Published in final edited form as:
J Meas Phys Behav. 2021 December ; 4(4): 321–332. doi:10.1123/jmpb.2021-0019.

Validity of a Global Positioning System-Based Algorithm and Consumer Wearables for Classifying Active Trips in Children and Adults

Chelsea Steel[1], **Katie Crist**[2], **Amanda Grimes**[3], **Carolina Bejarano**[4], **Adrian Ortega**[4], **Paul R. Hibbing**[1], **Jasper Schipperijn**[5], **Jordan A. Carlson**[1,6]

[1]Center for Children's Healthy Lifestyles & Nutrition, Children's Mercy Hospital, Kansas City, MO, USA;

[2]Department of Family Medicine, University of California, San Diego, La Jolla, CA, USA;

[3]School of Nursing and Health Studies, University of Missouri-Kansas City, Kansas City, MO, USA;

[4]Clinical Child Psychology Program, University of Kansas, Lawrence, KS, USA;

[5]Department of Sport Sciences and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark;

[6]Department of Pediatrics, Children's Mercy Hospital, University of Missouri Kansas City, Kansas City, MO, USA

Abstract

Objective: To investigate the convergent validity of a global positioning system (GPS)-based and two consumer-based measures with trip logs for classifying pedestrian, cycling, and vehicle trips in children and adults.

Methods: Participants ($N = 34$) wore a Qstarz GPS tracker, Fitbit Alta, and Garmin vivosmart 3 on multiple days and logged their outdoor pedestrian, cycling, and vehicle trips. Logged trips were compared with device-measured trips using the Personal Activity Location Measurement System (PALMS) GPS-based algorithms, Fitbit's SmartTrack, and Garmin's Move IQ. Trip- and day-level agreement were tested.

Results: The PALMS identified and correctly classified the mode of 75.6%, 94.5%, and 96.9% of pedestrian, cycling, and vehicle trips (84.5% of active trips, F1 = 0.84 and 0.87) as compared with the log. Fitbit and Garmin identified and correctly classified the mode of 26.8% and 17.8% (22.6% of active trips, F1 = 0.40 and 0.30) and 46.3% and 43.8% (45.2% of active trips, F1 = 0.58 and 0.59) of pedestrian and cycling trips. Garmin was more prone to false positives (false trips not logged). Day-level agreement for PALMS and Garmin versus logs was favorable across trip modes, though PALMS performed best. Fitbit significantly underestimated daily cycling. Results were similar but slightly less favorable for children than adults.

Carlson (jacarlson@cmh.edu) is corresponding author,.

Conclusions: The PALMS showed good convergent validity in children and adults and were about 50% and 27% more accurate than Fitbit and Garmin (based on F1). Empirically-based recommendations for improving PALMS' pedestrian classification are provided. Since the consumer devices capture both indoor and outdoor walking/running and cycling, they are less appropriate for trip-based research.

Keywords

cycling; fitbit; Garmin; transportation; walking

Pedestrian (e.g., walking) and cycling trips are important sources of physical activity in children and adults (Sahlqvist et al., 2012; Schoeppe et al., 2013). Increasing time spent walking and cycling among adults is recommended for population health and is a goal of the U.S.'s Healthy People 2030 initiative (U.S. Department of Health and Human Services, 2021) as well as global efforts (Giles-Corti et al., 2016; Global Advocacy for Physical Activity, & Advocacy Council of the International Society for Physical Activity Health, 2012; World Health Organization, 2019). However, active trips can be challenging to quantify. Identifying valid tools for objectively measuring active trips is critical for advancing physical activity research, particularly research involving the built environment and/or trip-based (e.g., walking) interventions. For example, objective data on physical activity accumulated through pedestrian and cycling trips are important to cite as evidence supporting public investment in built environment improvements (Saelens et al., 2003; Sallis et al., 2016). In addition, as numerous health interventions aim to increase pedestrian or cycling trips, device-based measures of these behaviors are becoming increasingly important for evaluation (Bird et al., 2013; Ogilvie et al., 2007; The Community Guide, 2021).

Time spent in active trips has been most commonly assessed in health research using self-report measures (Craig et al., 2003; Raza et al., 2020; Saunders et al., 2013). More recently, device-based measures have been developed that can address some limitations and inaccuracies of self-report measures (Prince et al., 2008). Device-based trip classification tools used in research have primarily involved wearable global positioning system (GPS) receivers, sometimes combined with accelerometers, and computational algorithms for processing the location information. The data processing algorithms generally use speed and distance information to identify when trips occur and classify trip mode (e.g., pedestrian, cycling, in vehicle; Carlson, Jankowska, et al., 2015; Chaix et al., 2014; Cho et al., 2011; Ellis et al., 2014; Kang et al., 2013). Only one set of processing algorithms is currently available to end users through an easily accessible computational interface. These GPS-based trip processing algorithms were originally part of the Personal Activity Location Measurement System (PALMS; Carlson, Jankowska, et al., 2015) and recently incorporated to the Human Activity Behavior Identification Tool and Data Unification System (HABITUS) (HABITUS, 2020). While research has shown that these algorithms have good validity in adult commuter cyclists (Carlson, Jankowska, et al., 2015), less is known about their validity in children and more generalizable samples of adults.

Consumer wearables, like those from Fitbit and Garmin, are also able to classify time spent in pedestrian and cycling trips. The trip classification features of these wearables, Fitbit SmartTrack and Garmin Move IQ, automatically apply proprietary algorithms to accelerometer data from the wearables and do not require input from the user. Consumer devices have some benefits over research devices, such as greater appeal to participants, long-term data collection, near real-time data transmission, and ease of data processing for end users, which make them important to investigate. Research has shown that wearable devices can have moderate to high validity for measuring energy expenditure, moderate to vigorous physical activity, and step counts, though less is known about their utility and validity for classifying active trips (Evenson et al., 2015; Ferguson et al., 2015; Fuller et al., 2020; O'Driscoll et al., 2020). While a previous study investigated Fitbit and Garmin trip classification features and found them to have good validity, the study was limited by using a controlled protocol with little variation in trip lengths, and children were not included (Dorn et al., 2019). A next step is to investigate these tools when applied to free-living activities in multiple age groups.

This study aimed to investigate the convergent validity of three trip classification tools by testing their agreement with trip logs in children and adults across multiple days of wear time during free-living activities. The three tools included the PALMS/HABITUS GPS-based algorithms, Fitbit SmartTrack feature, and Garmin Move IQ feature.

Methods

Participants and Procedures

Participants were 18 adults (77.8% were women), ages 18–60 years (mean age = 40.53, *SD* = 9.87), and 16 children (43.7% were girls), ages 9–17 years (mean age = 14.5, *SD* = 1.77), identified through postings to Children's Mercy Hospital internal website, word-of-mouth, and referrals. Some child participants were children of the adult participants. Data were collected in spring 2018. Inclusion criteria were the ability to speak and read English and be able to follow the study procedures described below. As cycling trips were required for participation, access to a bicycle was also an inclusion criterion, though the participant did not need to have a history of cycling or self-identify as a cyclist. Study procedures included wearing four activity monitors concurrently for at least 10 hr per day for up to 4 days, engaging in at least two pedestrian trips, two cycling trips, and two vehicle trips over the wear period, and logging all trips in a trip log. For the pedestrian trips, participants could choose to walk or run. Participants were asked that at least one of the pedestrian trips and one of the cycling trips last at least 10 min. The participant sample size was based on previous trip-based measurement studies with similar analyses and acceptable limits of agreement (Dorn et al., 2019), with the data collection goals for the present study being ≥60 trips per mode, ≥2 days of wear time per participant, and ≥30 participants. Participants were offered a $50 cash card as an incentive. The study was approved by the Children's Mercy Hospital protections committee.

Measures

Trip Log—A pocket-sized paper trip log was given to all participants with spaces to log the start time, end time, and mode of each trip they conducted. Participants were instructed to log all trips taken while wearing the devices if they occurred outdoors and spanned at least 100 yards. Participants were instructed to write the precise minute each trip started and ended. The accuracy of the log was emphasized during the consent appointment and upon enrollment as being the most important aspect of the study, as it provided the comparison measure for testing convergent validity. For all child participants, the parent or caregiver was asked to assist their child in completing the log accurately. In some cases, both the parent and the child participated in the study together. Automated text reminders were sent each evening to remind participants to log all trips taken while wearing the devices. Upon study completion, study staff reviewed the trip log with the participant to quality check for missing trip modes, end times, or retroactively log trips that were not recorded.

