

TRANSPORT FINDINGS

Validity, Reliability, and Usability of a Smartphone App to Measure Bicycling Location

Anna Porter , Kelly Evenson , Greg Griffin 

Keywords: Strava, GPS, fitness tracking, crowdsourced data, cycling, physical activity

<https://doi.org/10.32866/001c.57773>

Findings

This study evaluates concurrent validity, test-retest reliability, and usability of the Strava smartphone app for measuring bicycling locations in urban and rural field tests. Strava location data were inside an 11-meter buffer on average 64% of the time compared to Qstarz' 52%, over 100 evaluations (n participants=73). Most participants agreed or strongly agreed that the Strava app was useful (83%) and that they would prefer to use a smartphone app to track their bicycling (42%). Results indicate that the Strava app is reliable and valid for measuring bicycling locations in these field tests.

1. QUESTIONS

Global positioning systems (GPS) devices have been explored to measure bicycling, but substantial participant buy-in has historically been required – participants must remember to carry the device with them at all times, keep the battery charged, and may have to turn it on to record location. Smartphone apps like Strava offer a user-friendly alternative. Researchers and practitioners use crowdsourced bicycle trip data from Strava and other organizations to evaluate transportation outcomes (*e.g.*, Brum-Bastos et al. 2019; Ferster et al. 2021), so it is critical to ensure the validity and reliability of these measurement tools for bicycling. Strava processes the GPS data collected by the smartphone hardware to output trip information and makes de-identified data available as Strava Metro, with at least 125 transportation organizations using the service (Schneider 2017). Most studies aggregate Strava bicycle volumes at the block- or neighborhood-level geographies (Dadashova and Griffin 2020; Fischer, Nelson, and Winters 2020) but lack evaluation of spatial and usability qualities. We seek to answer the following question: what is the concurrent validity and test-retest reliability of the Strava smartphone app for measuring bicycling location in urban and rural environments, and what is the perceived usability of smartphone apps for measuring bicycling?

2. METHODS

Two field courses were created for bicycling—one in a small metropolitan setting (Ingram and Franco 2014) (termed “urban”, 3.3 km in length) with buildings and other sources of possible GPS interference shown in [Figure 1](#), and one in a “rural” setting (3.6 km) with little GPS interference. Towards the end of the data collection period in 2020, construction began on a portion of the urban route, necessitating an alternative course for the remaining rides (3.7 km). Individuals were recruited via university communications, social media, and word of mouth. Participants were required to be 18 years of age or older, have access to a bicycle they could ride safely, ride a bicycle at least once a

month, have a bicycle helmet that they were willing to wear, have a smartphone, and be willing to use the Strava app during the study visit. Individuals were invited to participate twice, either once per location or twice in the same location. This study was approved by the University Institutional Review Board (IRB-19-193).

The Qstarz BT-Q100XT data logger was used to assess the validity of the GPS data reported from the Strava app. The Qstarz has previously been shown to be valid in static (Duncan et al. 2013) and real-world settings (Schipperijn et al. 2014) and is frequently used for health research (Schipperijn et al. 2014). Research personnel bicycled with participants and facilitated participants' recording both Qstarz and Strava data logs trips, starting and ending at the red star mapped in [Figure 1](#).

The two study routes were created from shapefiles provided by the university (urban route) and the Rails to Trails conservancy TrailLink resource (rural route, data from www.traillink.com). These files were used to represent the true route. An 11 m buffer was created around the road network for each course to account for two standard 3.6 m road lanes (FHWA 2014) and an additional 3.6 m of error to accommodate a previously reported median error of the Qstarz GPS of 2.9 m (Schipperijn et al. 2014).

To assess the reliability of the Strava app and the Qstarz GPS, the proportion of location data within the buffer from experimental rides 1 and 2 were compared for each device using Spearman rank correlation and paired samples t-tests. To assess the difference in reliability between each device, the mean difference in the proportion of location points within the buffer between rides 1 and 2 was calculated for each device. Multivariable linear regression models were used to assess whether cloud cover, urbanicity, or phone make were associated with the Qstarz and Strava proportion in the buffer. Study rides with missing data due to user or device error were excluded. Researchers emailed participants a link to a short survey asking about their use of the Strava app approximately two weeks after the first data collection.

