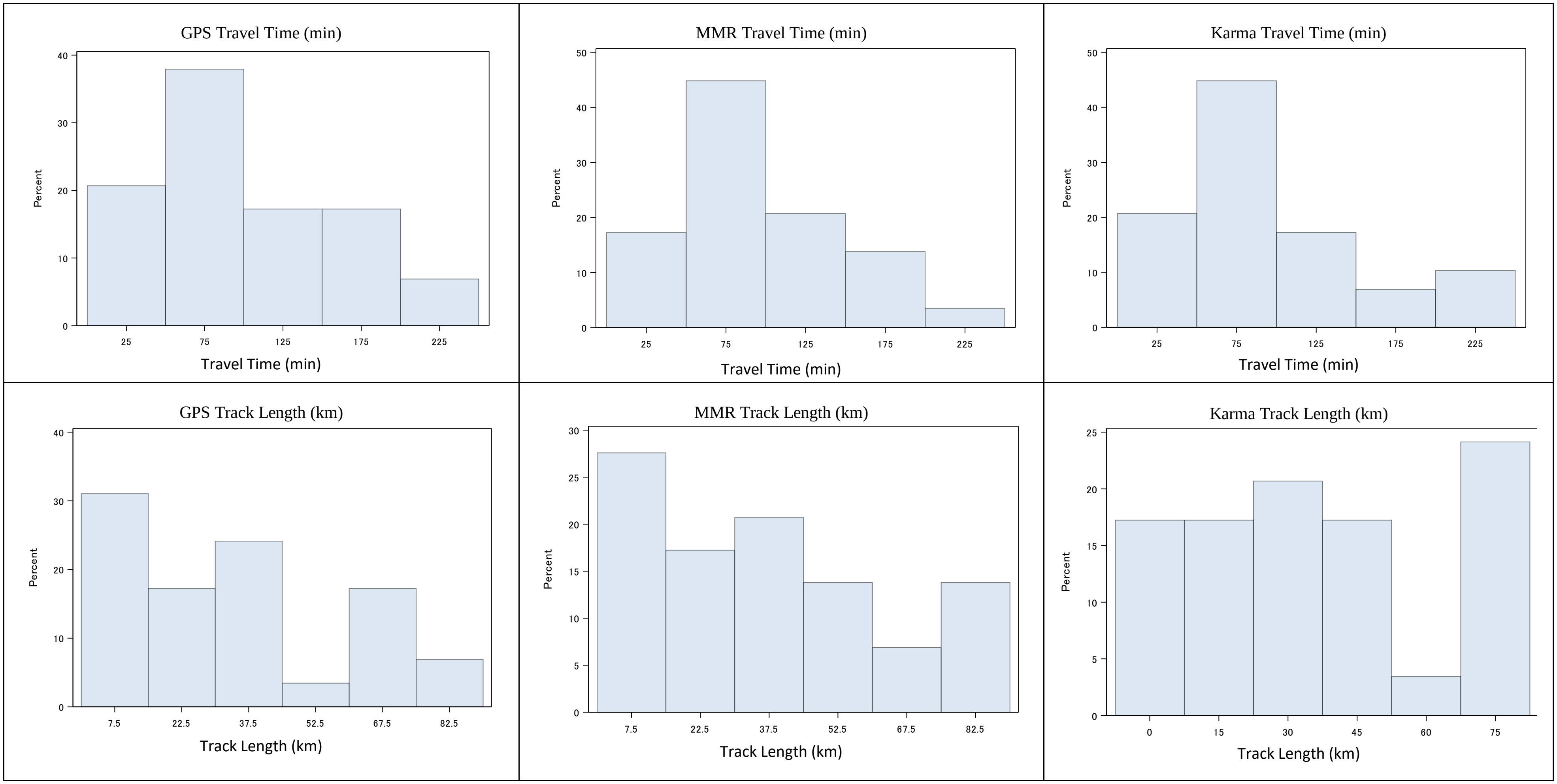
Acknowledgements

I would like to express my sincere thanks to everyone who helped me during this thesis work. Dr. Shilpi Gupta provided me training on how to handle study instruments from data collection to data downloading. Dr. Chelsea Rose assisted with statistical data analyses and data interpretation. The algorithm used to process the data into five-minute intervals was created by Ishin Iwasaki.

Appendices

Appendix 1 Histogram of sample distribution of primary outcome variables per instrument




GPS Travel Time (min)
Percent
Travel Time (min)
MMR Travel Time (min)
Percent
Travel Time (min)
Karma Travel Time (min)
Percent
Travel Time (min)
GPS Track Length (km)
Percent
Track Length (km)
MMR Track Length (km)
Percent
Track Length (km)
Karma Track Length (km)
Percent
Track Length (km)

Appendix 2 Distributions of primary outcome variables for food shopping (n=7) versus non-food shopping days (n=22) as computed from three instruments: GPS, MMR and Karma

	Dwell Count		***Active Dwell Duration (min)***		***Travel Time (minutes)***		***Track Length (km)***	
	Mean (SD)	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**
Food Shopping Day								
GPS	4.7 (1.4)	5.0 (2.0)	429.1 (276.1)	505.3 (554.5)	48.9 (38.5)	35.5 (79.1)	13.1 (14.3)	7.7 (18.6)
MMR	4.7 (1.4)	5.0 (2.0)	428.2 (278.8)	483.5 (614.9)	58.5 (41.7)	53.3 (87.9)	15.2 (15.7)	10.1 (21.7)
Karma	6.3 (2.7)	6.0 (4.0)	413.6 (251.7)	485.0 (445.0)	62.0 (36.3)	68.0 (78.0)	16.0 (15.0)	10.6 (22.9)
Non-Food Shopping Day								
GPS	6.3 (2.4)	6.0 (4.0)	559.3 (234.0)	701.3 (410.8)	110.0 (59.2)	84.2 (106.2)	40.7 (24.3)	36.7 (45.2)
MMR	6.5 (3.1)	5.0 (3.0)	575.6 (221.3)	709.4 (331.7)	108.6 (49.0)	90.9 (62.7)	43.8 (24.2)	38.8 (45.6)
Karma	5.9 (2.2)	6.0 (3.0)	550.0 (216.5)	637.5 (350.0)	106.1 (62.5)	86.5 (61.0)	42.7 (25.6)	41.3 (49.4)

Appendix 3 Distributions of primary outcome variables for weekday (n=17) versus weekend (n=12) as computed from three instruments: GPS, MMR and Karma