GPS Data Collection and Processing—A Qstarz BT-Q1000XT GPS receiver was worn on a belt at the hip. The device collected latitude, longitude, elevation, and satellite data in 15-s epochs. The data were processed in PALMS (Qstarz, Taipei, Taiwan) to classify pedestrian, cycling, and vehicle trips (Carlson, Jankowska, et al., 2015). The algorithms allow flexibility over several parameter settings. In the present study, the following settings were used based on our prior testing and use of the algorithms (Carlson et al., 2020; Carlson, Jankowska, et al., 2015; Carlson, Saelens, et al., 2015). Groups of sequential fixes (≥1 min) were considered trips if they spanned ≥100 m with an average speed ≥1 km/hr. Pauses of up to 3 min were allowed within a trip. Trips with a 90th percentile speed ≥25 km/hr were classified as vehicle trips. Trips with a 90th percentile speed between 10 and 24 km/hr were considered cycling trips, and trips with 90th percentile speeds lower than 10 km/hr were considered pedestrian trips. The signal-to-noise ratio (SNR) was used to remove trips that occurred indoors as well as eliminate false trips that occur when there is signal interference due to limited satellite access while indoors. The SNR was not used to remove indoor points from the beginning or end of trips, which is an option in PALMS. Points with an SNR < 225 were considered to be indoors (Lam et al., 2013), and "trips" with ≤50% of points outdoors were recoded as nontrips. This was to minimize periods of GPS scatter (i.e., inaccurate GPS fixes bouncing around a location) from being misclassified as trips. Post-PALMS processing involved aggregating the trip information to the minute level, with a given minute being classified as a trip if PALMS had indicated that any epochs during that minute were part of a trip.

Accelerometer Data Collection and Processing—Participants wore an ActiGraph GT3X+ (ActiGraph, Pensacola, FL) accelerometer at their right hip affixed to the same belt as the GPS device. Wear time for the ActiGraph was used to infer valid wear time for the GPS monitor and the consumer wearables, as participants were instructed to wear all at the same time. ActiLife software was used to download device data and identify valid wear time using the Choi algorithm applied to vector magnitude counts per minute using a 90-min window, 30-min stream frame, and 2-min tolerance (Choi et al., 2011, 2012).

Fitbit and Garmin Data Collection and Processing—Participants were asked to wear a Fitbit Alta device on their preferred wrist, in accordance with recommendations from the manufacturer (FitBit, 2020a). The Garmin vívosmart 3 was always worn on the nondominant wrist as recommended by the manufacturer (Garmin, 2020a). At the start of the measurement period, the Fitbit account was updated to reflect the chosen wrist of the participant (dominant vs. nondominant). The devices were connected to a Fitbit and Garmin account to retrieve the data at the end of each participant's wear period. The proprietary consumer wearable automatic activity detection feature, called SmartTrack for Fitbit and Move IQ for Garmin, automatically classifies activities, such as walking, running, cycling, and swimming based on algorithms applied to the accelerometer data. Vehicle time was not an activity type included in the classifications and thus was not investigated for the consumer devices. Participants were instructed not to use the user interface of the device to start an activity manually and instead allow the SmartTrack and Move IQ features to capture the trips. Fitbit and Garmin each state that the activity must occur for at least 10 consecutive minutes to be detected (FitBit, 2020b; Garmin, 2020b), with the exception for Garmin running activities, which have a minimum detection threshold of 5 min (Garmin, 2020b). We grouped walking and running together as pedestrian trips. The consumer device software provided the start and end time (date, hour, and minute) and mode (pedestrian or cycling) of each trip classified by each device.

Analysis

Trip-level agreement for trip detection and trip mode classification was calculated for each test measure (PALMS, Fitbit, and Garmin) as compared with the trip logs and presented using confusion matrices. Sensitivity (i.e., recall) and F1 scores were also calculated for each trip mode, with the latter used as a measure of accuracy based on the harmonic mean of sensitivity and positive predictive value. PALMS-, Fitbit-, and Garmin-detected trips were considered to be in agreement with the trip log if at least 1 min of the logged trip was detected by the test measure. A "false trip" (i.e., false positive) was defined as any trip detected by the test measure that did not have at least 1 min of overlap with a logged trip. A "missed trip" (i.e., false negative) was defined as any logged trip that did not have at least 1 min of overlap with a trip detected by the test measure. Fitbit and Garmin analyses were conducted both with and without logged trips under the minimum detection threshold (10 min) included in the log (comparison) data set. The trip-level analyses were conducted for the full sample (all ages) and separately for children and adults. The distributions of the duration of logged trips that failed to be detected by the consumer devices (i.e., missed trips) were plotted to identify whether most were close to the 10-min detection threshold or well over the threshold, after excluding logged trips <10 min.

For duration agreement, day-level variables were computed for time spent in each trip mode (in minutes per day) as indicated by each measurement method. Means and *SD*s were used to describe variables across measurement methods. Mean absolute error (MAE) and Spearman correlations were used to test day-level agreement between each test measure and the trip log. Spearman correlations were interpreted as small (≤.40), moderate (.41–.60), large (.61–.80), and very large (.81–1.0) (Landis & Koch, 1977). In addition, bias values and confidence limits representing the difference between measurement methods were assessed using mixed-effects linear regression, accounting for nesting of days within participants, and results were plotted across models. Within these models, interaction terms were used to test whether day-level agreement (i.e., bias) differed by age group (children vs. adults). For the day-level analyses among the consumer devices, separate models were conducted for (a) trips with logged durations over the minimum detection threshold and (b) all trips. All analyses were performed in SPSS Statistics (version 24.0; IBM Corp., Armonk, NY) (IBM Corp., 2016).

Results

The mean number of days devices were worn was 2.6 (*SD* = 1.1), for a total of 89 days of data (34 from children and 55 from adults). The devices were worn for a mean of 9.7 hr per day (*SD* = 4.2). Participants logged 380 total trips, with 151 trips conducted by children and 229 by adults (Figure 1). Over half of the trips were vehicle trips, comprising 62.3% of all trips conducted by children and 57.2% of all trips conducted by adults. Pedestrian trips had a mean duration of 10.7 min (*SD* = 11.4), cycling trips had a mean duration of 11.8 min (*SD* = 10.4), and vehicle trips had a mean duration of 21.4 min (*SD* = 20.1). Mean trip duration was similar between children and adults, within ≤0.6 min.

Trip-level agreement between PALMS and the logs was similar for children and adults (Table 1). The PALMS identified and correctly classified the mode of 75.6%, 94.5%, and 96.9% of pedestrian, cycling, and vehicle trips (84.5% of all active [pedestrian and cycling] trips) as compared with the trip log (i.e., sensitivity). The F1 scores were 0.84 for pedestrian, 0.87 for cycling, and 0.98 for vehicle trips. The PALMS had only one false trip (0.3% of all PALMS trips), which inspection of the GPS data suggested was caused by GPS scatter. The PALMS failed to identify (i.e., missed) 2.1% of all logged trips, most of which were pedestrian trips. The PALMS misclassified the mode of 6.1% of all logged trips, and the most common type of misclassification was classifying a logged pedestrian trip as a cycling trip. Inspection of the GPS data revealed that four of the logged pedestrian trips that were misclassified by PALMS as cycling trips had a high pedestrian speed, indicative of running. Two of the logged walking trips that were misclassified as cycling trips had a vehicle trip occur just before the pedestrian trip. For these pedestrian trips, PALMS grouped the latter part of the vehicle trip with the pedestrian trip, which led to a higher average speed for the pedestrian trip thus leading it to be classified as cycling. The remaining five logged pedestrian trips that were misclassified as cycling trips spanned both outdoors and indoors. The outdoor time led to a high SNR value, which prevented the trip from being removed, but the indoor time led to GPS scatter due to poor communication with the GPS satellite when indoors, which led to a higher average speed and thus the trip being classified as cycling.

The removal of logged pedestrian or cycling trips lasting <10 min (to match the consumer device manufacturer’s minimum detection threshold) resulted in the omission of 39 trips, with 56.4% of these removed trips being logged pedestrian trips (could include runs). Only one (2.6%) of these 39 logged trips had an associated trip classified by a consumer device, and it was a pedestrian trip that was identified by Garmin to last 15 min. Thus, the removal of logged trips lasting <10 min only contributed to one false trip in the Garmin data and no false trips in the Fitbit data set.

