The Journal of Strength and Conditioning Research™

The Validity and Reliability of Global Positioning System Units for Measuring Distance and Velocity During Linear and Team Sport Simulated Movements

Robert A. Huggins,[1] Gabrielle E.W. Giersch,[1] Luke N. Belval,[1] Courteney L. Benjamin,[1] Ryan M. Curtis,[1] Yasuki Sekiguchi,[1] Jussi Peltonen,[2] and Douglas J. Casa[1]

[1]Department of Kinesiology, Korey Stringer Institute, University of Connecticut, Storrs, Connecticut; and [2]Polar Electro Oy, Jyvaskyla, Finland

Abstract

Huggins, RA, Giersch, GEW, Belval, LN, Benjamin, CL, Curtis, RM, Sekiguchi, Y, Peltonen, J, and Casa, DJ. The validity and reliability of GPS units for measuring distance and velocity during linear and team sport simulated movements. *J Strength Cond Res* 34(11): 3070–3077, 2020—This experimental study aimed to assess the validity and reliability of shirt-mounted 10-Hz global positioning system (GPS) units (Polar Team Pro) for measuring total distance (TD), constant velocity (Vel_C), and instantaneous velocity (Vel_I) during linear running and a team sport simulation circuit (TSSC). Fifteen male soccer athletes completed linear tasks (40 and 100 m) at various velocities: walk (W) (4.8–7.9 km·h^{-1}), jog (J) (8.0–12.7 km·h^{-1}), run (R) (12.9–19.9 km·h^{-1}), and sprint (S) (>20.0 km·h^{-1}) and a 120-m TSSC. Global positioning system validity and reliability for TD, Vel_C, and Vel_I were compared with criterion measures using 2 methods (a and b) of GPS raw data extraction. When measuring TD for the Polar Team Pro device, validity and reliability measures were <5% error at all velocities during the 40-m (with the exception of the S [%CV = 8.03]) and 100-m linear trial (both extraction methods) and TSSC. The GPS mean difference (±*SD*) for TD during the TSSC using extraction methods (a) and (b) was 0.2 ± 1.2 and 2.2 ± 2.2 m, respectively. The validity of the device in measuring Vel_C was significantly different ($p < 0.05$) at all velocities during the 40 m (exception W) and the 100 m, with effect sizes ranging from trivial to small (exception of 100 m S). Vel_I was similar ($p > 0.05$) at all velocities, except for the W ($p = 0.001$). The reliability of the device when measuring Vel_C during the 40 and 100 m was <5% CV; however, during the 100 m, Vel_I ranged from 1.4 to 12.9%. Despite trivial to large effect sizes for validity of TD, this device demonstrated good reliability <5% CV during linear and TSSC movements. Similarly, effect sizes ranged from trivial to large for $Vel_{C,}$ and yet Vel_I reliability was good for Vel_C, but good to poor for Vel_I.

Key Words: performance analysis, monitoring, soccer, running, reproducibility of results

Introduction

Monitoring the movement demands associated with team sport athletes during match play and training sessions has become increasingly more common in sport, especially at the professional and collegiate levels (2,8,9,13,15). However; this was not the case until the late 2000s, following the release of the global positioning system (GPS) to the public by the US Department of Defense. At this time, the potential for GPS instruments in sport were truly realized and quickly became mainstream in sport. The advent of GPS instrumentation to the world of sport allowed exercise scientists and coaches to easily quantify individualized information from their players such as velocity zones (VZ), distances, accelerations, decelerations, and measurements related to training load (9,14,27–29). Since then, the monitoring and evaluation of these metrics has made a large impact on both the acute and chronic program variables and the overall periodization of teams and individual athletes' training (7,11,24). Not only did the ease of data acquisition allow for more real-time feedback but also it required less from the individual(s) collecting and managing the data. The ability for the individual(s) to implement a system with an entire team or squad while simultaneously ensuring (a) high-quality and coordinated data acquisition, (b) live "tagging or marking" of particular events within a training session, and (c) coordinated commencement or termination of the data collection period was a large advancement from previous methodologies. Additionally, in response to the enhanced practicality that GPS and external load monitoring provide, an entire industry with roots in data analytics and sport performance monitoring has developed, and now teams are able to use these data to better understand the impact that load have on fatigue, injury prevention, load or volume management, and return to participation after injury. However, although the benefits of GPS seemingly outweigh the detrimental effects, in order for those conducting the monitoring to accurately interpret these data, they must also understand the limitations, in particular, those related to the validity and reliability of the instrument. Without an understanding and appreciation of these limitations, one cannot truly deduce meaningful interpretations.

An appreciation for the mechanisms behind GPS and its validity and reliability is required to understand the contextual limitations both outside and within sport. The location of a particular GPS receiver or device is triangulated from the

Address correspondence to Robert A. Huggins, Robert.huggins@uconn.edu.

Journal of Strength and Conditioning Research 34(11)/3070–3077

numerous satellites orbiting the earth that are equipped with onboard atomic clocks (25). A minimum of 4 satellites continuously sending and receiving information related to distance and time from the satellites are required to triangulate the position of the receiver or individual unit. That being said, there are limitations that interfere or disrupt the transmission of these signals such as tall structures and clouds within the atmosphere proximal to the recording area. Furthermore, increased error in the measurement of total distance (TD) has been reported when there is a reduction in the number of satellites available for triangulation of the GPS signal (16).