3. FINDINGS

A total of 73 unique participants were screened for participation in the study, of which 61 (84%) completed at least one study ride. There were 100 unique study visits, 41 of them taking place in the urban location and 59 taking place in the rural location. Sixty-nine visits were recorded via iPhone, and 31 on Android. Just over half took place in cloudy weather (54%). Strava location data were inside the buffer on average 64% of the time, as compared to Qstarz 52% of the time, as shown in [Table 1](#).

In the urban context, the difference between Qstarz and Strava data accuracy was insignificant. However, in the rural location, Strava location data was inside the buffer 56% of the time, as compared to 40% of the time for Qstarz. Trip time was also examined for concurrent validity, comparing the time

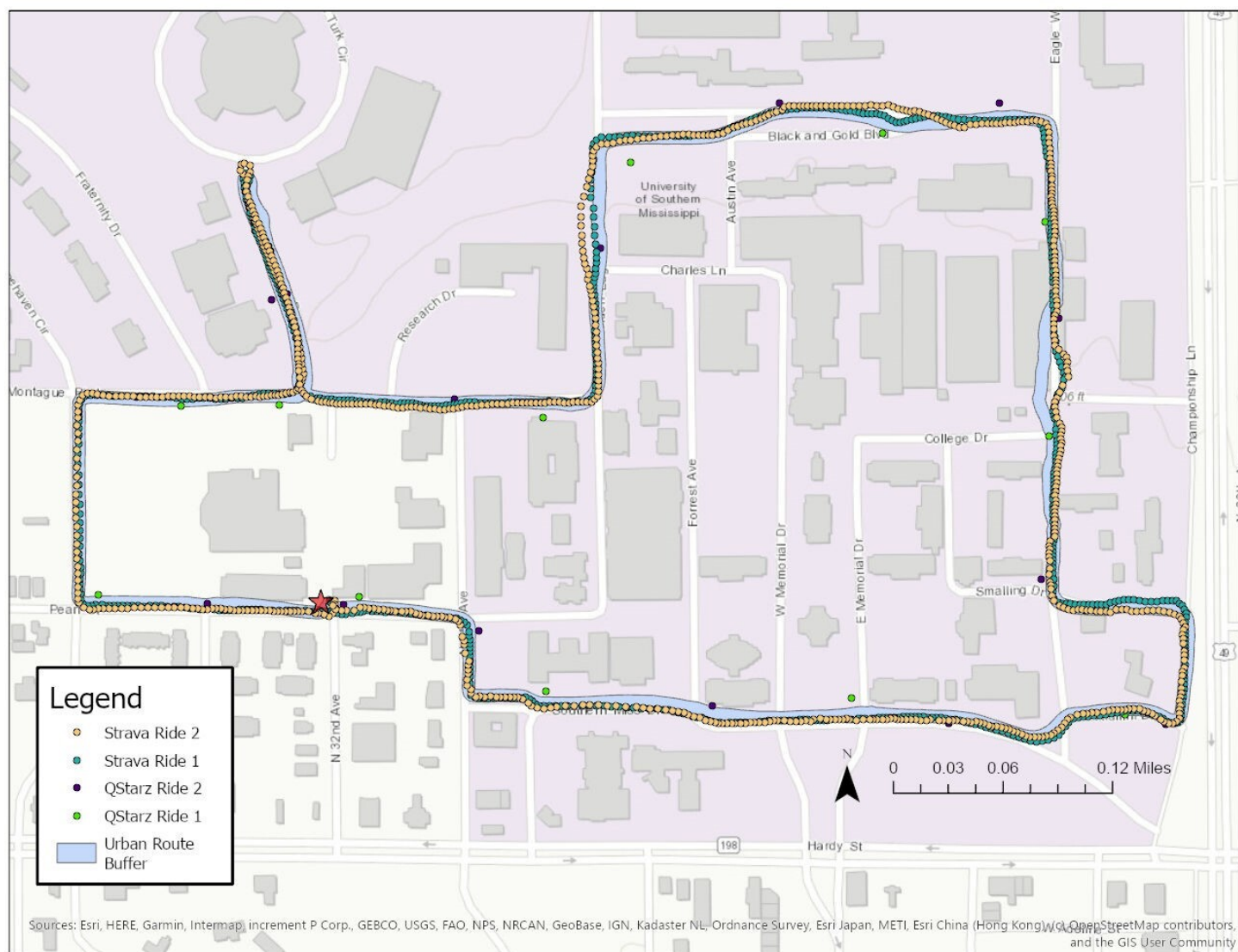


Figure 1. Example of data collected from a study ride on the urban route. The star indicates the start and end of the route.

Table 1. Concurrent validity results comparing location data accuracy of Strava and Qstarz GPS Device

	n	Mean ^a (SD)	95% CI	p-value
<i>Urban and Rural combined</i>				
Qstarz	194	0.52 (0.30)		
Strava	199	0.64 (0.26)		
Difference		-0.12	-0.17, -0.06	<0.001
<i>Urban</i>				
Qstarz	78	0.69 (0.23)		
Strava	81	0.75 (0.15)		
Difference		-0.06	-0.12, 0.0003	0.051
<i>Rural</i>				
Qstarz	116	0.40 (0.28)		
Strava	118	0.56 (0.29)		
Difference		-.15	-0.22, -0.08	<0.001

^a Mean proportion of points within the buffer.

recorded by Strava ($M=12.4$ min., $SD=2.1$ min.) to the time recorded by the research assistant ($M=12.7$ min., $SD=2.3$ min.), and the difference was not statistically significant for the total sample, or when stratified by urbanicity.

Test-retest reliability results are reported in [Table 2](#), showing a moderate to strong correlation across study rides for both Qstarz and Strava's proportion of location data within the buffer. For the total sample, both Qstarz and Strava did not produce significantly different results in the proportion of location data that fell within the buffer when comparing the first and second rides. Reliability results were similar for the rural location, but Strava produced significantly different results when comparing the two rides for the urban location, with the first ride having more location data within buffer as compared to the second.