	Dwell Count		***Active Dwell Duration (min)***		***Travel Time (minutes)***		***Track Length (km)***	
	Mean (SD)	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**
Weekday								
GPS	6.0 (2.5)	6.7 (3.0)	726.9 (179.7)	632.4 (245.8)	110.0 (62.7)	121.9 (101.8)	37.3 (25.2)	42.7 (44.4)
MMR	5.0 (2.3)	6.1 (3.0)	727.3 (165.5)	634.5 (245.9)	108.9 (53.4)	112.4 (62.7)	41.8 (26.1)	44.0 (45.6)
Karma	6.0 (2.4)	5.9 (3.0)	700.0 (178.9)	606.5 (235.0)	102.0 (65.6)	119.2 (108.0)	42.3 (26.4)	43.4 (49.4)
Weekend Day								
GPS	4.8 (1.4)	5.0 (2.0)	400.4 (257.9)	379.8 (469.7)	63.1 (30.3)	57.4 (38.8)	15.1 (19.9)	21.7 (26.5)
MMR	5.0 (3.6)	6.0 (1.0)	453.3 (270.5)	406.1 (552.8)	79.5 (40.6)	74.1 (48.0)	27.3 (21.5)	26.8 (29.2)
Karma	6.0 (2.3)	6.1 (3.3)	485.0 (237.8)	390.4 (398.8)	67.0 (27.5)	61.9 (32.5)	18.8 (22.5)	26.1 (26.3)

Appendix 4 Chi-square goodness-of-fit test p-value of dwell count between instruments

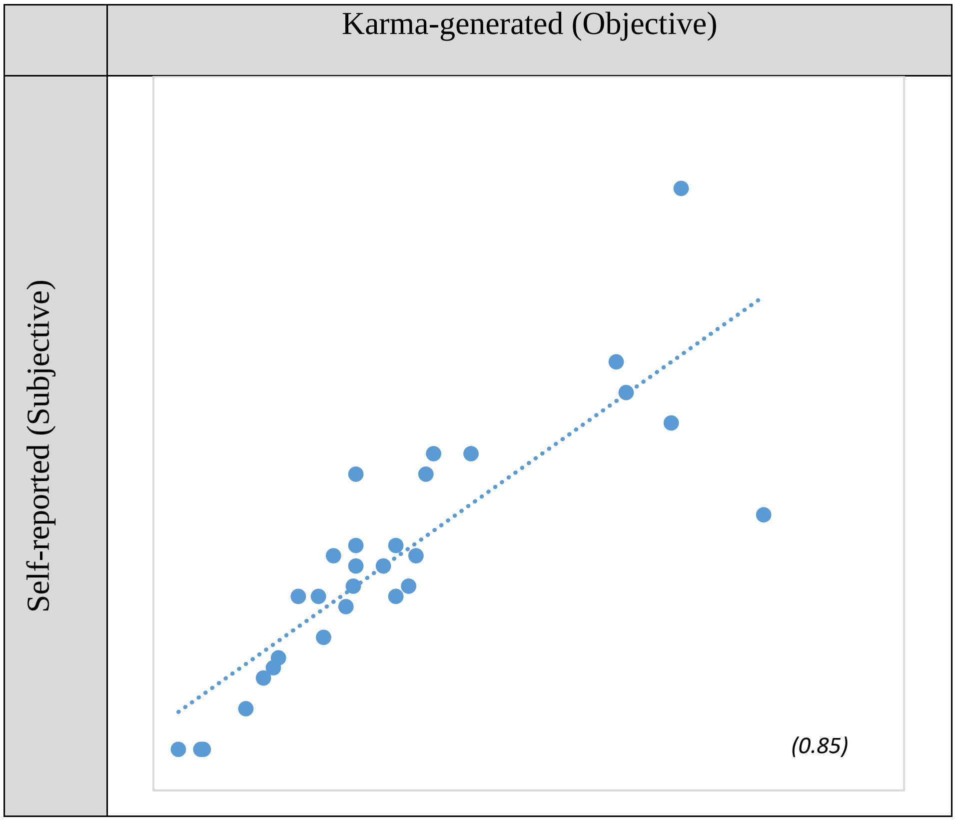
Chi-Square p-value

	GPS	**MMR**	**Karma**
GPS	1.0	0.18	0.32
MMR	0.18	1.0	0.91
Karma	0.32	0.91	1.0

Appendix 5 Travel time and track length by self-reported travel mode in Karma

	Travel Time (min)		***Track Length (km)***	
	Mean (SD)	**Median (IQR)**	**Mean (SD)**	**Median (IQR)**
Walking (n=58)	9.3 (6.6)	7.5 (8.8)	0.8 (0.5)	0.6 (0.8)
Biking (n=8)	5.6 (1.3)	5.0 (1.3)	1.2 (0.1)	1.2 (0.2)
Public Transportation (n=31)	42.9 (22.2)	37.0 (20.5)	14.5 (10.1)	9.8 (16.7)
Driving (n=76)	11.2 (6.5)	11.0 (10.0)	7.2 (6.7)	6.2 (8.9)

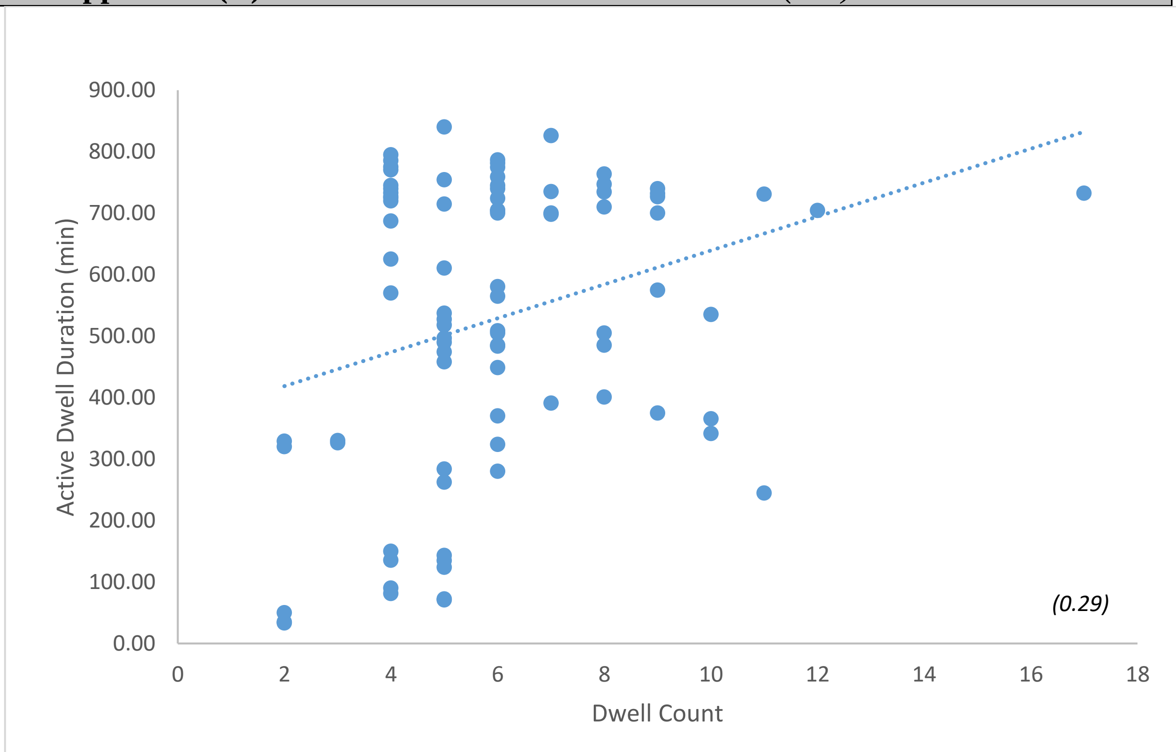
Appendix 6 Plot for Karma-generated and participant reported travel times in minutes (Pearson's Coefficient)