Within the context of sport, the validity and reliability of GPS units for measuring distance in team sports has been conducted previously in a variety of sports and at various levels of competition (9,21–23,28). Furthermore, the validity of 1- (9,21) 5- (2,8,9,12,15,21,23,28), 10- (4,22,28), 15- (22), and 18-Hz (3,20) GPS devices has been examined. Although data suggest that 10-, 15-, and 18-Hz units demonstrate improved measures of movement demands compared with 1 and 5-Hz units (3,20,22), other investigations have demonstrated that GPS is a valid and reliable measure of TD with <5% *SEE* and <5% coefficient of variation (CV); however, large variability in the reported validity and reliability exist between brands and models (5,9,22). Furthermore, previous investigations determined that validity and reliability reported from various GPS device models and manufacturers should not be extrapolated to others (1); thus, each device manufacturer using GPS requires independent validity and reliability testing. The validation process is a critical step in the application of this technology in sport and research settings to understand the limitations of the device.

Polar Team Pro units are currently used internationally in a variety of team sports and competitive levels, and these devices warrant validation. Therefore, the purpose of this investigation was to examine the validity and reliability of one such device manufacturer (Polar Team Pro GPS device) in the assessment of TD, instantaneous velocity (Vel_I), and constant or average velocity (Vel_C) during linear and sport-specific movement patterns. Furthermore, we examined the interunit reliability of the GPS devices in the measurement of TD. Specifically, we aim to determine the validity and reliability of a novel shirt-mounted GPS device for TD, Vel_I, and Vel_C compared with the criterion measure during linear tasks (40 and 100 m) at various velocities walk (W), jog (J), run (R), sprint (S), and during a team sports simulation circuit (TSSC). Additionally, from a data methodology perspective, we aim to analyze TD using 2 methods of data extraction to determine if differences exist between methods.

Methods

Experimental Approach to the Problem

To determine the ability of a 10-Hz GPS device of measuring TD, athletes completed linear tasks of 40 and 100 m at 4 different VZ for a total of 8 repetitions. The measurement recorded by the device was compared with the criterion measure (tape measure). To determine the ability of the device to measure both Vel_I and Vel_c during the same linear tasks, the measures were compared with the criterion: laser for Vel_I and timing gates for Vel_C. Finally, to determine the ability of the device to measure team sport movement demands, athletes performed 2 repetitions of a TSSC used previously by Coutts and Duffield (9), which was used to test the TD. To examine the various methodologies related to TD data extraction, we chose to compare 2 methods that have been reported. One used GPS-based initiation of movement and timing gate end of movement (method A), whereas the second method used the activation of the onset of first GPS movement and the end of GPS movement (method B). The rationale for reporting both was to elucidate the possibility of potential differences between these methods previously reported, which may reveal important differences when reporting GPS validity and reliability for future devices.

Subjects

Fifteen male club collegiate soccer athletes ([mean ± *SD*], 20 ± 1 years of age, 177.0 ± 7.5-cm tall, and 71.57 ± 7.17 kg of body mass) volunteered to participate in this study. Athletes donned a shirt with an integrated 10-Hz GPS and 200-Hz MEMs-enabled athlete tracking device (Polar Team Pro; Polar Electro, Kempele, Finland). All athletes gave written informed consent before participation in the study, and the University of Connecticut Institutional Review Board approved the study. The sponsor did not have the right to approve or disapprove publication of this article.

Procedures

All athletes were familiarized with the testing procedures by verbal instruction, demonstration by a researcher, and subsequent physical practice of the linear and TSSC task. Height was recorded to the nearest centimeter using a standard tape measure, whereas body mass was assessed using a digital scale (T51P; Ohaus, Pine Brook, NJ). Athletes were provided a the 10-Hz GPS device. Traditionally, this device is worn on a chest strap, and the device is located just below the xiphoid process; however, athletes wore a custom fit GPS shirt (Polar Team Pro Sensor and Shirt; Polar Electro, Kempele, Finland) designed to house the unit posteriorly between the first and fourth thoracic vertebrae above the scapula with heart rate leads embedded in the shirt rather than across the chest. The shirt was fitted per manufacturer's suggested instructions, ensuring a "tight fit" to ensure that the GPS sensor was snug on the upper back between the superior and medial portions of the scapulae. Given the previous research examining both chest and back worn GPS devices, it is not likely that the physical location of the GPS had any impact on the validity or reliability of the device; therefore, the Polar shirt was used rather than the chest strap. Upon arrival to the outdoor testing field, GPS signal acquisition was visually confirmed via mobile tablet with Polar software, which requires a minimum of 4 satellites for the athlete's unit number to appear on the tablet (dilution of precision is not reported by the device therefore not reported). Before the participation, athletes conducted a standardized 10-minute dynamic warm-up. Next, athletes performed at least 1 linear task at each VZ and a TSSC. This was done to orient them to the velocities at which they would be running for the linear task and the varying velocities within the TSSC. Verbal instruction and feedback were provided by the researchers on the linear, sport-specific movement tasks and VZ required during each bout. Before the start of each bout of activity at all velocities and tasks, athletes remained stationary with 1 foot on the starting line for a minimum of 5 seconds before and upon completion of each bout to allow for GPS and accelerometry data to equilibrate and to allow for more uniform extraction of data from the raw files upon export.