A limited sample ($n=36$, 59%) completed the post-ride user survey. Most participants (83%) agreed or strongly agreed that the Strava app was useful for measuring bicycling. In a research study scenario, 64% preferred the method allowed by Strava in which they start and stop the app to record their bicycling activity rather than automatic tracking (Table S1).

As compared to the Qstarz BT-Q100XT data logger, the Strava app was a reliable and valid tool for measuring bicycling location in our study. Location data reported by the Strava app performed as well or better than the criterion in validity tests, yet the location data collected fell within the designated buffer on average only 64% of the time. This indicates that the location data is not within our set margin of error 37% of the time, which could be considered a substantial error based on the size of the buffer, 11 meters (36 feet). Larger buffer sizes were not tested in this study. Participants considered Strava a useful tool and acceptable for use in a research study. The spatial accuracy of Strava datasets is useful for research and practice on specific transportation and recreation corridors but not sufficient to evaluate rider placement within the

Table 2. Test re-test reliability results comparing location data accuracy of Strava and Qstarz GPS Device

Measurement Tool	n	Mean ^a (SD) Ride 1	Mean (SD) Ride 2	Spearman's rho	Mean Difference	95% CI	p- value
<i>Urban and Rural combined</i>							
Qstarz	97	0.52 (0.03)	0.52 (0.03)	0.68	-0.004	-0.052, 0.043	0.86
Strava	99	0.63 (0.03)	0.64 (0.03)	0.69	-0.006	-0.046, 0.035	0.79
Difference in reliability of Qstarz vs. Strava	196			0.06	-0.001	-0.026, 0.036	0.97
<i>Urban</i>							
Qstarz	39	0.72 (0.22)	0.67 (0.24)	0.80	0.046	-0.012, 0.104	0.12
Strava	40	0.77 (0.15)	0.74 (0.15)	0.74	0.036	0.005, 0.066	0.03
Difference in reliability of Qstarz vs. Strava	79			0.06	-0.01	-0.075, 0.054	0.75
<i>Rural</i>							
Qstarz	58	0.38 (0.28)	0.42 (0.28)	0.56	-0.038	-0.107, 0.031	0.26
Strava	59	0.54 (0.29)	0.57 (0.29)	0.64	-0.033	-0.098, 0.031	0.31
Difference in reliability of Qstarz vs. Strava	117			0.13	0.005	-0.089, 0.098	0.92