When considering all logged pedestrian and cycling trips, including those lasting <10 min, Fitbit identified and correctly classified the mode of 26.8% and 17.8% of pedestrian and cycling trips (22.6% of all active trips) as compared with the trip log. When considering only logged trips lasting at least 10 min, Fitbit identified and correctly classified the mode of 36.7% and 23.2% of pedestrian and cycling trips (30.2% of all active trips) as compared with the trip log. Fitbit did not misclassify the mode of any logged trips. Seven trips that were "identified" by Fitbit were not true trips, as they did not appear in the trip logs (i.e., 16.7% of all Fitbit trips were false trips). Fitbit failed to identify 77.4% of all logged trips and 69.8% of logged trips lasting at least 10 min. Figure 2a and 2b shows that the logged pedestrian and cycling trips lasting at least 10 min that Fitbit failed to identify had a mean duration of 16.9 min (*SD* = 11.1) and 14.5 min (*SD* = 7.0), respectively. While a large proportion of the logged trips lasting between 10 and 14.9 min were missed by Fitbit (79.3% and 81.6% of all logged trips between 10 and 14.9 min), and a large proportion of the missed trips were between 10 and 14.9 min (60.5% and 72.1% of all missed pedestrian and cycling trips), Fitbit also missed many trips that lasted longer than 15 min. Trip-level agreement between Fitbit and the trip logs was sometimes poorer in children as compared with adults.

When considering all logged pedestrian and cycling trips, including those lasting <10 min, Garmin identified and correctly classified the mode of 46.3% and 43.8% of pedestrian and cycling trips (45.2% of all active trips) as compared with the trip log. When considering only logged trips lasting at least 10 min, Garmin identified and correctly classified the mode of 61.7% and 57.1% of pedestrian and cycling trips (59.5% of all active trips) as compared with the trip log. Garmin did not misclassify the mode of any logged trips. Fifteen trips, or 17.6%, of all trips "identified" by Garmin were false trips, as they did not appear in the trip logs. (This includes logged trips under 10 min.) Garmin failed to identify 54.8% of all logged trips and 40.2% of logged trips lasting at least 10 min. Figure 2c and 2d shows that the logged trips lasting at least 10 min that Garmin failed to identify had a mean duration of 18.3 min (*SD* = 13.7) for pedestrian and 14.3 min (*SD* = 9.1) for cycling trips. A large proportion of the logged trips lasting between 10 and 14.9 min were missed by Garmin (41.4% and 75.0% of all logged trips between 10 and 14.9 min), and a large proportion of the missed trips were between 10 and 14.9 min (52.2% and 87.5% of all missed pedestrian and cycling trips), though Garmin also missed many pedestrian trips that lasted longer than 15 min. Regarding cycling trips, Garmin missed few that were longer than 15 min. Trip-level agreement between Garmin and the trip logs was sometimes poorer in children as compared with adults.

Table 2 presents day-level agreement between each test measure and the trip logs. For PALMS versus the trip logs, MAEs were 40.6% of the log mean for time spent in pedestrian trips, 21.0% for time spent in cycling trips, and 13.0% for times spent in vehicle trips. All Spearman correlations were very large (.81, .85, and .94). For Fitbit versus trip log, after excluding trips <10 min from the log-based calculations, MAEs were 44.0% of the log mean for time spent in pedestrian trips and 75.7% for time spent in cycling trips. Spearman correlations were large (.71) and moderate (.40). For Garmin versus trip log, after excluding trips <10 min from the log-based calculations, MAEs were 62.6% of the log mean for time spent in pedestrian trips and 51.9% for time spent in cycling trips. Spearman correlations were large (.68 and .70). Agreement between the consumer device features and trip logs was similar but generally slightly poorer when logged trips with durations <10 min were included in the analyses.

Figure 3 shows the direction and magnitude of bias for each test measure relative to the trip logs. The bias values for PALMS were nonsignificant for all trip modes, though on average PALMS slightly underestimated time spent in pedestrian trips by 3.0 min/day (19.6%). Fitbit slightly underestimated time spent in pedestrian trips by an average of 4.3–5.4 min/day (30.9%–35.7%), but these biases were nonsignificant. Fitbit significantly underestimated time spent cycling by an average of 6.1–7.2 min/day (58.4%–62.1%). The bias values for Garmin were nonsignificant for all trip modes, though on average Garmin slightly underestimated time spent cycling by an average of 2.5–3.1 min/day (23.5%–27.2%).

The bias estimates for each measurement method shown in Table 2 did not differ significantly between children and adults as indicated by nonsignificant interactions. The following regression coefficients and CIs reflect the difference in the level of overestimation or underestimation of time (in minutes per day) in each trip mode for children as compared with adults: PALMS (walk, $B = -2.53$ [16.7% of log mean], 95% CI [−14.03, 8.97]; cycle, $B = -2.11$ [18.2% of log mean], 95% CI [−11.71, 7.47]; vehicle, $B = -1.48$ [2.5% of log mean], 95% CI [−32.76, 29.81]); Fitbit, excluding trips under 10 min (walk, $B = 0.40$ [2.9% of log mean], 95% CI [−10.60, 11.40]; cycle, $B = -2.51$ [24.2% of log mean], 95% CI [−11.22, 6.20]); and Garmin, excluding trips under 10 min (walk, $B = -8.50$ [61.1% of log mean], 95% CI [−21.02, 4.03]; cycle, $B = -3.98$ [38.3% of log mean], 95% CI [−12.95, 4.98]). An example interpretation is that PALMS underestimated time in pedestrian trips by about 4.53 min/day (29.8%) in children ($12.22 - 15.19 - 2.53 \times 0.62$) and 2.01 min/day (13.2%) in adults ($12.22 - 15.19 + 2.53 \times 0.38$). The values 0.62 and 0.38 reflect the proportion of the 89 wear days that were from adults and children, respectively.

Discussion

Present findings indicated the PALMS GPS-based trip algorithms have good validity among children and adults for detecting trips and accurately classifying trip mode across pedestrian, cycling, and vehicle trips. The PALMS was about 50% and 27% more accurate than the Fitbit SmartTrack and Garmin Move IQ consumer device features, respectively, based on differences in trip-level F1 scores. These consumer device features captured indoor walking/running and cycling, which led to false trips, and were unable to detect active trips under 10 min and all vehicle trips, as designed by the manufacturers. Given that increasing engagement in active trips is a public health priority that has been highlighted in the United States (U.S. Department of Health and Human Services, 2021) and globally (Giles-Corti et al., 2016; Global Advocacy for Physical Activity, & Advocacy Council of the International Society for Physical Activity Health, 2012; World Health Organization, 2019) and targeted by numerous interventions and policy efforts (Bird et al., 2013; Ogilvie et al., 2007; The Community Guide, 2021), there is a need to more widely incorporate trip-based measures into research and practice. Objective measures in particular are advantageous when accurate and detailed trip information is desired while minimizing participant burden. The PALMS trip algorithms provide an opportunity to incorporate trip-based measures with good validity more widely into research and practice through the use of GPS trackers, which have been shown to be acceptable for use in numerous population groups (Kerr et al., 2011).

The PALMS algorithms had good validity at both the trip and day level. Their ability to accurately capture the duration, mode, and approximate timing of each trip while minimizing false positives (false trips) allows them to support investigation of research questions that require day level or within-day information. The PALMS algorithms performed similarly in the present study as they did in a previous validation study that used annotated images from person-worn cameras (SenseCams; the SenseCam; Microsoft, Redmond, WA) as the comparison measure. The slight underestimation of daily time spent in each trip mode by PALMS in the present study, which is in contrast to the previously mentioned study showing slight overestimation by PALMS, may have been due to overestimation of trip time in the logs, which has been shown previously (James et al., 2016). There has been one other study that tested the validity of the PALMS algorithms against trip diaries (Kang et al., 2018). That study found poorer validity, mostly caused by overestimation of trips by PALMS (i.e., false trips), which was likely due to the use of an older GPS device that did not capture SNR leading to GPS scatter when indoors to mimic trip speeds and distances. In the present study, the PALMS algorithms performed as well as or better than other GPS-based trip processing algorithms, some of which incorporate additional information, such as from geographic information systems and/or accelerometers (Brondeel et al., 2015; Cho et al., 2011; Ellis et al., 2014; Gong et al., 2012; Kang et al., 2013; Troped et al., 2008). An advantage of the PALMS algorithms is that they are available to end users without computational or programming expertise through a user-friendly interface that supports parameter customization (i.e., HABITUS). The present study extends previous findings on PALMS by showing good validity among children and a more generalizable sample of adults, as the previous study was limited to adults who self-identified as commuter cyclists and thus may have different cycling patterns than other population groups (e.g., higher speeds and longer distances). Relatedly, in the study of commuter cyclists, using a speed of ≥35 km/hr to infer cycling resulted in misclassification of some cycling trips and vehicle trips, whereas, in the present study, only one cycling trip was misclassified as a vehicle trip despite a lower speed of ≥25 km/hr being used. This highlights the importance of using population-appropriate speed thresholds in PALMS, which are user adjustable, for distinguishing between trip modes. The current findings also add to previous evidence by showing there is room for improvement in PALMS' validity for classifying pedestrian trips.