Linear Movement Task. Athletes performed a linear movement task consisting of 2 trials at 4 different velocity ranges for a total of 8 repetitions: walking (W) (4.8–8.0 $km \cdot h^{-1}$), jogging (J) (8.1–12.7 $km \cdot h^{-1}$), running (R) (12.8–20.1 $km \cdot h^{-1}$), and sprinting (S) (>20.1 $km \cdot h^{-1}$) for 100 m. Timing gates were stationed at 0, 40, and 100 m to assess 40- and 100-m differences. Total distance, Vel_C, and Vel_I recorded by the device were compared with criterion measure. The criterion for TD was standard tape measure. Criterion TD for method A (Figure 1A) was defined as the measured distance from the timing gate at the start line to timing gate at the finish line (100 m). For method B, it was the measured distance from timing gate start line to a cone placed precisely 1 and 5 m beyond the finish line. This was done to account for the deceleration during the W and J (TD = 101 m) and R and S (TD = 105 m) (Figure 1A). Vel_C was determined by calculating the average velocity of the selected segment based on the velocity reported from the device file. The criterion measure for Vel_C was obtained from the electronic timing gates (TC Timing System; Brower Timing Systems Inc., Draper, UT). The device Vel_I was compared with the criterion Vel_I as measured by a tripod mounted laser with swivel handle (UltraLyte 20 20 200 LR; Laser Technology Inc., Centennial, CO) placed 10 m beyond the 100-m finish. The validity and reliability of laser technology at various velocities and distances have been reported in previous studies (6,10,17,27,28). The laser continuously aimed on the athlete's trunk for the duration of the trial, and the peak velocity was recorded. Vel_I was assessed continuously during the entire 100-m task at each VZ and during the first 20 m of the TSSC. If an error occurred during a repetition, the athlete was asked to perform another attempt. Errors included a deviation from the course, interference with a cone, lack of standing still at the 2 required locations, and lack of achievement of the correct VZ.

Multidirectional Movement Task. Device validity and reliability during multidirectional movement was assessed using the TSSC, which has been previously studied by Johnston et al. (22) The total measured distance of the circuit was 120 m in length (as determined by a standard tape measure) and involved sprinting, fast running, jogging, walking, cutting, accelerating, decelerating, and standing still (Figure 1B). Each athlete completed the drill twice. If an error occurred, a third trial was performed. Errors included a deviation from the course, interference with a cone, lack of standing still at the 2 required locations, and lack of achievement of the correct VZ.

Data Extraction Methodologies. Raw data files for each athlete were exported from the Polar software to visualize the collected variables with a timestamp. Once exported, data were clipped using 2 different methods. One method of data extraction used the GPS-based initiation of movement and timing gate end of movement (method A), which was exactly 100 m, whereas the second method used the activation of the onset of first GPS movement and the end of GPS movement (method B). As a frame of reference, method B is what is reported within the Polar Team Pro Software that the end user would use when examining TD. Athletes ran the 40 and 100 m in the same bout; therefore, method A that used the timing gate as an end point was able to calculate TD for both the 40 and 100 m. However, method B was only able to calculate TD for the 100 m because the athlete did not stop at the 40 m. To reiterate for the method B, the actual TD used for the criterion are 101 m for the W and J and 105 m for the R and S (Figure 1A). During the bouts, Vel_C and Vel_I were extracted from the device and compared with the respective criterion measure, timing gate and laser, respectively.

Statistical Analyses

At the present time, there are no standard recommendations for the acceptable error for validity measures in the context of GPS variables, such as TD, Vel_I, and Vel_C. However, a previous review (27) has suggested that for the sake of consistency and alignment with previous investigations, the recommendations for reliability be used in place of validity. As such, measures of validity were rated as good (<5%), moderate (5–10%), or poor (>10%) as interpreted by Hopkins et al. (19). Furthermore, it has been suggested that errors of up to 10% would be seen as an acceptable level for time efficiency and ease of use in field-based sports. Therefore, the

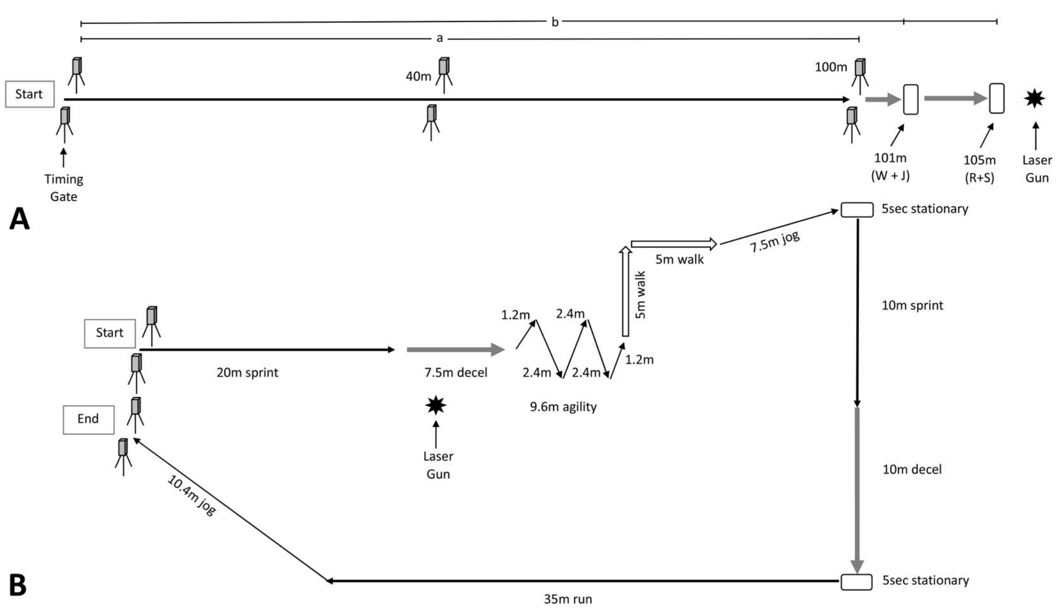

Figure 1. Linear (A) and team sport simulation circuit (B) schematics. W = walk, J = jog, R = run, S = sprint. Decel = deceleration, [a]Method A GPS start to timing gate finish location (100 m); [b]Method B GPS start to GPS finish locations for W + J (101 m) and R + S (105 m), respectively.