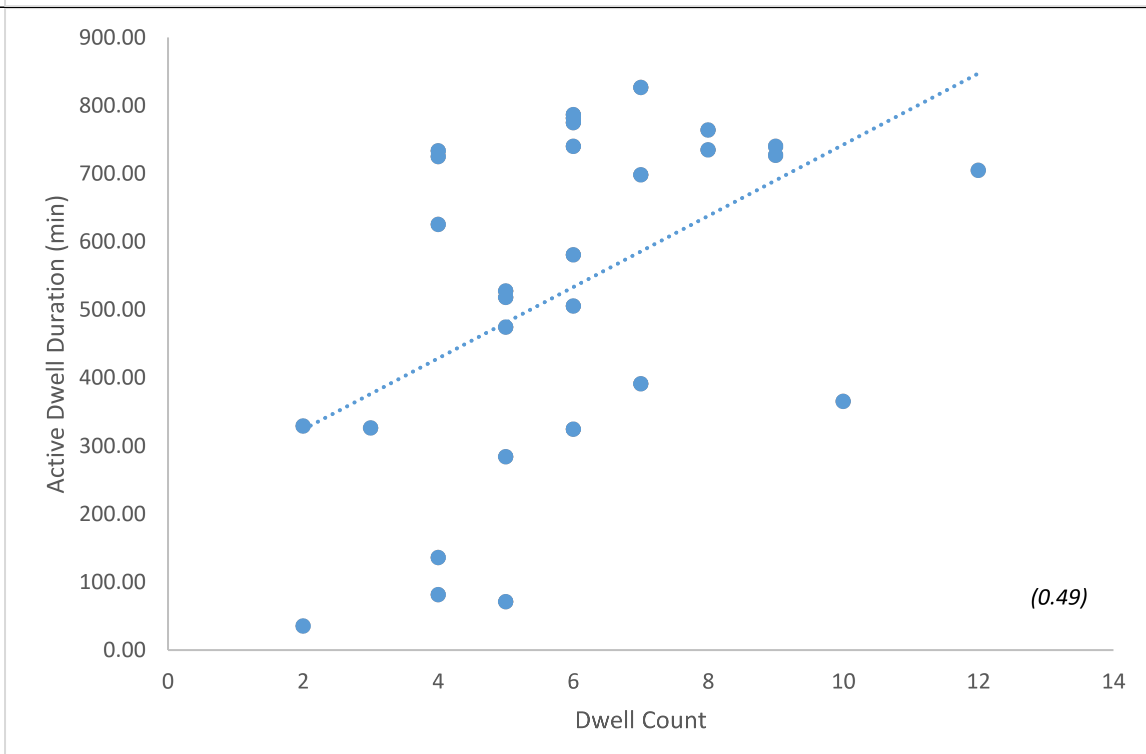
Appendix 7 Regression plots of dwell counts in comparison to other primary outcome variables across instruments (Pearson's coefficient)

