

## ORIGINAL SCIENTIFIC PAPER

# Validity and Reliability of Wearable Devices during Self-Paced Walking, Jogging and Overground Skipping

James W. Navalta<sup>1</sup>, Dustin W. Davis<sup>1</sup>, Bryson Carrier<sup>1</sup>, Elias M. Malek<sup>1</sup>, Nicole Vargas<sup>1</sup>, Jorge Perdomo Rodriguez<sup>1</sup>, Bianca Weyers<sup>1</sup>, Katherine Carlos<sup>1</sup>, Miranda Peck<sup>1</sup>

<sup>1</sup>University of Nevada, Las Vegas, Department of Kinesiology and Nutrition Sciences, Las Vegas, Nevada, USA

## Abstract

Wearable technology can track unusual exercise, providing data for improving fitness. The aim of the study was to determine validity and reliability during walking, jogging, and skipping. Eighteen volunteers completed 5 min self-paced activities interspersed with 5 min rest. Variables and devices were step count (Garmin Instinct), estimated energy expenditure (Garmin Instinct, Polar Vantage M2), and heart rate (Garmin Instinct, Polar Vantage M2, Polar OH1, Polar Verity Sense). Validity measures were mean absolute percent error (MAPE) and Lin's Concordance (CCC), and reliability were coefficient of variation (CV), and intraclass correlation (ICC). Thresholds were MAPE  $\leq 5\%$ , CCC  $\geq 0.90$ , CV  $\leq 10\%$ , ICC  $\geq 0.70$ . Garmin Instinct step count during skipping was not considered valid (MAPE=90.2%, CCC=0.008) or reliable (CV=6%, ICC CI=0.4). Energy expenditure during skipping was not valid or reliable in the Garmin Instinct (MAPE=28%, CCC=0.27; CV=19%, ICC=0.61) or the Polar Vantage M2 (MAPE=19%, CCC=0.57; CV=13%). While the Polar Vantage M2 was reliable for estimated energy expenditure during walking and jogging activities, wrist-worn devices (Garmin Instinct, Polar Vantage M2) were neither valid nor reliable in returning estimated energy expenditure during overground skipping. From a wider perspective, wearable device algorithms for estimating energy expenditure should continue to be refined until they return the same level of accuracy as what is currently observed for heart rate, and to a lesser extent step count. Skipping may be an excellent unusual activity for testing wearable devices.

**Keywords:** wearable activity tracker, unusual exercise, accuracy, consistency

## Introduction

Wearable technology, particularly activity trackers, have sprung up in popularity in recent years. These wearable devices are designed to monitor and track physical activity, including steps taken, estimated energy expenditure, and heart rate (Bunn, Navalta, Fountaine, & Reece, 2018). Systematic reviews consistently report that using wearable activity trackers leads to improvements in weight loss, waist circumference, and body mass index (Yen & Chiu, 2019). Wearing an activity tracker and receiving daily feedback on progress can lead to increased physical activity and reduced sedentary behavior

(Jakicic et al., 2016). However, the boundless enthusiasm for wearable devices should be tempered, as the benefits of using activity trackers may depend on motivation and level of engagement with the device (Cadmus-Bertram, Marcus, Patterson, Parker, & Morey, 2015).

Wearable technology has been increasingly used to monitor and track physical activity beyond traditional exercise activities, such as in water-based sports, dancing, and video gaming. Activity trackers have been proposed to obtain metrics during water-based activities such as kayaking and canoeing (Umek & Kos, 2018; Liu, Wang, Qiu, Zhang, & Hao,



Correspondence:

James W. Navalta  
University of Nevada, Department of Kinesiology and Nutrition Sciences, 4505 S. Maryland Parkway Box 453034, Las Vegas, Nevada, 89154-3034  
E-mail: james.navalta@unlv.edu

2021). Similarly, wearable sensors can be used to assess activity during both social and jazz dancing activities (Stančin & Tomažič, 2022), making it a useful tool for dancers who must perform repetitive leaping motions to track their progress and improve their performance. Furthermore, virtual reality gaming has gained popularity as an alternative form of exercise, and wearable technology such as virtual reality headsets and motion capture devices can be used to monitor and track physiological responses during these activities (Cao, Xie, & Chen, 2019). Overall, wearable technology has shown potential in monitoring and tracking physical activity in a variety of unusual exercise activities, providing useful data for individuals to optimize their performance and improve their overall fitness levels. However, further research is needed to explore the accuracy and reliability of these devices in diverse settings and unusual activities.