The primary source of error among PALMS' classification of pedestrian trips was that pedestrian trips were sometimes misclassified as cycling trips due to three circumstances (a) fast running speeds, (b) a pedestrian trip spanning outdoors and indoors with indoor satellite interference, and (c) the grouping of a pedestrian trip immediately following or preceding a vehicle trip. Addressing these circumstances would primarily result in improved accuracy in pedestrian and cycling trip classification but also in vehicle trip classification, because misclassification of one mode impacts other modes. Our recommendation for the first circumstance is to identify habitual runners using a baseline survey and employ a higher speed threshold among those participants to distinguish between pedestrian and cycling trips (e.g., 10–12 km/hr). The second circumstance may be addressable by manipulating user adjustable PALMS parameters. Increasing the proportion of the trip that is required to be outdoors for the trip to be considered valid (50% was used in the present study), and/or removing indoor points from the beginning and end of trips, both of which are based on SNR, may result in the outdoor portion of the trip being appropriately classified as pedestrian. However, we have observed previously (Carlson et al., 2020) that a higher threshold can result in failure to detect short vehicle trips (e.g., to school), which can have lower SNR values due to being inside a vehicle and/or immediately leaving an indoor location (e.g., home). Although the thresholds for these parameters are applied to all trip modes, a potential solution may be to run the algorithms with two different thresholds (a lower, e.g., 50%, and higher, e.g., 75%–90%) and merge the output data, using the higher threshold for pedestrian and cycling trips and lower threshold for vehicle trips. However, caution should be used when employing a higher threshold in urban areas, as research has shown that the presence of large buildings can impact SNR, falsely indicating indoor time (Schipperijn et al., 2014). Using a threshold lower than 50% is not recommended because it is likely to result in many false trips, as we and others have shown previously (Carlson, Jankowska, et al., 2015; Kang et al., 2018). The third circumstance listed above may be the most challenging to address and is likely to require modifications to the PALMS algorithms themselves or post-PALMS processing approaches, rather than manipulating user adjustable parameters. In the present study, we observed that these pedestrian trips, in addition to being misclassified as cycling, immediately preceded or followed a true vehicle trip. Thus, part of the vehicle trip (very beginning or very end), though not the entire vehicle trip, was being grouped with the pedestrian trip. Such circumstances could be flagged relatively easily in the PALMS output data set by identifying when a predicted cycling trip immediately preceded or followed a predicted vehicle trip. Each circumstance could then be visualized, or a speed-based algorithm could be created, to determine the proper point of mode shift as well as the proper mode of each trip segment. This would result in an extension of the predicted vehicle trip, shortening of the predicted cycling trip, and reclassification of the predicted cycling trip as pedestrian. Another solution may be to incorporate accelerometer data into the PALMS algorithms, which currently only utilize GPS information. Future research is needed to test these recommendations, and users should use caution and visual data inspection when considering parameter adjustments and/or post-PALMS corrections.

Though some sample-level validity estimates for the consumer device features were acceptable (e.g., some biases), individual trip-level validity was generally poor, even when considering only trips that met the minimum duration of 10 min. Both Fitbit's SmartTrack and Garmin's Move IQ failed to detect a substantial number of trips over 10 min while also exhibiting many false positives (false trips), primarily for pedestrian trips. While it is possible that some of the trips that the consumer devices failed to detect were overestimated in the logs (i.e., a trip that truly lasted <10 min, under the detectable threshold, may have been recorded in the log as lasting more than 10 min), both consumer devices also failed to detect many trips that lasted more than 15 min, except that Garmin missed few cycling trips over 15 min. Thus, overestimation of trip duration in the trip logs was not likely a major cause of the missed trips in general. Since SmartTrack and Move IQ aim to detect walking and running in any location, rather than only outdoor walking and running, it is likely that at least some of the false pedestrian trips were pedestrian activities that occurred indoors. Interestingly, missed trips were more common among children, whereas false trips were more common among adults. The former may have been due to children's unique movement patterns (Welk et al., 2000), whereas the latter may have been due to adults walking or moving around indoors, such as while shopping or doing household work (which could meet the consumer device manufacturer's conceptualization of walking). This highlights the challenge of solely using movement information from accelerometers to infer pedestrian and cycling trips, as is done by the consumer device features. Though day-level validity estimates were more favorable than the trip-level estimates, this was largely due to false trips contributing to (i.e., increasing) the daily time spent in each trip mode. Taken together, these consumer device features do not appear to have acceptable validity for many research applications, particularly when aiming to capture outdoor trips. We recommend exerting caution even when using these features for providing day-level information on time spent in outdoor pedestrian and cycling trips, as present findings suggest these measures can overestimation or underestimate daily time spent in pedestrian or cycling trips by ~50%–75%. When investigators are interested in all walking and/or running, rather than only outdoor pedestrian trips as is typically the case in trip-based or active transportation research, consumer wearables may be more appropriate. Incorporating GPS information may be promising for improving the validity of trip classification from consumer devices and better matching an outdoor definition of pedestrian trips.

Though the selection of measures should be informed by their validity, there are several additional factors to consider and trade-offs between research and consumer devices. Although the PALMS algorithms are provided in a user-friendly interface, some expertise is required to collect and process the data and to understand the implications of manipulating parameter settings. Fees are likely to be associated with using the algorithms to offset computing costs. The consumer devices require little to no expertise or computing/ processing fees unless an application programming interface (API) is being used, though we recommend using the APIs for determining wear time as detailed in the following section. Consumer devices can be particularly advantageous for monitoring and feedback, and through the use of an API, information can be transmitted in near real time to inform interventions (Balbim et al., 2021; Brickwood et al., 2019). Such capabilities are not available through GPS trackers like the one used in the present study but are becoming available through the use of GPS and accelerometer information from smartphones (Hurvitz et al., 2014; King et al., 2016; Wang et al., 2021). Although consumer devices have the advantage of measuring overall physical activity and step counts, combined GPS and accelerometer data provide substantially more information about physical activity locations that can support investigation of innovative and important location-based research questions not possible with consumer device data (Duncan et al., 2009; Jankowska et al., 2015). In addition to an inability to capture pedestrian and cycling trips lasting <10 min and distinguish between indoor and outdoor activities, the Fitbit SmartTracks and Garmin Move IQ features do not capture vehicle trips. Vehicle trips are important to investigate in health research because they can be an important source of sedentary time and are important to consider when targeting mode shifts (e.g., switching from vehicle to pedestrian). Another advantage of PALMS is that users can customize parameters for their target population or study purpose, whereas consumer device algorithms generally are inflexible and algorithms can evolve or drastically change over time with little transparency.

Strengths and Limitations

Study strengths included the comparison of validity estimates between children and adults and across research and consumer measures. The focus on free-living activities supports generalizability of the current findings over previous research that investigated the consumer trip classification features using trips of prescribed and uniform durations occurring in the same locations (treadmill and university track; Dorn et al., 2019). Study limitations included the reliance on trip logs for the comparison measure, which can have inaccuracies, though the importance of the logs was emphasized to participants and logs have been shown to have good validity (James et al., 2016). It is possible that some of the trips that failed to be detected by the consumer devices occurred when the participant was not wearing the device, because nonwear time was not measured for the consumer devices. Future studies can address this limitation by using the companies' APIs, which provide epoch-level (e.g., every 15 min for Garmin) information that can be used to infer wear time. Since the present study focused on only three trip modes, it is unclear how the measures would perform on other modes, such as skate-board or scooter trips. The use of public transit, although not directly assessed, was likely low in the present sample, which limits generalizability of findings to public transit users. Public transit is known to create challenges with GPS-based trip classification due to the high frequency of multimode trips (e.g., walking to transit) and potential for satellite interference (e.g., when riding subway). Differentiating between transit and private vehicle trips is also an important measurement area to explore in future studies, as increasing transit use is a public health priority for both health and environmental reasons. Similarly, future studies should aim to differentiate between walking and running, as the latter is less likely to be a target of interventions to increase active transportation. It is important to note that the present study defined trips as occurring outdoors and spanning at least 100 yards. Though walking, and to a lesser extent running and cycling, can occur indoors, indoor trips are often missed by GPS-based classification measures due to poor communication with GPS satellites resulting in GPS scatter. Since the PALMS was recently discontinued and transferred to the HABITUS, it is important for future research to confirm comparability of the algorithms' performance between platforms.