Appendix 7(A): *Dwell Count vs Active Dwell Duration (min) across Instruments*

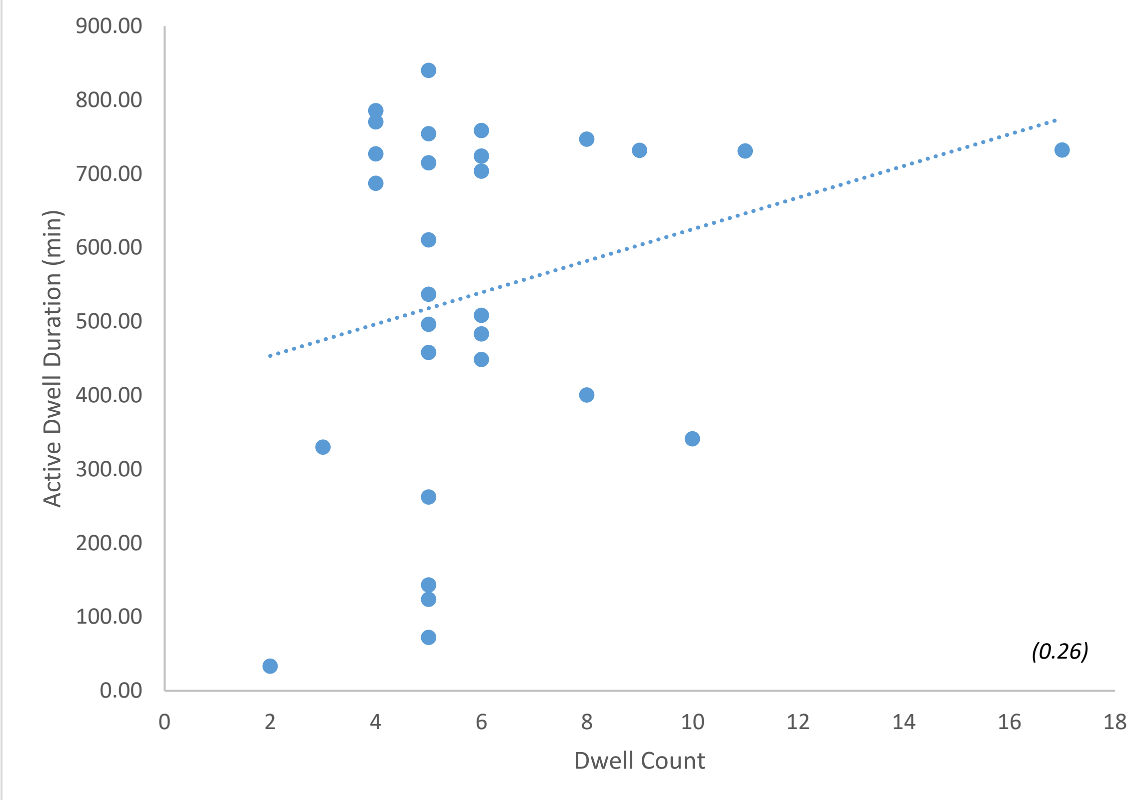
All Instruments



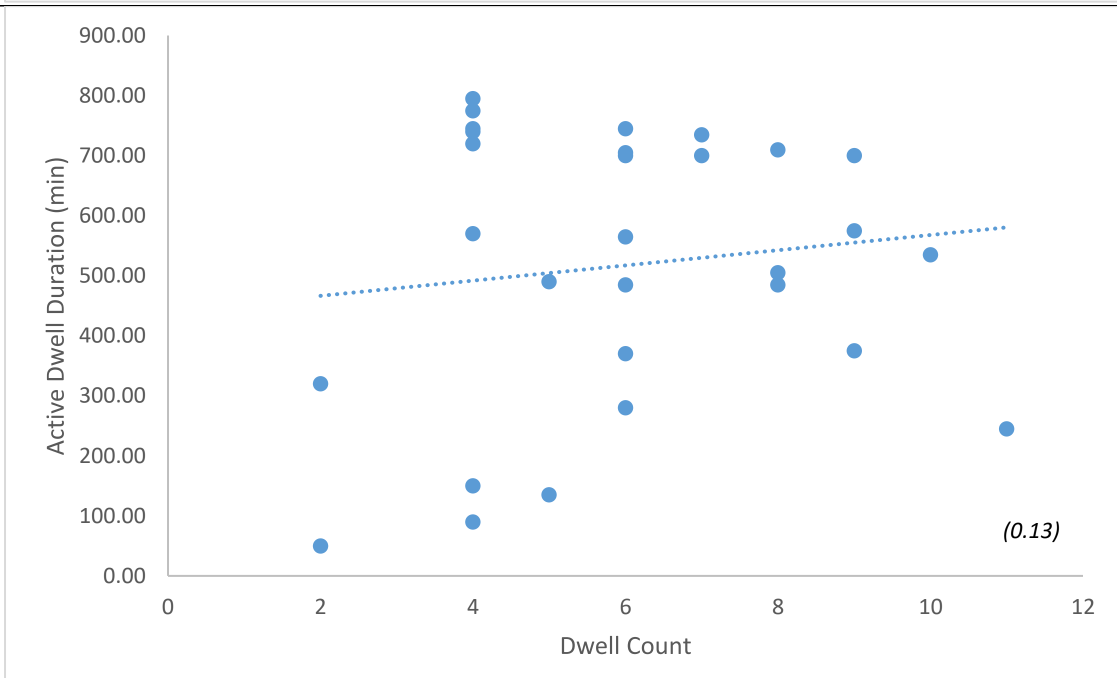
GPS

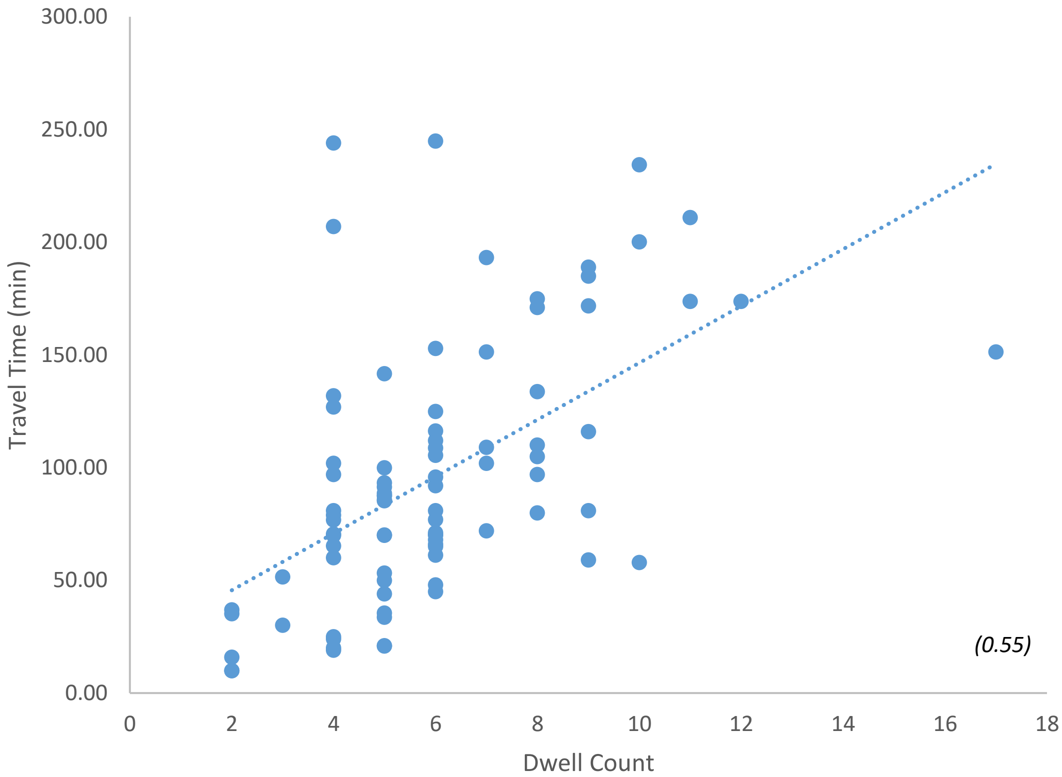
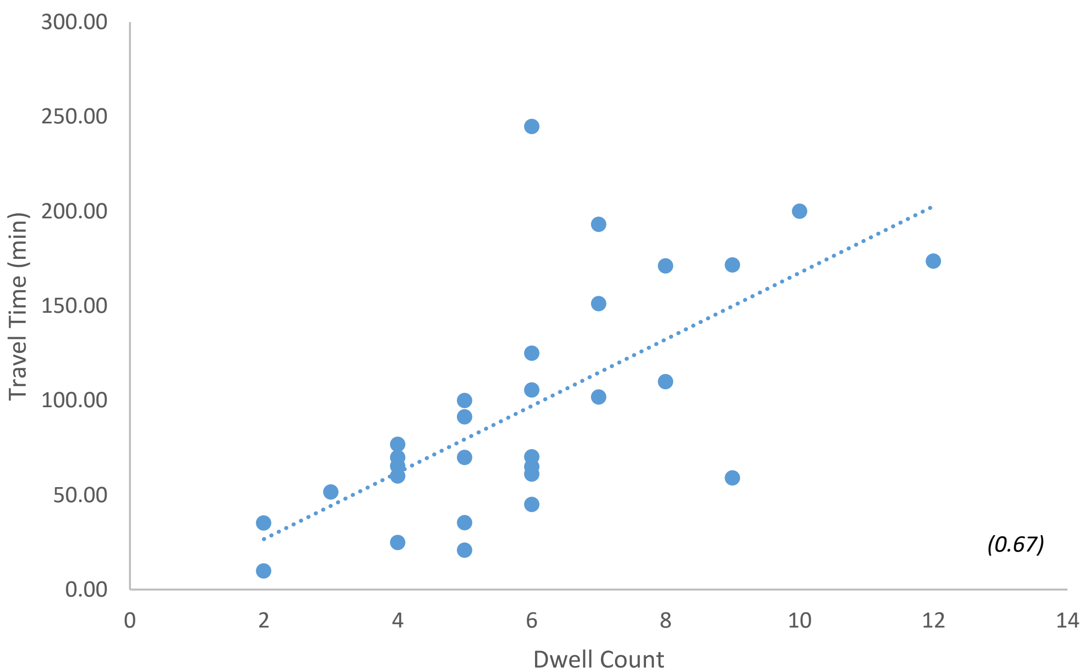