validity of TD during the linear tasks and TSSC was established using the standard error of the estimate (%*SEE*) (±90% confidence interval [CI]). Additionally, for the purpose of future comparison with other investigations, validity was also reported as mean difference (MD ± *SD*), %MD, effect size (ES) (±90% CI), and qualitative interpretation. Effect size was interpreted as trivial (<0.2), small (0.2–0.6), moderate (0.6–1.2), large (1.2–2.0), and very large (>2.0) (19). The Shapiro-Wilk test confirmed the assumption of normality of the statistical distribution. Total distance was compared with the criterion value, which was either 40 or 100 m for method A, 101 or 105 m for method B, and 120 m for the TSSC. The validity of Vel_I and Vel_C was determined using paired samples t tests between the device and the laser for Vel_I and the device and timing gates for Vel_C. Reliability for TD was determined during linear and TSSC using typical error (TE) (±90% CI), %CV, and smallest worth-while change (±90% CI). Smallest worth-while change as defined by Hopkins et al. (18) as $0.2ST = 0.2\sqrt{(S2 - e2)}$ or the minimal (clinical) important difference. Statistical analyses were performed using SPSS (Statistics 20.0) for Mac OS Catalina (version 10.15.2).

Results

All validity and reliability measures for TD during the 40 and 100 m (methods A and B) as well as TSSC are depicted in Table 1.

Validity and Reliability for Measuring Total Distance

The validity and reliability during linear and sport-specific movements for method A for TD during the 40 and 100 m, independent of VZ, ranged from (mean difference (MD ± *SD*) [device − criterion]) −2.8 ± 3.0 to 0.8 ± 1.1 and −2.3 ± 2.6 to 1.2 ± 1.2 m and *SEE* ranged from 2.6 to 7.5% and 1.2–2.6%, respectively (Table 1). Independent of VZ, the average mean difference indicated that the device underestimated TD by −1.3 ± 1.2 m during the 40 m and −0.5 ± 1.4 m during the 100 m and mean *SEE* was 4.0 and 1.7%, respectively. Effect sizes during the 40 m ranged from 0.1 to 1.7 m with qualitative interpretations of these differences ranging from trivial to large. Similar findings were observed during the 100 m (ES range; 0.1–1.4 m; trivial to large). Conversely, using method A, during the TSSC, the device overestimated TD by 0.2 ± 1.2 m, yet there were small effects (ES = 0.28 m), and *SEE* was good (1.0 ± 0.3%). Independent of VZ, the average magnitude of the effect size during the 40-m task was moderate (ES = 1.1), and during the 100-m task, it was moderate (ES = 0.8). Independent of VZ, the average reliability (CV) during the 40 and 100 m was 4.2 and 1.7%, respectively, and CV during the TSSC was 1.0%.

For the calculation using the initiation and final movement based on GPS start and GPS finish (method B), MD ± *SD* for TD during the 101 or 105 m ranged from 0.9 ± 0.6 to 2.3 ± 1.0 m independent of VZ. Effect sizes ranged from 0.1 to 1.4 m, yielding qualitative interpretations that varied from trivial to large. Independent of VZ, the average MD indicated that the device overestimated TD by 1.5 ± 0.7 m, although the average magnitude of this difference was very large (ES = 2.9 m) and average *SEE* was 2.9%. During the TSSC, the device overestimated TD by 2.2 ± 2.2 m and *SEE* was 1.8%. Independent of VZ, mean reliability (CV) during the 101 or 105 m was 0.7%, whereas CV during the TSSC was 1.8%.

Validity and Reliability for Measuring Velocity (Constant and Instantaneous)

All validity and reliability measures for Vel_C and Vel_I during the 40 m, 100 m, and TSSC are depicted in Table 2. The MD for Vel_C

Table 1
Validity and reliability of GPS [device − criterion] for total distance.*†

	Validity						Reliability				
Velocity zone	N	Mean difference ± *SD* (m)	% MD	ES (90% CI) (m) qualitative	% *SEE* (±90% CI)	Error interpretation	N	TE (m)	% CV	Error interpretation	SWC (m)
Total distance 40 m (a)											
Walk	25	0.1 ± 1.1	0.2	0.1 (0.5 to 0.7); trivial	2.6 ± 0.9	Good	25	0.8 (0.5 to 1.0)	2.6	Good	0.2
Jog	25	−1.0 ± 1.2	−2.6	1.2 (0.6 to 1.8); large	3.0 ± 1.0	Good	25	0.8 (0.56 to 1.11)	3.0	Good	0.2
Run	26	−1.5 ± 1.2	−3.7	1.7 (2.3 to 1.1); large	3.04 ± 1.0	Good	26	0.9 (0.58 to 1.14)	3.1	Good	0.2
Sprint	24	−2.8 ± 3.0	−6.9	1.3 (0.7 to 1.9); large	7.5 ± 2.5	Mod.	24	2.1 (1.40 to 2.82)	8.0	Mod.	0.6
Total distance 100 m (a)											
Walk	26	1.1 ± 1.1	1.1	1.4 (0.8 to 2.0); large	1.2 ± 0.4	Good	26	0.8 (0.6 to 1.1)	1.2	Good	0.2
Jog	25	−0.1 ± 1.6	−0.1	0.1 (−0.6 to 0.5); trivial	1.6 ± 0.5	Good	25	1.1 (0.8 to 1.5)	1.6	Good	0.3
Run	25	−0.7 ± 1.4	−0.7	0.7 (0.1 to 1.3); moderate	1.4 ± 0.5	Good	25	1.0 (0.7 to 1.3)	1.4	Good	0.3
Sprint	25	−2.3 ± 2.6	−2.3	1.2 (0.6 to 1.8); large	2.6 ± 0.9	Good	25	1.9 (1.3 to 2.5)	2.7	Good	0.5
Total distance 101 and 105 m (b)											
Walk	26	2.3 ± 1.0	1.3	3.2 (2.3 to 4.0); very large	1.0 ± 0.3	Good	26	0.7 (0.5 to 1.0)	1.0	Good	0.2
Jog	26	1.3 ± 0.6	1.3	3.8 (2.9 to 4.7); very large	0.6 ± 0.2	Good	26	0.4 (0.3 to 0.6)	0.6	Good	0.1
Run	26	1.0 ± 0.5	1.0	2.5 (1.8 to 3.2); very large	0.5 ± 0.2	Good	26	0.4 (2.7 to 5.3)	0.5	Good	0.1
Sprint	25	0.9 ± 0.6	1.0	2.3 (1.6 to 3.0); very large	0.5 ± 0.2	Good	25	0.4 (0.3 to 0.5)	0.5	Good	0.1
Total distance (TSSC)											
Multizone (a)	23	0.2 ± 1.2	0.2	0.3 (−0.3 to 0.9); small	1.0 ± 0.3	Good	23	0.8 (0.5 to 1.1)	1.0	Good	0.2
Multizone (b)	23	2.2 ± 2.2	1.8	1.4 (0.8 to 2.1); large	1.8 ± 0.6	Good	23	1.5 (1.0 to 2.0)	1.8	Good	0.4