Bounding straight to the issue, overground skipping is a non-traditional exercise that has not been investigated to date in the context of wearable technology devices. Most humans transition from walking to running as a faster method of travel and exercise (Navalta, Davis, Carrier, Sertic, & Cater, 2021), however, running requires greater energy expenditure than walking (Harrell et al., 2005). Some people, particularly children, perform skipping as a transitory movement as an alternative to walking and running (Navalta et al., 2021). The available literature centers on rope skipping, rather than overground skipping (Verdel et al., 2022; Yongmao & Yuxin, 2023). One study evaluated skipping in 20-sec bouts as obtained from a novel smart patch (Verdel et al., 2022), and a recent investigation was conducted on rope skipping and deep learning algorithms (Yongmao & Yuxin, 2023). However, overground skipping is different in that the action involves forward propulsion that is absent in rope skipping, which may be responsible for half of the energy cost of walking (Gottschall & Kram, 2005). Thus, the evaluation of wearable devices to return valid and reliable measurements during overground skipping should be conducted.

The purpose of this investigation was to determine the validity and reliability of wearable technology devices during an overground skipping activity. While many studies are designed to determine validity, there is a need to also obtain reliability metrics (Carrier, Barrios, Jolley, & Navalta, 2020). Previous research has shown wearable devices to have varied levels of validity, with heart rate generally being acceptable, energy expenditure estimation unacceptable, and step count falling in between (Bunn et al., 2018). Based on previous literature and testing performed in our laboratory, it was hypothesized that during overground skipping, step count would be reliable but not valid, estimates of energy expenditure would be neither valid nor reliable, and heart rate would be both valid and reliable. These findings are important, as this is the first study to report validity and reliability of wearable devices during an overground skipping activity.

## Methods

### Participants

Eighteen participants (female  $n=10$ , male  $n=8$ , transgender, intersex, or other  $n=0$ ) provided written informed consent that was approved by the Institutional Review Board (IRB approval# UNLV-2022-80) and volunteered for the study. A power analysis was performed using pilot data, indicating the need for at least eleven participants (coefficient of

determination  $r^2=0.57$ , correlation  $r$  effect size=0.755,  $\alpha=0.05$ ,  $b=0.80$ ). Participants were screened and deemed not to require medical clearance to complete exercise according to the American College of Sports Medicine preparticipation health screening recommendations. The mean demographic information included: age= $26\pm 8$  years, body mass= $73.3\pm 11.9$  kg, height= $66.4\pm 3.5$  cm, and Body Mass Index= $25.6\pm 3.0$   $\text{kg}\cdot\text{m}^{-2}$ . Participants self-identified ethnic group included African American ( $n=1$ ), White ( $n=10$ ), Hispanic ( $n=3$ ), Polynesian ( $n=1$ ), and Southeast Asian ( $n=3$ ).

### Device Setup

Demographic information was obtained and input into the criterion and wearable devices prior to testing each participant. Participants were outfitted with criterion devices (K5 [COSMED, Rome, Italy] for energy expenditure; H10 [Polar Electro, Kempele, Finland] for heart rate) and wearable devices (described below), and a secure Bluetooth connection was confirmed. In all cases, devices were affixed to the body according to manufacturer recommendations. Briefly, the K5 was secured on the back of the participant, the H10 attached around the chest. The biceps-worn experimental devices were the Polar OH1 (Polar Electro, Kempele, Finland) and Polar Verity Sense (Polar Electro, Kempele, Finland), placed on the right and left biceps each. The wrist-worn experimental devices were the Garmin Instinct (Garmin Limited, Olathe, Kansas) and Polar Vantage M2 (Polar Electro, Kempele, Finland), placed on the right and left wrist each. Two of the same models of each experimental device were used simultaneously so that concurrent reliability could be obtained (Pinedo-Jauregi, Garcia-Tabar, Carrier, Navalta, & Camara, 2022).