Conclusions

The use of valid and appropriate measures of trips and trip mode is critical for advancing research and practice in physical activity promotion. There is growing evidence in support of the validity and utility of GPS-based trip measures, such as the PALMS algorithms, which present findings indicate have good validity for a range of trip- and day-level applications in both children and adults. Consumer devices may be useful for measuring pedestrian and cycling activity in some circumstances but concerns over validity for measuring outdoor trips limit their overall utility for active transportation and other trip-based research. Given the importance of active trips to overall physical activity (Sahlqvist et al., 2012; Schoeppe et al., 2013), more widespread application of trip-based measures is needed to accelerate physical activity promotion interventions and policies.

References

Balbim GM, Marques IG, Marquez DX, Patel D, Sharp LK, Kitsiou S, & Nyenhuis SM (2021). Using Fitbit as an mHealth Intervention tool to promote physical activity: Potential challenges and solutions. JMIR mHealth and uHealth, 9(3), e25289. 10.2196/25289 [PubMed: 33646135]

Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, & Powell J (2013). Behavior change techniques used to promote walking and cycling: A systematic review. Health Psychology, 32(8), 829. 10.1037/a0032078 [PubMed: 23477577]

Brickwood K-J, Watson G, O'Brien J, & Williams AD (2019). Consumer-based wearable activity trackers increase physical activity participation: Systematic review and meta-analysis. JMIR mHealth and uHealth, 7(4), e11819. [PubMed: 30977740]

Brondeel R, Pannier B, & Chaix B (2015). Using GPS, GIS, and accelerometer data to predict transportation modes. Medicine & Science in Sports & Exercise, 47(12), 2669–2675. 10.1249/MSS.0000000000000704 [PubMed: 25984892]

Carlson JA, Jankowska MM, Meseck K, Godbole S, Natarajan L, Raab F, Demchak B, Patrick K, & Kerr J (2015). Validity of PALMS GPS scoring of active and passive travel compared to SenseCam. Medicine & Science in Sports & Exercise, 47(3), 662. 10.1249/MSS.0000000000000446 [PubMed: 25010407]

Carlson JA, Saelens BE, Kerr J, Schipperijn J, Conway TL, Frank LD, Chapman JE, Glanz K, Cain KL, & Sallis JF (2015). Association between neighborhood walkability and GPS-measured walking, bicycling and vehicle time in adolescents. Health & Place, 32, 1–7. 10.1016/j.healthplace.2014.12.008 [PubMed: 25588788]

Carlson JA, Steel C, Bejarano C, Grimes A, Shook RP, Reddy A, Green M, Morefield T, Steele L, & Campbell K (2020). Impacts of an earn-a-bike cycling education program on children's time spent cycling. Journal of Transport & Health, 16, 100826. 10.1016/j.jth.2020.100826

Chaix B, Kestens Y, Duncan S, Merrien C, Thierry B, Pannier B, Brondeel R, Lewin A, Karusisi N, & Perchoux C (2014). Active transportation and public transportation use to achieve physical activity recommendations? A combined GPS, accelerometer, and mobility survey study. The International Journal of Behavioral Nutrition and Physical Activity, 11(1), 124. 10.1186/s12966-014-0124-x [PubMed: 25260793]

Cho G-H, Rodriguez DA, & Evenson KR (2011). Identifying walking trips using GPS data. Medicine & Science in Sports & Exercise, 43(2), 365–372. 10.1249/MSS.0b013e3181ebec3c [PubMed: 20581721]

Choi L, Liu Z, Matthews CE, & Buchowski MS (2011). Validation of accelerometer wear and nonwear time classification algorithm. Medicine & Science in Sports & Exercise, 43(2), 357–364. 10.1249/MSS.0b013e3181ed61a3 [PubMed: 20581716]

Choi L, Ward SC, Schnelle JF, & Buchowski MS (2012). Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Medicine & Science in Sports & Exercise, 44(10), 2009. 10.1249/MSS.0b013e318258cb36 [PubMed: 22525772]

Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, & Sallis JF (2003). International physical activity questionnaire: 12-country reliability and validity. Medicine & Science in Sports & Exercise, 35(8), 1381–1395. 10.1249/01.MSS.0000078924.61453.FB [PubMed: 12900694]

Dorn D, Gorzelitz J, Gangnon R, Bell D, Koltyn K, & Cadmus-Bertram L (2019). Automatic identification of physical activity type and duration by wearable activity trackers: A validation study. JMIR mHealth and uHealth, 7(5), e13547. 10.2196/13547 [PubMed: 31124470]

Duncan MJ, Badland HM, & Mummery WK (2009). Applying GPS to enhance understanding of transport-related physical activity. Journal of Science and Medicine in Sport, 12(5), 549–556. 10.1016/j.jsams.2008.10.010 [PubMed: 19237315]

Ellis K, Godbole S, Marshall S, Lanckriet G, Staudenmayer J, & Kerr J (2014). Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. Frontiers in Public Health, 2, 36. 10.3389/fpubh.2014.00036 [PubMed: 24795875]

Evenson KR, Goto MM, & Furberg RD (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. The International Journal of Behavioral Nutrition and Physical Activity, 12(1), 1–22. [PubMed: 25592201]

Ferguson T, Rowlands AV, Olds T, & Maher C (2015). The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: A cross-sectional study. The International Journal of Behavioral Nutrition and Physical Activity, 12(1), 1–9. 10.1186/s12966-015-0201-9 [PubMed: 25592201]

FitBit. (2020a). Help. https://myhelp.fitbit.com/s/home?language=en_US

FitBit. (2020b). SmartTrack. https://www.fitbit.com/eu/smarttrack#how-it-works

Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, Buote R, Van Heerden D, Luan H, & Cullen K (2020). Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: Systematic review. JMIR mHealth and uHealth, 8(9), e18694. 10.2196/18694 [PubMed: 32897239]

Garmin. (2020a). Instinct owner's manual. https://www8.garmin.com/manuals/webhelp/instinct/EN-US/GUID-CDDD370D-57F3-4D56-9405-9E508FA82E0E.html

Garmin. (2020b). Support center. https://support.garmin.com/en-US/?faq=zgFpy8MShkArqAxCug5wC6

Giles-Corti B, Vernez-Moudon A, Reis R, Turrell G, Dannenberg AL, Badland H, Foster S, Lowe M, Sallis JF, & Stevenson M (2016). City planning and population health: A global challenge. The Lancet, 388(10062), 2912–2924. 10.1016/S0140-6736(16)30066-6

Global Advocacy for Physical Activity, & Advocacy Council of the International Society for Physical Activity Health. (2012). NCD prevention: Investments [corrected] that work for physical activity. British Journal of Sports Medicine, 46(10), 709–712. [PubMed: 22869788]

Gong H, Chen C, Bialostozky E, & Lawson CT (2012). A GPS/GIS method for travel mode detection in New York City. Computers, Environment and Urban Systems, 36(2), 131–139. 10.1016/j.compenvurbsys.2011.05.003

HABITUS. (2020). About. https://www.habitus.eu/about

Hurvitz PM, Moudon AV, Kang B, Saelens BE, & Duncan GE (2014). Emerging technologies for assessing physical activity behaviors in space and time. Frontiers in Public Health, 2, 2. 10.3389/fpubh.2014.00002 [PubMed: 24479113]

IBM Corp. (2016). IBM SPSS statistics for windows, Version 24.0.