*ES = Effect size; TE = Typical Error; MD = Mean difference; CV = Coefficient of variation; SEE = standard error of the estimate; SWC = Smallest worthwhile change; GPS = global positioning system; CI = confidence interval; CV = coefficient of variation; SWC = smallest worth-while change; TSSC = team sport simulation circuit.

†(a) Validity and reliability of GPS device based off initiation of movement and timing gates. (b) Validity of GPS device based off initiation of GPS movement and end of the GPS movement. Absolute mean difference ± SD. Standard error of the estimate of linear running task at different locomotions and distances is shown by percentage difference (±90% CI) between the reference distance and GPS distance to indicate bias and standard deviation (±90% CI) of the percent difference between the known distance and the GPS recorded distance for each trial.

Table 2

Validity and reliability of GPS [device − criterion] for instantaneous and constant velocity.*†

	Validity							Reliability				
Velocity zone	N	Mean difference ± *SD* (m·s^{1})	MD (%)	ES (90% CI) (m·s^{1}) qualitative	*SEE* (±90% CI) (%)	Error interpretation	T-test *p*	N	TE (m·s^{1})	CV (%)	Error interpretation	SWC (m·s^{1})
Constant velocity (40 m)												
Walk	24	−0.1 ± 0.0	1.3	0.1 (−0.7 to 0.7); trivial	1.4 ± 0.4	Good	0.165	24	0.07 (0.05 to 0.09)	1.15	Good	0.02
Jog	28	0.0 ± 0.1	−1.2	0.0 (−0.6 to 0.5); trivial	2.5 ± 0.8	Good	<0.001*	28	0.07 (0.05 to 0.09)	4.91	Good	0.02
Run	30	−0.7 ± 0.1	−2.4	0.5 (−1.0 to 0.1); small	1.7 ± 0.6	Good	<0.001*	30	0.37 (0.26 to 0.48)	0.67	Good	0.11
Sprint	26	−0.7 ± 0.2	−5.7	0.5 (−1.0 to 0.1); small	2.6 ± 0.8	Good	<0.001*	26	0.33 (0.22 to 0.43)	0.50	Good	0.09
Constant velocity (100 m)												
Walk	26	−0.0 ± 0.1	−2.9	0.2 (−0.4 to 0.7); trivial	3.3 ± 1.1	Good	<0.001*	26	0.01 (0.01 to 0.02)	1.06	Good	0.01
Jog	25	0.0 ± 0.1	−1.0	0.1 (−0.6 to 0.5); trivial	5.0 ± 1.5	Good	0.01*	25	0.05 (0.03 to 0.06)	2.01	Good	0.01
Run	25	−0.1 ± 0.5	−10.3	0.2 (−0.7 to 0.4); trivial	7.0 ± 2.1	Mod.	<0.001*	25	0.06 (0.04 to 0.08)	0.72	Good	0.02
Sprint	25	−0.4 ± 0.5	−12	1.3 (−1.9 to −0.7); large	6.0 ± 1.9	Mod.	<0.001*	25	0.13 (0.09 to 0.18)	0.45	Good	0.04
Instantaneous velocity (100 m)												
Walk	22	−0.1 ± 0.1	−5.7	0.7 (−1.3 to −0.1); moderate	7.9 ± 2.8	Mod.	0.001*	22	0.09 (0.06 to 0.13)	1.38	Good	0.03
Jog	26	0.0 ± 0.4	1.0	0.0 (−0.5 to 0.6); trivial	12.3 ± 4.0	Poor	1.000	26	0.27 (0.18 to 0.36)	12.86	Poor	0.08
Run	24	−0.1 ± 0.6	−1.5	0.1 (−0.6 to 0.5); trivial	9.9 ± 3.3	Mod.	0.456	24	0.42 (0.28 to 0.57)	6.59	Moderate	0.12
Sprint	25	0.1 ± 0.3	1.2	0.2 (−0.3 to 0.8); small	2.9 ± 0.9	Good	0.088	25	0.18 (0.12 to 0.24)	2.43	Good	0.06
Instantaneous velocity (TSSC)												
Sprint	23	0.4 ± 0.3	5.3	1.0 (0.4 to 1.6); moderate	4.7 ± 1.6	Good	<0.001*	23	0.28 (0.19 to 0.37)	5.17	Moderate	0.07

*ES = Effect size; TE = Typical Error; MD = Mean difference; CV = Coefficient of variation; SEE = standard error of the estimate; SWC = Smallest worthwhile change; GPS = global positioning system; CI = confidence interval; CV = coefficient of variation; SWC = smallest worthwhile change; TSSC = team sport simulation circuit.