The chest-worn (H10) and biceps-worn devices (OH1, Verity Sense) were connected via Bluetooth to an iPad mini (Apple Inc., Cupertino, CA) with the PerformTek application (Valencell, Inc., Raleigh, NC) which returns heart rate of all connected devices on a single csv file. The wrist-worn devices (Instinct, Vantage M2) collected information directly onto the device and were downloaded at a later time using recommended applications (Instinct through the Garmin Connect application and Bluetooth connection; Vantage M2 through the Polar FlowSync desktop application and Polar Flow online database).

### Exercise Bouts

After confirming connection to the criterion and experimental devices, participants performed a self-paced walk back and forth through an indoor hallway along a 61-meter track, followed by 5 minutes (min) of seated rest. All devices were reset during the period of rest (i.e., new activity recording session with the same demographic and anthropometric data in the watch; not a factory reset). Participants then completed 5 min of self-paced running back and forth through the same indoor hallway, followed by 5 min of seated rest. All devices were again reset during the period of rest. Finally, participants performed 5 min of self-paced overground skipping back and forth through the same indoor hallway. Step count was collected via manual clicker step counters by two independent observers during the overground skipping bout only. The step count was the arithmetic mean of the observer's manual counts. A step was defined as the completion of a unilateral stride before transferring motion to the contralateral

side of the body. The protocol is similar in terms of timing and administration to other investigations conducted by the laboratory group (Montes & Navalta, 2019; Navalta et al., 2019; Montes, Tandy, Young, Lee, & Navalta, 2022).

*Devices*

**Polar H10:** Although the use case specific to overground skipping has not been determined, the Polar H10 chest strap has been shown in other settings to have acceptable reliability (Speer, Semple, Naumovski, & McKune, 2020) and to be valid compared to electrocardiography (Gilgen-Ammann, Schweizer, & Wyss, 2019). The Polar H10 is an electrocardiogram-based heart rate sensor secured around the chest at the level of the xyphoid process. The H10 contains plastic electrodes on the underside of the strap that detects heart rate. The sensor materials include acrylonitrile butadiene styrene (ABS), ABS plus glass fiber (ABS+GF), polycarbonate, and stainless steel, and the strap material is composed of 38% polyamide, 29% polyurethane, 20% elastane, 13% polyester, and silicone prints. The H10 has a sampling frequency of 1000 Hertz (Hz).

**Polar OH1:** The Polar OH1 is a photoplethysmography (PPG) device that uses an optical heart rate sensor worn on the upper arm. The OH1 is an optical heart rate sensor worn on the upper arm. The sensor materials include ABS, ABS+GF, poly(methyl methacrylate) (PMMA), and steel use stainless (SUS) 316. The device was positioned so that the sensor was on the underside of the armband and firmly against the skin. The OH1 has a sample rate of 135 Hz.

**Polar Verity Sense:** The Polar Verity Sense is also a PPG device worn on the upper arm. The sensor materials include ABS, ABS+GF, PMMA, and SUS 316. The device was positioned with the sensor on the underside of the armband and firmly against the skin. The Verity Sense has a sample rate of 135 Hz.

**Garmin Instinct:** The Garmin Instinct is a PPG device. The physical size is 46 x 46 x 12.5 millimeters (mm) with a mass of 45 grams (g). The optical heart rate sensor is worn on the wrist and employs Garmin Elevate™ heart rate sensor technology. The sample rate is unknown.