MMR

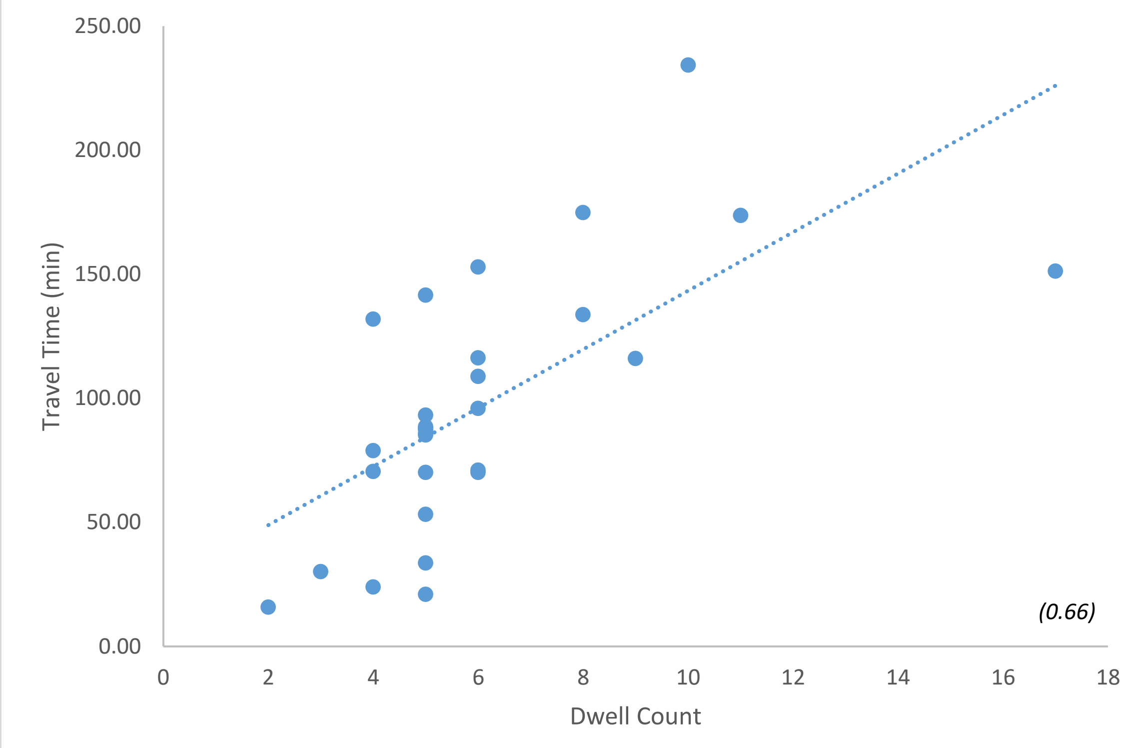

Active Dwell Duration (min)
900.00
800.00
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
0
2
4
6
8
10
12
14
16
18
Dwell Count
(0.26)

Karma

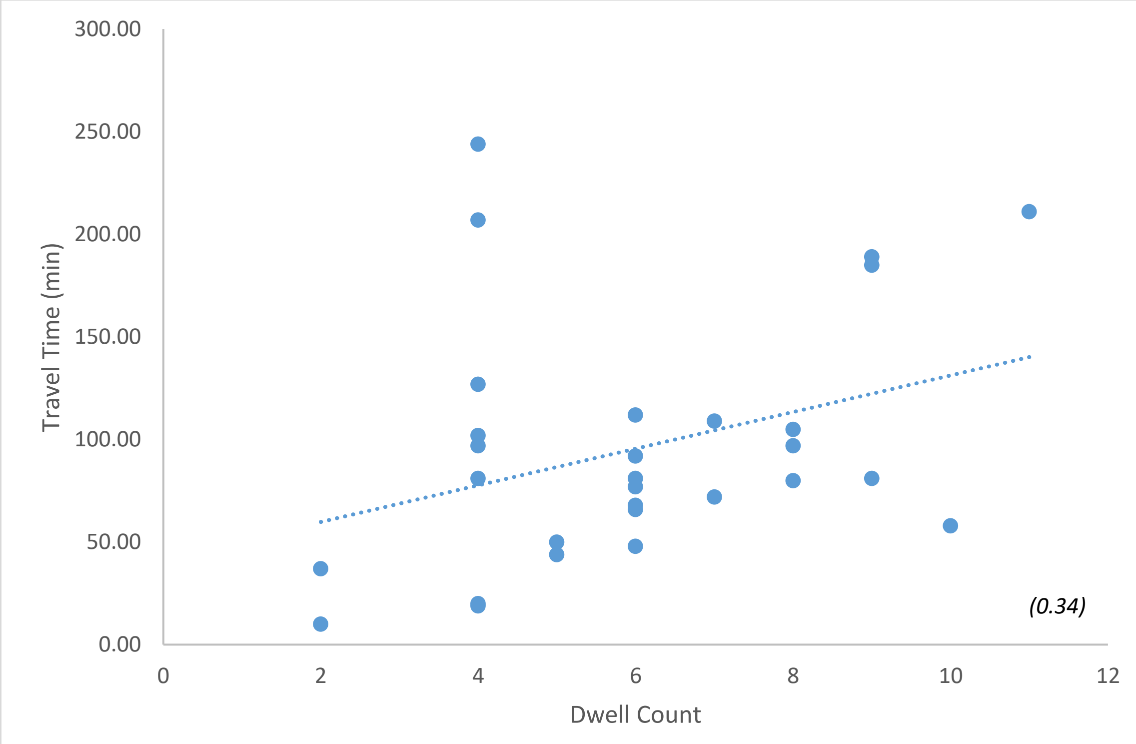

Active Dwell Duration (min)
900.00
800.00
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
0
2
4
6
8
10
12
Dwell Count
(0.13)