†Validity and reliability of GPS device based off known instantaneous velocity and constant velocity at different velocity bands and distances. Absolute mean difference ± *SD*. Standard error of the estimate of linear running task at different locomotions and distances is shown by percentage difference (±90% CI) between the reference distance and GPS distance to indicate bias and standard deviation (±90% CI) of the percent difference between the known distance and the GPS recorded distance for each trial.

during the 40 and 100 m, independent of VZ, ranged from -0.7 ± 0.2 to 0.0 ± 0.1 and -0.1 ± 0.3 to 0.0 ± 0.1 m·s^{-1}, and *SEE* ranged from 1.4 to 2.6 and 3.3–7.0%, respectively. For Vel_C, effect sizes during the 40 m ranged from 0.1 to 0.5 m·s^{-1} with qualitative interpretations of these differences ranging from trivial to small. Similar findings were observed for Vel_c during the 100 m (ES range, 0.1–1.3 m; trivial to large). Independent of VZ, the average MD indicated the device underestimated Vel_C by -0.38 ± 0.37 m·s^{-1} during the 40 m and -0.14 ± 0.20 m·s^{-1} during the 100 m. The average magnitude of the effect sizes during both the 40 m and the 100 m was small (ES = 0.3 and 0.4), and mean *SEE* was 2.1 and 5.3%, respectively. Independent of VZ, mean Vel_C reliability (CV) during the 40 and 100 m was 1.8 and 3.6%, respectively.

The MD for Vel_I during the 100 m, independent of VZ, ranged from -0.1 ± 0.6 to 0.1 ± 0.3 m·s^{-1} and *SEE* ranged from 2.9 to 12.3%. Independent of VZ, during the 100 m, the average MD indicates the device underestimated Vel_I by -0.0 ± 0.1 m·s^{-1}, whereas the average magnitude of the effect size was small (ES = 0.3), and the average *SEE* was 8.3%. Independent of VZ, average Vel_I reliability (CV) during the 100 m was 5.8%. During the first 20 m of the TSSC sprint (Figure 1B), the device overestimated Vel_I by 0.4 ± 0.3 m·s^{-1}, the magnitude of the difference was moderate (ES = 1.0), and *SEE* was 4.7 ± 1.6%. Reliability (CV) during the TSSC was 5.2%.

Discussion

The purpose of this study was to examine the validity of a commercially available GPS device (10 Hz) that uses a novel shirt-based technology for the assessment of TD during linear running at various velocity thresholds or VZs (using 2 differing methods of data extraction) and sport-specific movement patterns. Furthermore, this study examined the validity and reliability of the device in assessing Vel_C and Vel_I.

The findings from this study determined that this GPS device demonstrated trivial to large differences in the magnitudes of effect with small errors (<5%, good) for TD at various velocities during 40- and 100-m tasks. One exception to this was the 40 m S, where a large difference in the magnitude of the effect and moderate difference in the error (%*SEE* = 7.5%) was observed. Similarly, reliability of the measurement error as indicated by the %CV during the 40 and 100 m at various velocities was small (<5%, good) with the exception of the 40 m S, (8.0%, moderate). When comparing the reliability of the GPS device during 100-m linear running, other studies have reported TD error measurements ranging from CV = 0.4–1.9% in 10-Hz GPS units (23,26). These error measures are similar to our findings (1.2–2.7% and 0.5–1.0%) for methods A and B, respectively. Indirect comparison of measures would suggest that the current GPS device examined in this study was similar to that of other 10-Hz devices when measuring 100-m TD. For example, previous research examining 3 GPS devices (SPI-10 = −5.3 m, SPI Elite = −2.4 m, and WiSPI = 0.6 m; GPSports, Canberra, Australia) from one manufacturer during the TSSC (9) observed that mean differences for TD were similar to our findings for both method A (0.23 m) and method B (2.2 m). In the same study, reliability measures for the 3 devices were 6.4, 4.0, and 7.2%, respectively, during the TSSC, which are increased compared with our reliability measures (A = 1.0% and B = 1.8%). Although there was not a direct comparison of these devices in our study and there may have been subtle methodological differences, it would appear that the device in this study demonstrated increased reliability during team sport–simulated movement patterns.