James P, Weissman J, Wolf J, Mumford K, Contant CK, Hwang W-T, Taylor L, & Glanz K (2016). Comparing GPS, log, survey, and accelerometry to measure physical activity. American Journal of Health Behavior, 40(1), 123–131. 10.5993/AJHB.40.1.14 [PubMed: 26685821]

Jankowska MM, Schipperijn J, & Kerr J (2015). A framework for using GPS data in physical activity and sedentary behavior studies. Exercise and Sport Sciences Reviews, 43(1), 48. 10.1249/JES.0000000000000035 [PubMed: 25390297]

Kang B, Moudon AV, Hurvitz PM, Reichley L, & Saelens BE (2013). Walking objectively measured: Classifying accelerometer data with GPS and travel diaries. Medicine & Science in Sports & Exercise, 45(7), 1419. 10.1249/MSS.0b013e318285f202 [PubMed: 23439414]

Kang M, Moudon AV, Hurvitz PM, & Saelens BE (2018). Capturing fine-scale travel behaviors: A comparative analysis between personal activity location measurement system (PALMS) and travel diary. International Journal of Health Geographics, 17(1), 1–14. 10.1186/s12942-018-0161-9 [PubMed: 29329535]

Kerr J, Duncan S, & Schipperijn J (2011). Using Global positioning systems in health research: A practical approach to data collection and processing. American Journal of Preventive Medicine, 41(5), 532–540. 10.1016/j.amepre.2011.07.017 [PubMed: 22011426]

King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, Banerjee B, Robinson TN, & Cirimele JJ (2016). Effects of three motivationally targeted mobile device applications on initial physical activity and sedentary behavior change in midlife and older adults: A randomized trial. PLoS One, 11(6), e0156370. 10.1371/journal.pone.0156370 [PubMed: 27352250]

Lam MS, Godbole S, Chen J, Oliver M, Badland H, Marshall SJ, Kelly P, Foster C, Doherty A, & Kerr J (2013). Measuring time spent outdoors using a wearable camera and GPS. Proceedings of the 4th International SenseCam & Pervasive Imaging Conference, San Diego, CA.

Landis JR, & Koch GG (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. 10.2307/2529310 [PubMed: 843571]

O'Driscoll R, Turicchi J, Hopkins M, Gibbons C, Larsen S, Palmeira A, Heitmann B, Horgan G, Finlayson G, & Stubbs R (2020). The validity of two widely used commercial and research-grade activity monitors, during resting, household and activity behaviours. Health and Technology, 10(3), 637–648. 10.1007/s12553-019-00392-7

Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, & Mutrie N (2007). Interventions to promote walking: Systematic review. The BMJ, 334(7605), 1204. 10.1136/bmj.39198.722720.BE [PubMed: 17540909]

Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, & Tremblay M (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. The International Journal of Behavioral Nutrition and Physical Activity, 5(1), 56–24. 10.1186/1479-5868-5-56 [PubMed: 18990237]

Raza W, Krachler B, Forsberg B, & Sommar JN (2020). Health benefits of leisure time and commuting physical activity: A meta-analysis of effects on morbidity. Journal of Transport & Health, 18, Article 100873. 10.1016/j.jth.2020.100873

Saelens BE, Sallis JF, & Frank LD (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine, 25(2), 80–91. 10.1207/S15324796ABM2502_03 [PubMed: 12704009]

Sahlqvist S, Song Y, & Ogilvie D (2012). Is active travel associated with greater physical activity? The contribution of commuting and non-commuting active travel to total physical activity in adults. Preventive Medicine, 55(3), 206–211. 10.1016/j.ypmed.2012.06.028 [PubMed: 22796629]

Sallis JF, Bull F, Burdett R, Frank LD, Griffiths P, Giles-Corti B, & Stevenson M (2016). Use of science to guide city planning policy and practice: How to achieve healthy and sustainable future cities. The Lancet, 388(10062), 2936–2947. 10.1016/S0140-6736(16)30068-X

Saunders LE, Green JM, Petticrew MP, Steinbach R, & Roberts H (2013). What are the health benefits of active travel? A systematic review of trials and cohort studies. PLoS One, 8(8), e69912. 10.1371/journal.pone.0069912 [PubMed: 23967064]

Schipperijn J, Kerr J, Duncan S, Madsen T, Klinker CD, & Troelsen J (2014). Dynamic accuracy of GPS receivers for use in health research: A novel method to assess GPS accuracy in real-world settings. Frontiers in Public Health, 2, 21. 10.3389/fpubh.2014.00021 [PubMed: 24653984]

Schoeppe S, Duncan MJ, Badland H, Oliver M, & Curtis C (2013). Associations of children's independent mobility and active travel with physical activity, sedentary behaviour and weight status: A systematic review. Journal of Science and Medicine in Sport, 16(4), 312–319. 10.1016/j.jsams.2012.11.001 [PubMed: 23219100]

The Community Guide. (2021). Combined built environment features help communities get active. https://www.thecommunityguide.org/content/combined-built-environment-features-help-communities-get-active

Troped PJ, Oliveira MS, Matthews CE, Cromley EK, Melly SJ, & Craig BA (2008). Prediction of activity mode with global positioning system and accelerometer data. Medicine & Science in Sports & Exercise, 40(5), 972–978. 10.1249/MSS.0b013e318164c407 [PubMed: 18408598]

U.S. Department of Health and Human Services. (2021). Physical activity. Office of Diseases Prevention and Health Promotion. https://health.gov/healthypeople/objectives-and-data/browse-objectives/physical-activity

Wang Y, König LM, & Reiterer H (2021). A Smartphone App to Support Sedentary Behavior Change by Visualizing Personal Mobility Patterns and Action Planning (SedVis): Development and Pilot Study. JMIR Formative Research, 5(1), e15369. 10.2196/15369 [PubMed: 33502322]

Welk GJ, Corbin CB, & Dale D (2000). Measurement issues in the assessment of physical activity in children. Research Quarterly for Exercise and Sport, 71(Suppl. 2), S59–S73. 10.1080/02701367.2000.11082788 [PubMed: 10925827]

World Health Organization. (2019). Global action plan on physical activity 2018–2030: More active people for a healthier world.

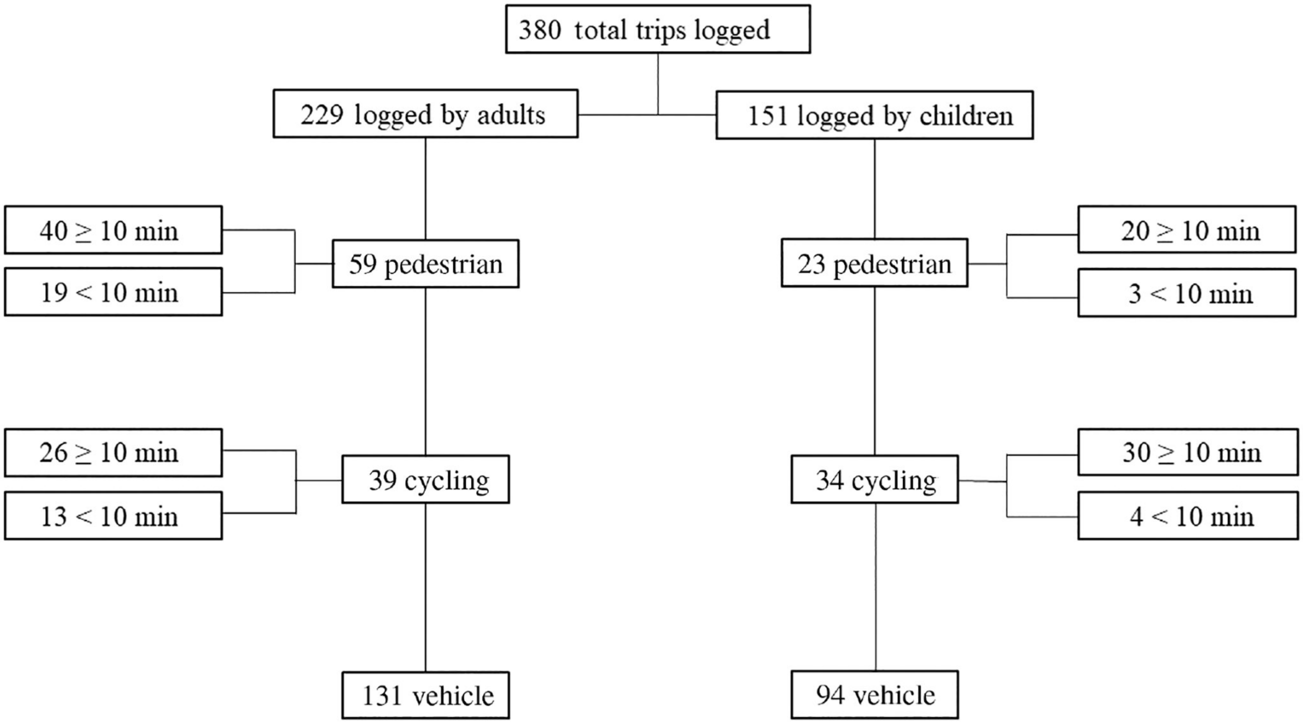

Figure 1 —.
Description of trips logged.