	Appendix 7(B): *Dwell Count vs Travel Time (min) across Instruments*
All Instruments	
GPS	

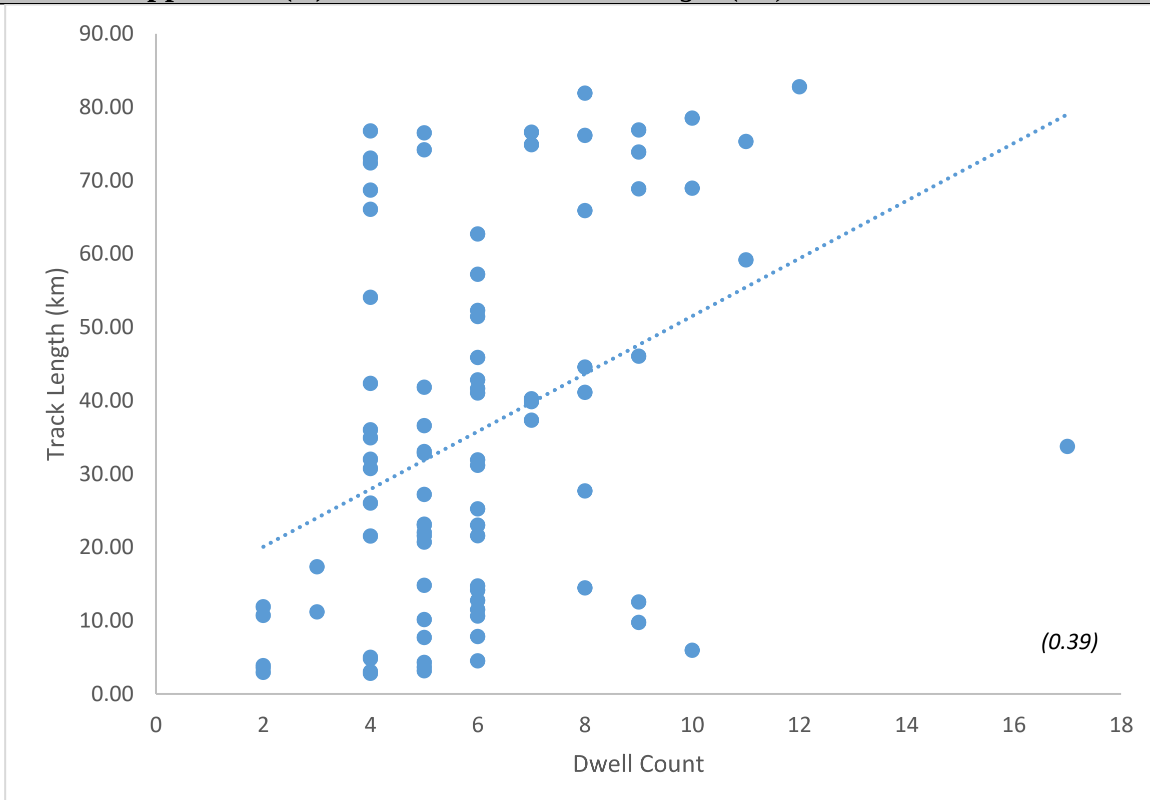
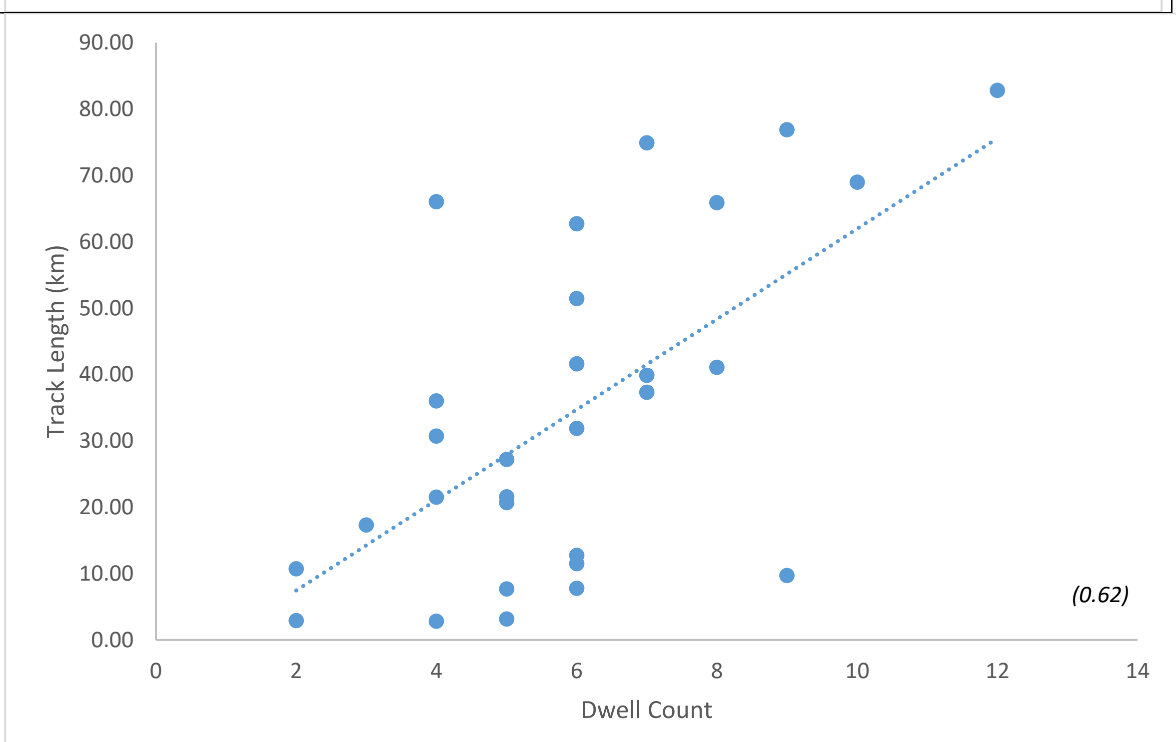
MMR