^a Mean proportion of points within the buffer.

roadway or adjacent bikeway. Based on technology circa 2020, evaluations of rider placement on a facility should consider survey-grade GPS, telemetry, or videography.

ACKNOWLEDGMENTS

This work was supported by the Collaborative Sciences Center for Road Safety (<https://www.roadsafety.unc.edu>), a United States Department of Transportation National University Transportation Center award no. 69A3551747113 and the UNC Injury Prevention Research Center (#R49/CE0042479) from the Centers for Disease Control and Prevention (KRE); The University of Southern Mississippi (AKP); and a Lutcher Brown Fellowship from The University of Texas at San Antonio (GPG). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Submitted: October 01, 2022 AEDT, Accepted: December 22, 2022 AEDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for more information.

REFERENCES

- Brum-Bastos, Vanessa, Colin J. Ferster, Trisalyn Nelson, and Meghan Winters. 2019. "Where to Put Bike Counters? Stratifying Bicycling Patterns in the City Using Crowdsourced Data." *Transport Findings*. <https://doi.org/10.32866/10828>.
- Dadashova, Bahar, and Greg P. Griffin. 2020. "Random Parameter Models for Estimating Statewide Daily Bicycle Counts Using Crowdsourced Data." *Transportation Research Part D: Transport and Environment* 84: 102368. <https://doi.org/10.1016/j.trd.2020.102368>.
- Duncan, Scott, Tom I. Stewart, Melody Oliver, Suzanne Mavoa, Deborah MacRae, Hannah M. Badland, and Mitch J. Duncan. 2013. "Portable Global Positioning System Receivers." *American Journal of Preventive Medicine* 44 (2): e19–29. <https://doi.org/10.1016/j.amepre.2012.10.013>.
- Ferster, Colin, Trisalyn Nelson, Karen Laberee, and Meghan Winters. 2021. "Mapping Bicycling Exposure and Safety Risk Using Strava Metro." *Applied Geography* 127 (January): 102388. <https://doi.org/10.1016/j.apgeog.2021.102388>.
- FHWA. 2014. "Lane Width." https://safety.fhwa.dot.gov/geometric/pubs/mitigationstrategies/chapter3/3_lanewidth.cfm.
- Fischer, Jaimy, Trisalyn Nelson, and Meghan Winters. 2020. "Comparing Spatial Associations of Commuting versus Recreational Ridership Captured by the Strava Fitness App." *Findings*. <https://doi.org/10.32866/001c.16710>.
- Ingram, Deborah D., and Sheila J. Franco. 2014. "2013 NCHS Urban-Rural Classification Scheme for Counties." *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research* 166: 1–73.
- Schipperijn, Jasper, Jacqueline Kerr, Scott Duncan, Thomas Madsen, Charlotte Demant Klinker, and Jens Troelsen. 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. <https://doi.org/10.3389/fpubh.2014.00021>.
- Schneider, Benjamin. 2017. "The Social Network That Helps Planners Understand Pedestrians and Cyclists." Bloomberg City Lab. <https://www.bloomberg.com/news/articles/2017-11-07/how-cities-use-strava-to-plan-for-pedestrians-and-bikes>.

SUPPLEMENTARY MATERIALS

Supplemental Information

Download: <https://findingspress.org/article/57773-validity-reliability-and-usability-of-a-smartphone-app-to-measure-bicycling-location/attachment/123878.pdf>