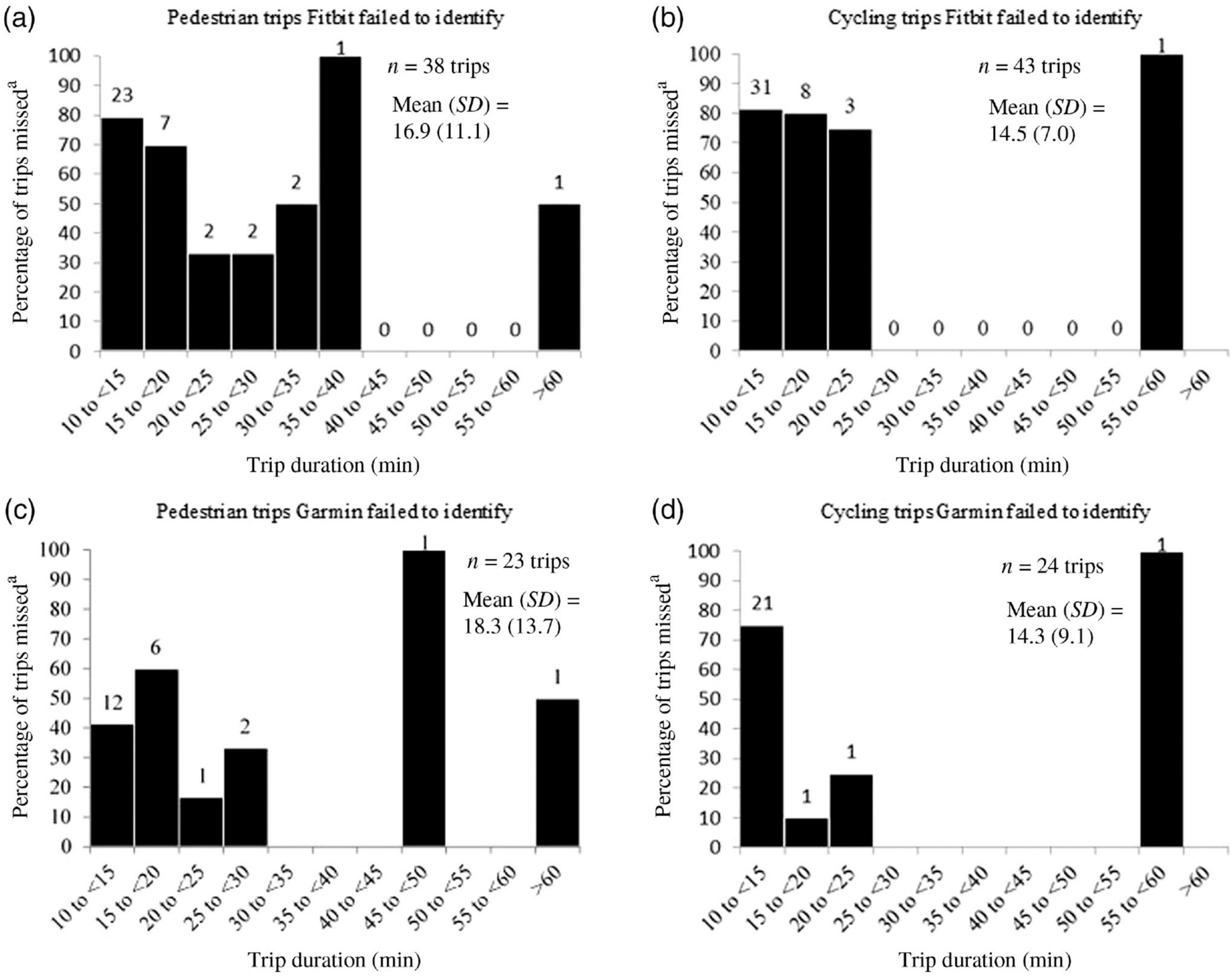

Figure 2 —.

Percentage of trips of different lengths failed to be identified by the Fitbit and Garmin devices across children and adults. [a]Logged pedestrian and cycling trips with a duration <10 min were excluded from analysis due to manufacturer's minimum detection threshold. Values above each bar represent the number of missed trips within the given trip duration range. The percentage of trips missed was calculated as the number of missed trips within the given duration range divided by the total number of logged trips within the given duration range. Within each chart, the *n* reflects the total number of missed trips across all trip duration ranges, and the overall mean and *SD* duration of these missed trips are provided (in minutes per day).

Author Manuscript
Author Manuscript
Author Manuscript

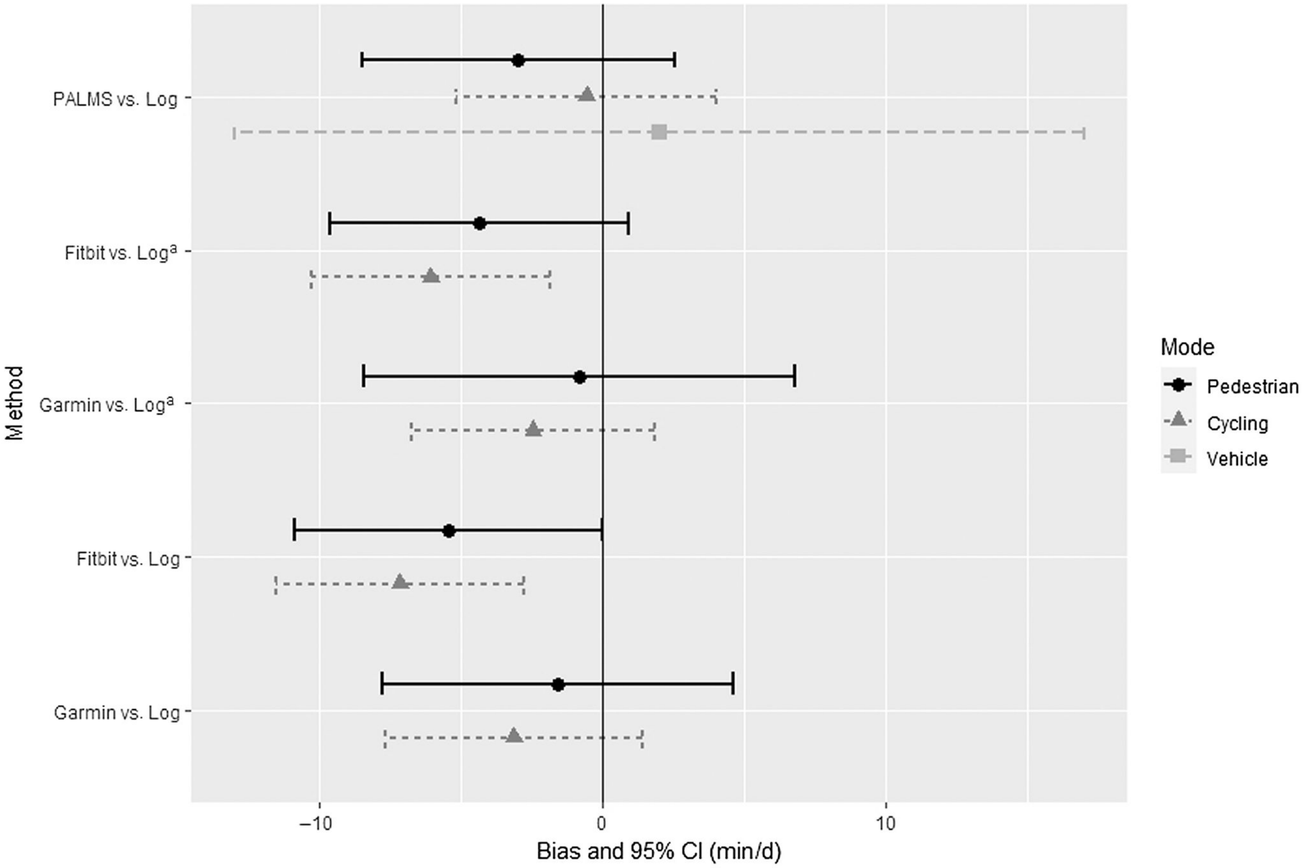

Figure 3 —.