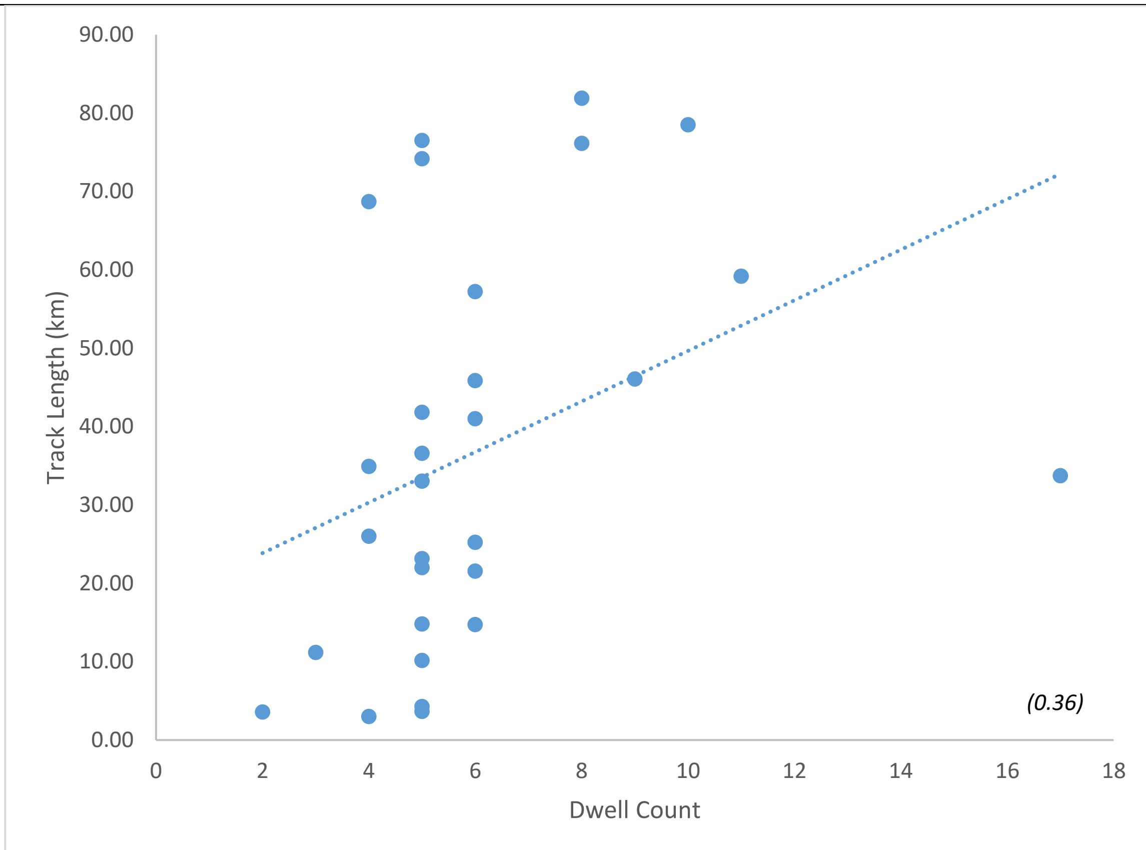
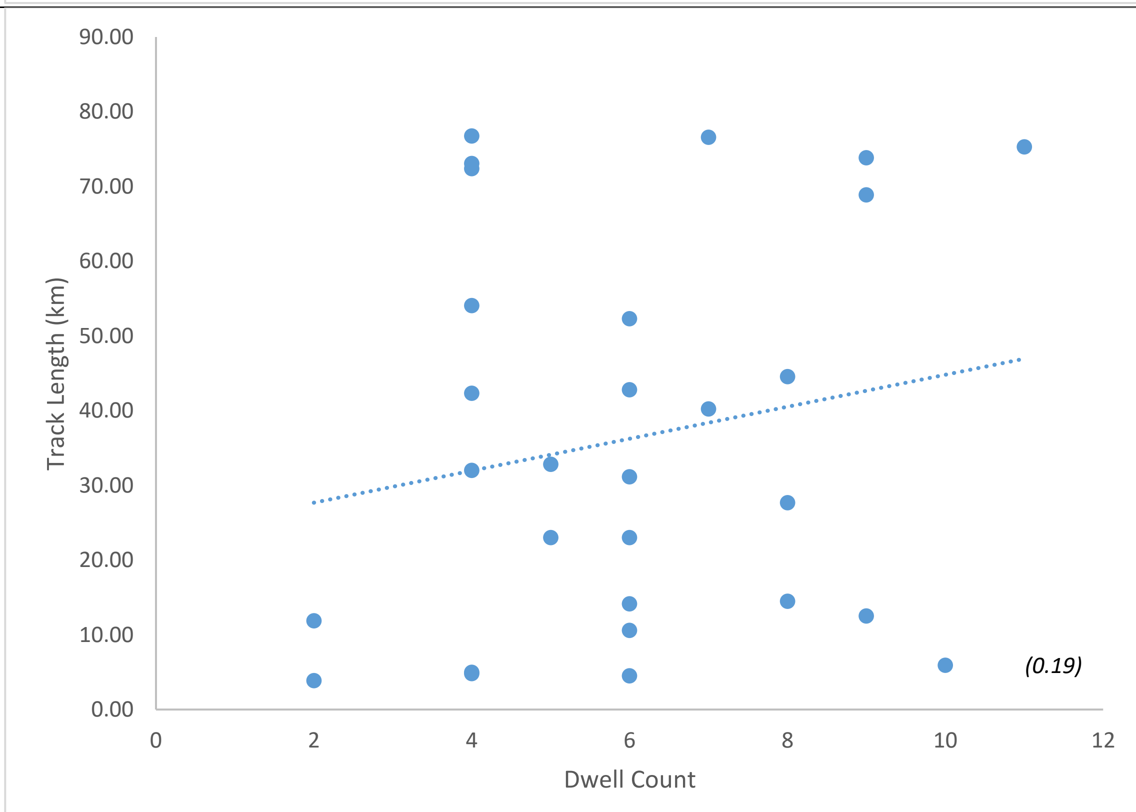

250.00
200.00
150.00
100.00
50.00
0.00
Travel Time (min)
0
2
4
6
8
10
12
14
16
18
Dwell Count
(0.66)

Karma


300.00
250.00
200.00
150.00
100.00
50.00
0.00
Travel Time (min)
0
2
4
6
8
10
12
Dwell Count
(0.34)

(0.34)

	Appendix 7(C): *Dwell Count vs Track Length (km) across Instruments*
All Instruments	
GPS	

MMR	90.00 80.00 70.00 60.00 50.00 40.00 30.00 20.00 10.00 0.00 Track Length (km) 0 2 4 6 8 10 12 14 16 18 Dwell Count (0.36)
Karma	90.00 80.00 70.00 60.00 50.00 40.00 30.00 20.00 10.00 0.00 Track Length (km) 0 2 4 6 8 10 12 Dwell Count (0.19)