When examining Vel_C, paired samples *t* tests indicate that there are statistically significant differences between the device and the criterion, indicating that the device overestimated between 0.0 and -0.7 m·s^{-1}. Although these differences are significant, it is important to note that the magnitude of these differences as measured by the effect size and qualitative interpretation in the measurement of Vel_C are trivial to small during the 40 m and trivial to large in the 100 m. The reasoning for the significant differences can be best explained by the very narrow standard deviations surrounding the mean differences. Hence, small differences yielded statistically significant differences when comparing the device and the criterion. Thus, it is imperative for the researcher and clinician to understand that although these statistically significant differences exist, the clinical meaningfulness of these differences are quite small. When examining Vel_I, our study determined that this device demonstrated small-to-moderate effects related to the magnitude of the difference between the device and criterion with good-to-poor error reliability for Vel_I at all VZs during the 100 m. Although Varley et al. (28) did not report significant difference between the means for validity compared with the criterion, they too examined the validity of 10-Hz GPS units in measuring Vel_I during linear running using %MD. Varley et al. (28) determined that as velocity increased, the bias or %MD decreased. We observed similar findings where the %MD also decreased as velocity increased. However, the previous investigation observed a reduction in the percent bias overall in their 10-Hz device (MinimaxX 4.0; Catapult Innovations, Scoresby VIC, Australia) than the device in this study. These velocity findings are interesting because we observed an increased *SEE* of Vel_C during the 100 m (range, 3.3–7.0%) compared with the first 40 m (range, 1.4–2.6%) at all VZs, suggesting that with increase distance, there may be increased error, independent of VZ. Although not statistically supported, Vel_I error seems to reduce as the velocity increases, whereas error in Vel_C seems to increase as velocity increases. These findings are similar to one recent study examining the within- and between-unit reliability of 2 other GPS devices (5). Further research behind the mechanism or calculation for these differences in error with increasing speed and velocity are warranted. Alternatively, previous research examining the accuracy of velocity measures using laser technology identified that the distribution of errors in measurement increases with respect to the distance the target object is from the laser, thus demonstrating that the general spread of the velocity data increases as the distance from the laser increases (17). As mentioned in the methods, our laser was placed beyond the finish line and subjects ran toward the device rather than away, potentially providing an explanation for the measurement error we observed for Vel_I in Table 2. The reliability of the device when measuring Vel_c was very good with the %CV never exceeding 5% at any of the VZs. However, during the jog, we did note that the devices were the least reliable and the same went for the %CVs during the jog for Vel_I. We are unable to provide any potential rationale for this observed value other than the larger errors that were observed during the jog relative to the other VZs.

Although both methods of data extraction for TD demonstrated small error (<5% *SEE*), with the exception of the 40 m S, subtle differences were observed between the 2 calculation methods used for the error and effects reported for TD. Also, *SEE* was greatly improved using method B vs. A. However, the ES and the qualitative interpretation of the magnitude of these effects were much larger. These findings would suggest that the export of raw data files for precise TD measures when the TD is known (method A) may be necessary for research purposes and that using GPS start to GPS finish (method B) is less valid and has the potential for larger magnitude-based effects. These contrasting findings shed light on previously reported methodology and raw GPS data extraction and clipping/cropping. Frequently, the clipping process is either not stated or vaguely described in previous studies and clearly the method of data extraction, as is evident in the current study, can be the difference between reporting valid vs. invalid and "good" vs. "poor" validity and reliability measures. Future studies examining the validity and reliability of GPS devices should carefully describe raw GPS data extraction methods for the purpose of replication. This will assist with cross-device comparisons for a given metric (e.g., TD) as technology improves.

This study was limited in its ability to be applied to sport beyond the linear and team sport movements examined in this study. Certainly, athletes are required to perform linear running beyond 100 m and with more changes in direction, acceleration, and deceleration than demonstrated in the TSSC. It is possible that increased distance may result in improved or worsening validity and reliability of the device in question. Furthermore, only one device was assessed during this study; therefore, direct comparisons with other devices and their reported validity and specificity are limited because they did not undergo the same testing battery and conditions. Another potential limitation is in reference to the physical location of the device on the body (i.e., chest vs. upper back). This location we chose has the potential to be viewed as a limitation influencing its practicality given that in most use cases, this device is worn on the anterior surface of the body. That being said, there is no research to suggest that any of the variables reported would be influenced by changing the physical location of the device to the posterior aspect of the body via the shirt mount; thus, we chose to use the shirt rather than the chest strap. Furthermore, we hypothesized that in this location, the device would be less likely to slide or move or fall off while on the athlete.

Practical Applications

This study assessed the validity and reliability of the shirt-mounted 10-Hz Polar Team Pro GPS units during linear and team sport movement patterns. Our findings are critical to be aware of so that coaches and researchers can self-assess the validity and reliability of the units they use and the magnitude of these effects for measuring distance and speed during linear and team sport–simulated movements. Our findings determined that TD data reported by the Polar Team Pro units demonstrated different findings depending on the method of data analysis used. One method (method A) demonstrated trivial to large magnitudes of differences of validity with good reliability (0.5–2.7%) at all velocities during 100-m linear running, whereas the other reported very large differences with good reliability (0.5–1.0%) and similar findings, which were true for team sport–specific movements. Shorter linear distance (40 m) at various velocities ranged from trivial to large for validity and good to moderate (2.6–8.0%) for reliability. Although average speed data reported by this device overestimated statistically compared with criterion and the magnitude of those differences was trivial to large, "good" reliability during linear running during the 40 and 100 m at all VZs was observed. With the exception of walking, when measuring instantaneous velocity, the device was similar to the criterion, and the magnitude of the effect ranged from small to trivial and the reliability was poor to good. Coaches can use this technology to evaluate and compare between and within athletes across multiple training sessions for the improved management of player load, optimal recovery, and periodization. Researchers should be aware of the measurement error of these units when assessing variability of athletes and sessions for various movement demands. The GPS data extraction processes and methods when examining the validity of a device to measure TD can differ by extraction methodology used; therefore, clear explanation of these methods are critical to the applicability and interpretation. Finally, we suggest that the levels of validity and reliability reported should be taken into account when purchasing GPS systems and interpreting the movement demands of team sports from them.

Acknowledgments

The authors thank the University of Connecticut Men's Club Soccer athletes for their participation. Special thanks to Dr. William M. Adams for his contribution to the research design, The University of Connecticut Division of Public Safety, specifically the Police Department for the use of their laser equipment, and Chris A. West from the University of Connecticut Division of Athletics for use of the timing system. Funding for this study was provide by Polar Electro, Oy (Jyvaskyla Finland). Jussi Peltonen is a senior researcher employed by Polar Electro, Oy. He did not assist with the design, collection, data analysis, or interpretation of the results. His role was to assist with the review of the manuscript. All other coauthors have no financial or other interest in the product or distributor of the product. The results of the present study do no constitute endorsement of the product by the authors of the National Strength and Conditioning Association.