Mean bias and 95% CIs from mixed-effects models. [a]Logged pedestrian and cycling trips with a duration <10 min were excluded from analysis due to manufacturer's minimum detection threshold. The log was the referent measure, so negative values reflect an underestimation by the test measure. The mean difference between measurement methods was similar between children and adults as indicated by no significant age Group × Measurement method interactions ($p < .05$).

Table 1

Trip-Level Agreement Between Each Test Measure and the Trip Logs for Identifying Trips and Classifying Trip Mode in Children and Adults (N = 380 Trips)

Sample and test measure	Trip log classification			
All participants				
PALMS classification	No trip	Pedestrian (n = 82)	Cycling (n = 73)	Vehicle (n = 225)
No trip	—	7 (8.5%)	0 (0%)	1 (0.4%)
Pedestrian	0	62 (75.6%)	3 (4.1%)	0 (0.0%)
Cycling	1	11 (13.4%)	69 (94.5%)	6 (2.7%)
Vehicle	0	2 (2.4%)	1 (1.4%)	218 (96.9%)
F1		0.84	0.87	0.98
Fitbit[a] classification	No trip	Pedestrian (n = 60)	Cycling (n = 56)	—
No trip	—	38 (63.3%)	43 (76.8%)	—
Pedestrian	7	22 (36.7%)	0 (0%)	—
Cycling	0	0 (0%)	13 (23.2%)	—
F1		0.49	0.38	—
Garmin[a] classification	No trip	Pedestrian (n = 60)	Cycling (n = 56)	—
No trip	—	23 (38.3%)	24 (42.9%)	—
Pedestrian	12	37 (61.7%)	0 (0%)	—
Cycling	4	0 (0%)	32 (57.1%)	—
F1		0.68	0.70	—
Fitbit classification	No trip	Pedestrian (n = 82)	Cycling (n = 73)	—
No trip	—	60 (73.2%)	60 (82.2%)	—
Pedestrian	7	22 (26.8%)	0 (0%)	—
Cycling	0	0 (0%)	13 (17.8%)	—
F1		0.40	0.30	—
Garmin classification	No trip	Pedestrian (n = 82)	Cycling (n = 73)	—
No trip	—	44 (53.7%)	41 (56.2%)	—
Pedestrian	11	38 (46.3%)	0 (0.0%)	—
Cycling	4	0 (0.0%)	32(43.8%)	—
F1		0.58	0.59	—

Sample and test measure	Trip log classification			
Adults				
PALMS classification	No trip	Pedestrian (n = 59)	Cycling (n = 39)	Vehicle (n = 131)
No trip	—	7 (11.9%)	0 (0%)	1 (0.8%)
Pedestrian	0	44 (74.6%)	2 (5.1%)	0 (0.0%)
Cycling	1	7 (11.9%)	37 (94.9%)	3 (2.3%)
Vehicle	0	1 (1.7%)	0 (0%)	127 (96.9%)
F1		0.83	0.85	0.98
Fitbit[a] classification	No trip	Pedestrian (n = 40)	Cycling (n = 26)	—
No trip	—	24 (60.0%)	19 (73.1%)	—
Pedestrian	6	16 (40.0%)	0 (0%)	—
Cycling	0	0 (0%)	7 (27.9%)	—
F1		0.52	0.42	—
Garmin[a] classification	No trip	Pedestrian (n = 40)	Cycling (n = 26)	—
No trip	—	10 (25.0%)	7 (26.9%)	—
Pedestrian	10	30 (75.0%)	0 (0%)	—
Cycling	2	0 (0%)	19 (73.1%)	—
F1		0.75	0.81	—
Fitbit classification	No trip	Pedestrian (n = 59)	Cycling (n = 39)	—
No trip	—	43 (72.9%)	32 (82.1%)	—
Pedestrian	6	16 (27.1%)	0 (0%)	—
Cycling	0	0 (0%)	7 (17.9%)	—
F1		0.40	0.30	—
Garmin classification	No trip	Pedestrian (n = 59)	Cycling (n = 39)	—
No trip	—	28 (47.5%)	20 (51.3%)	—
Pedestrian	9	31 (52.5%)	0 (0%)	—
Cycling	2	0 (0%)	19 (48.7%)	—
F1		0.63	0.63	—
Children				
PALMS classification	No trip	Pedestrian (n = 23)	Cycling (n = 34)	Vehicle (n = 94)
No trip	—	0 (0%)	0 (0%)	0 (0.0%)
Pedestrian	0	18 (78.3%)	1 (2.9%)	0 (0.0%)

Sample and test measure	Trip log classification			
Cycling	0	4 (17.4%)	32 (94.1%)	3 (3.2%)
Vehicle	0	1 (4.3%)	1 (2.9%)	91 (96.8%)
F1		0.86	0.88	0.97
Fitbit[a] classification	No trip	Pedestrian ($n = 20$)	Cycling ($n = 30$)	—
No trip	—	14 (70.0%)	24 (80.0%)	—
Pedestrian	1	6 (30.0%)	0 (0%)	—
Cycling	0	0 (0%)	6 (20.0%)	—
F1		0.44	0.33	—
Garmin[a] classification	No trip	Pedestrian ($n = 20$)	Cycling ($n = 30$)	—
No trip	—	13 (65.0%)	17 (56.7%)	—
Pedestrian	2	7 (35.0%)	0 (0%)	—
Cycling	2	0 (0%)	13 (43.3%)	—
F1		0.48	0.58	—
Fitbit classification	No trip	Pedestrian ($n = 23$)	Cycling ($n = 34$)	—
No trip	—	17 (73.9%)	27 (79.4%)	—
Pedestrian	1	6 (26.1%)	0 (0%)	—
Cycling	0	0 (0%)	7 (20.6%)	—
F1		0.40	0.34	—
Garmin classification	No trip	Pedestrian ($n = 23$)	Cycling ($n = 34$)	—
No trip	—	16 (69.6%)	19 (55.9%)	—
Pedestrian	2	7 (30.4%)	0 (0%)	—
Cycling	2	0 (0%)	15 (44.1%)	—
F1		0.44	0.59	—

Note. The percentages in each column total to 100% within each test measure. PALMS = Personal Activity Location Measurement System.

[a]Logged pedestrian and cycling trips with a duration <10 min were excluded from analysis due to minimum detection threshold used by the Fitbit and Garmin classification features.

Table 2

Day-Level Agreement Between Each Test Measure and the Trip Logs for Estimating Time Spent in Each Trip Mode in Children and Adults (*N* = 89 Days)

	Mean (*SD*), min/day			
Measure comparison and trip mode	**Log**	**Test measure**	**Mean absolute error, min/day (percentage of log mean)**	r_s
Log vs. PALMS				
Pedestrian	15.19 (20.48)	12.22 (18.19)	6.16 (40.6%)	.806
Cycling	11.55 (18.08)	10.98 (17.57)	2.42 (21.0%)	.847
Vehicle	58.56 (60.54)	60.56 (61.97)	7.60 (13.0%)	.936
Log vs. Fitbit[a]				
Pedestrian	13.92 (19.66)	9.58 (18.63)	6.13 (44.0%)	.707
Cycling	10.39 (17.18)	4.32 (15.19)	7.87 (75.7%)	.403
Log vs. Garmin[a]				
Pedestrian	13.92 (19.66)	13.11 (25.37)	8.72 (62.6%)	.683
Cycling	10.39 (17.18)	7.93 (16.51)	5.39 (51.9%)	.703
Log vs. Fitbit				
Pedestrian	15.19 (20.48)	9.76 (18.61)	7.22 (47.5%)	.694
Cycling	11.55 (18.08)	4.38 (15.22)	8.97 (77.7%)	.407
Log vs. Garmin				
Pedestrian	15.19 (20.48)	13.62 (25.65)	9.30 (61.2%)	.687
Cycling	11.55 (18.08)	8.40 (17.24)	5.98 (51.8%)	.704

Note. PALMS = Personal Activity Location Measurement System.

[a] Logged pedestrian and cycling trips with a duration <10 min were excluded from analysis due to minimum detection threshold used by the Fitbit and Garmin classification features.