References

1. Wan, N., & Lin, G. (2016). Classifying human activity patterns from smartphone collected GPS data: A fuzzy classification and aggregation approach. Transactions in GIS, 20(6), 869-886.
2. Ball, K., Timperio, A. F., & Crawford, D. A. (2006). Understanding environmental influences on nutrition and physical activity behaviors: where should we look and what should we count?. International Journal of Behavioral Nutrition and Physical Activity, 3(1), 33.
3. Lytle, L. A. (2009). Measuring the food environment: state of the science. American journal of preventive medicine, 36(4), S134-S144.
4. Drewnowski, A., Moudon, A. V., Jiao, J., Aggarwal, A., Charreire, H., & Chaix, B. (2014). Food environment and socioeconomic status influence obesity rates in Seattle and in Paris. International journal of obesity, 38(2), 306.
5. Drewnowski, A., Aggarwal, A., Cook, A., Stewart, O., & Moudon, A. V. (2016). Geographic disparities in healthy eating index scores (HEI–2005 and 2010) by residential property values: findings from Seattle obesity study (SOS). Preventive medicine, 83, 46-55.
6. Drewnowski, A. (2004). Obesity and the food environment: dietary energy density and diet costs. American journal of preventive medicine, 27(3), 154-162.
7. Glanz, K., Sallis, J. F., Saelens, B. E., & Frank, L. D. (2005). Healthy nutrition environments: concepts and measures. American journal of health promotion, 19(5), 330-333.
8. Story, M., Kaphingst, K. M., Robinson-O'Brien, R., & Glanz, K. (2008). Creating healthy food and eating environments: policy and environmental approaches. Annu. Rev. Public Health, 29, 253-272.
9. Wolf, J. L. (2000). Using GPS data loggers to replace travel diaries in the collection of travel data (Doctoral dissertation, School of Civil and Environmental Engineering, Georgia Institute of Technology).
10. Wu, L., Yang, B., & Jing, P. (2016). Travel mode detection based on GPS raw data collected by smartphones: a systematic review of the existing methodologies. Information, 7(4), 67.
11. Gong, L., Morikawa, T., Yamamoto, T., & Sato, H. (2014). Deriving personal trip data from GPS data: a literature review on the existing methodologies. Procedia-Social and Behavioral Sciences, 138, 557-565.
12. Gonzalez, P. A., Weinstein, J. S., Barbeau, S. J., Labrador, M. A., Winters, P. L., Georggi, N. L., & Perez, R. (2010). Automating mode detection for travel behaviour analysis by using global positioning systems-enabled mobile phones and neural networks. IET Intelligent Transport Systems, 4(1), 37-49.
13. Cetateanu, A., & Jones, A. (2016). How can GPS technology help us better understand exposure to the food environment? A systematic review. SSM-population health, 2, 196-205.
14. Syed, S., & Cannon, M. E. (2004, January). Fuzzy logic-based map matching algorithm for vehicle navigation system in urban canyons. In ION National Technical Meeting, San Diego, CA (Vol. 1, pp. 26-28).

15. Kestens, Y., Thierry, B., Shareck, M., Steinmetz-Wood, M., & Chaix, B. (2018). Integrating activity spaces in health research: Comparing the VERITAS activity space questionnaire with 7-day GPS tracking and prompted recall. Spatial and spatio-temporal epidemiology, 25, 1-9.
16. Nitsche, P., Widhalm, P., Breuss, S., Brändle, N., & Maurer, P. (2014). Supporting large-scale travel surveys with smartphones–a practical approach. Transportation Research Part C: Emerging Technologies, 43, 212-221.
17. Hirsch, J. A., James, P., Robinson, J. R., Eastman, K. M., Conley, K. D., Evenson, K. R., & Laden, F. (2014). Using MapMyFitness to place physical activity into neighborhood context. Frontiers in public health, 2, 19.
18. Empson, R. (2013, November 14). Under Armour Buys Mobile Workout Veteran MapMyFitness For $150M As It Looks To Keep Pace In Digital Fitness. Retrieved from https://techcrunch.com/2013/11/14/under-armour-puts-down-150m-to-acquire-fitness-tracking-platform-maymyfitness-and-its-20m-users/
19. Miller, C. (2018, July 02). Facebook to shutter three of its mobile applications, including 'Moves' fitness tracking. Retrieved from https://9to5mac.com/2018/07/02/facebook-shutting-down-moves-tbh-hello/
20. QStarz: 'QSTARZ Travel Recorder XT'. Retrieved from http://www.qstarz.com/GPS_products.html#TravelRecorder
21. QStarz: 'Download QTravel'. Available from http://www.qstarz.com/download.php?t=4&m=BT-Q1000XT
22. R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria. Available from https://www.R-project.org/.
23. Carey, Nick. (2005) Establishing Pedestrian Walking Speeds. Retrieved from https://www.westernite.org/datacollectionfund/2005/.
24. Ashbrook, D., & Starner, T. (2003). Using GPS to learn significant locations and predict movement across multiple users. Personal and Ubiquitous computing, 7(5), 275-286.
25. Gong, L., Morikawa, T., Yamamoto, T., & Sato, H. (2014). Deriving personal trip data from GPS data: a literature review on the existing methodologies. Procedia-Social and Behavioral Sciences, 138, 557-565.
26. Wan, N., & Lin, G. (2013). Life-space characterization from cellular telephone collected GPS data. Computers, Environment and Urban Systems, 39, 63-70.
27. Microsoft: 'Microsoft Excel'. Available from https://products.office.com/en-us/excel
28. SAS: 'SAS'. Available from https://www.sas.com/en_us/software/all-products.html
29. McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia Medica, 22(3), 276–282.
30. Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. Anesthesia & Analgesia, 126(5), 1763-1768.
31. Bewick, V., Cheek, L., & Ball, J. (2003). Statistics review 7: Correlation and regression. Critical Care, 7(6), 451–459.
32. Giavarina, D. (2015). Understanding Bland Altman analysis. Biochemia Medica, 25(2), 141–151. http://doi.org/10.11613/BM.2015.015
33. Zahradnik, F. (2018, May 17). What Is Assisted GPS? Retrieved August 15, 2018, from https://www.lifewire.com/assisted-gps-1683306

34. Can I Use the App Without a Data Plan?. Retrieved August 15, 2018, from https://support.mapmyfitness.com/hc/en-us/articles/200118024-Can-I-Use-the-App-Without-a-Data-Plan-
35. Wan, N., Qu, W., Whittington, J., Witbrodt, B. C., Henderson, M. P., Goulding, E. H., ... & Lin, G. (2013). Assessing smart phones for generating life-space indicators. Environment and Planning B: Planning and Design, 40(2), 350-361.