This study was funded in part by Polar Electro Oy.

References

1. Akenhead R, French D, Thompson KG, Hayes PR. The acceleration dependent validity and reliability of 10 Hz GPS. *J Sci Med Sport* 17: 562–566, 2014.
2. Aughey RJ. Applications of GPS technologies to field sports. *Int J Sports Physiol Perform* 6: 295–310, 2011.
3. Beato M, Coratella G, Stiff A, Iacono AD. The validity and between-unit variability of GNSS units (STATSports apex 10 and 18 Hz) for measuring distance and peak speed in team sports. *Front Physiol* 9: 1288, 2018.
4. Beato M, Devereux G, Stiff A. Validity and reliability of global positioning system units (STATSports viper) for measuring distance and peak speed in sports. *J Strength Cond Res* 32: 2831–2837, 2018.
5. Beato M, de Keijzer KL. The inter-unit and inter-model reliability of GNSS STATSports Apex and Viper units in measuring peak speed over 5, 10, 15, 20 and 30 meters. *Biol Sport* 36: 317–321, 2019.
6. Bezodis NE, Salo AIT, Trewartha G. Measurement error in estimates of sprint velocity from a laser displacement measurement device. *Int J Sports Med* 33: 439–444, 2012.
7. Bourdon PC, Cardinale M, Murray A, et al. Monitoring athlete training loads: Consensus statement. *Int J Sports Physiol Perform* 12: S2161–S2170, 2017.
8. Brewer C, Dawson B, Heasman J, Stewart G, Cormack S. Movement pattern comparisons in elite (AFL) and sub-elite (WAFL) Australian football games using GPS. *J Sci Med Sport* 13: 618–623, 2010.
9. Coutts AJ, Duffield R. Validity and reliability of GPS devices for measuring movement demands of team sports. *J Sci Med Sport* 13: 133–135, 2010.
10. Cummins C, Orr R, O'Connor H, West C. Global positioning systems (GPS) and microtechnology sensors in team sports: A systematic review. *Sports Med Auckl NZ* 43: 1025–1042, 2013.

11. Curtis RM, Huggins RA, Benjamin CL, et al. Seasonal accumulated workloads in collegiate Men's soccer: A comparison of starters and reserves. *J Strength Cond Res*, 2019. doi: 10.1519/JSC.0000000000003257.
12. Duffield R, Coutts AJ, Quinn J. Core temperature responses and match running performance during intermittent-sprint exercise competition in warm conditions. *J Strength Cond Res* 23: 1238–1244, 2009.
13. Duffield R, Reid M, Baker J, Spratford W. Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. *J Sci Med Sport* 13: 523–525, 2010.
14. Dwyer DB, Gabbett TJ. Global positioning system data analysis: Velocity ranges and a new definition of sprinting for field sport athletes. *J Strength Cond Res* 26: 818–824, 2012.
15. Farrow D, Pyne D, Gabbett T. Skill and physiological demands of open and closed training drills in Australian football. *Int J Sports Sci Coach* 3: 489–499, 2008.
16. Gray AJ, Jenkins D, Andrews MH, Taaffe DR, Glover ML. Validity and reliability of GPS for measuring distance travelled in field-based team sports. *J Sports Sci* 28: 1319–1325, 2010.
17. Harrison AJ, Jensen RL, Donoghue O. A comparison of laser and video techniques for determining displacement and velocity during running. *Meas Phys Educ Exerc Sci* 9: 219–231, 2005.
18. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 30: 1–15, 2000.
19. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 41: 3–13, 2009.
20. Hoppe MW, Baumgart C, Polglaze T, Freiwald J. Validity and reliability of GPS and LPS for measuring distances covered and sprint mechanical properties in team sports. *PLoS One* 13: e0192708, 2018.
21. Jennings D, Cormack S, Coutts AJ, Boyd L, Aughey RJ. The validity and reliability of GPS units for measuring distance in team sport specific running patterns. *Int J Sports Physiol Perform* 5: 328–341, 2010.
22. Johnston RJ, Watsford ML, Kelly SJ, Pine MJ, Spurrs RW. Validity and interunit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. *J Strength Cond Res* 28: 1649–1655, 2014.
23. Johnston RJ, Watsford ML, Pine MJ, et al. The validity and reliability of 5-Hz global positioning system units to measure team sport movement demands. *J Strength Cond Res* 26: 758–765, 2012.
24. Malone JJ, Lovell R, Varley MC, Coutts AJ. Unpacking the black box: Applications and considerations for using GPS devices in sport. *Int J Sports Physiol Perform* 12: S218–S226, 2017.
25. Misra P, Enge P. GPS coordinate frames, time reference, and orbits (Chapter 4). In: *Global Positioning System: Signals, Measurements, and Performance* Lincoln, MA: Ganga-Jamuna Press, 2011. pp. 91–133.
26. Rampinini E, Impellizzeri FM, Castagna C, Coutts AJ, Wisløff U. Technical performance during soccer matches of the Italian serie A league: Effect of fatigue and competitive level. *J Sci Med Sport Sports Med Aust* 12: 227–233, 2009.
27. Scott MTU, Scott TJ, Kelly VG. The validity and reliability of global positioning systems in team sport: A brief review. *J Strength Cond Res* 30: 1470–1490, 2016.
28. Varley MC, Fairweather IH, Aughey RJ. Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. *J Sports Sci* 30: 121–127, 2012.
29. Wehbe GM, Hartwig TB, Duncan CS. Movement analysis of Australian national league soccer players using global positioning system technology. *J Strength Cond Res* 28: 834–842, 2014.