**Polar Vantage M2:** The Polar Vantage M2 is a PPG device. The physical size is 45 x 45 x 15.3 mm with a mass of 52 g. The optical heart rate sensor is worn on the wrist and em-

ploy Precision Prime™ heart rate sensor technology. The Vantage M2 is reported to have a sample rate of 135 Hz.

*Statistical Analysis*

Reported measures associated with validity include mean absolute percent error (MAPE), Lin’s Concordance Correlation Coefficient (CCC), and the mean absolute error (MAE). The equations for these metrics were input into an Excel spreadsheet (Microsoft Excel for Mac version 16.66.1, Redmond, WA). For validity thresholds our laboratory group has used a MAPE value ≤ 5%, and a CCC ≥ 0.90 (J. W. Navalta et al., 2020). A device meeting both thresholds was considered evidence in support of validity. The Bland-Altman analysis was used to determine agreement. Bias and limits of agreement were determined using the blandr analysis in jamovi (version 2.3.19.0).

Reported measures associated with reliability include the coefficient of variation (CV), and two-way mixed model with absolute agreement intraclass correlation coefficient (ICC). The CV was determined using Excel (Microsoft Excel for Mac version 16.66.1, Redmond, WA), and the ICC (single measures) using SPSS Statistics (IBM SPSS Statistics, version 28.0.1.0, Chicago, IL). For non-laboratory settings a threshold of ≤ 10% for CV, and ≥ 0.70 for ICC, with a lower bound of the 95% CI ≥ 0.70 has been used (Navalta et al., 2019). A device meeting both thresholds was considered evidence in support of reliability.

**Results**

*Step Count*

Step count was obtained only from the Garmin Instinct during the skipping trial because the other devices did not return the measure. Manual step count was not obtained from one participant and was not recorded from the Garmin Instinct for two other participants. Thus, validity measures reflect data points from 30 distinct bouts, and reliability measures from 16 pairs (see table 1). Based on the predefined definition of a step, the device did not meet any of the threshold measures for validity and overestimated the step count by nearly double (see Table 1). The Garmin Instinct met the reliability thresholds for CV and ICC but was not considered reliable because the lower end of the 95% CI did not meet the established threshold.

**Table 1.** Step count during the self-paced skipping bout. Criterion (manually counted steps), and validity and reliability measures of the Garmin Instinct.

	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
Skipping	Manual Count	30	507.3 (32.9)				
	Garmin Instinct	30	957.2 (139.9)	90.2	-499.8	181.6 to -718.5	0.008
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	16	933.7 (165.3)	1000.9 (103.4)	6.0	0.74	0.40 to 0.90

Average is the arithmetic mean (standard deviation) measurement of step count. MAPE = mean absolute percent error, CCC = Lin’s Concordance Correlation Coefficient, CV = coefficient of variation, ICC = Intraclass correlation coefficient, CI = Confidence interval of the ICC. Values noted in bold and italics meet the predetermined threshold for validity or reliability.

*Energy Expenditure*

There was no missing energy expenditure data from the criterion device or any of the wearable activity trackers. Validity and reliability measures for wrist-worn devices (Garmin Instinct, Polar Vantage M2) during walking, jogging, and skipping are provided in Table 2. The biceps-worn devices

did not return an estimate of energy expenditure. Neither the Garmin Instinct nor the Polar Vantage M2 met the predetermined validity thresholds when returning estimates of energy expenditure across any bout of exercise. The Polar Vantage M2 met the reliability thresholds during self-paced walking and jogging, but not overground skipping.

**Table 2.** Energy expenditure during self-paced walking, jogging, and skipping bouts. Criterion (COSMED K5), and validity and reliability measures of the Garmin Instinct, and Polar Vantage M2.

	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
Walking	COSMED K5	36	26.2 (5.6)				
	Garmin Instinct	36	37.1 (13.6)	57.0	-10.9	19.3 to -41.0	-0.06
	Polar Vantage M2	36	37.3 (9.0)	46.7	-11.1	6.1 to -28.3	0.15
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	18	38.7 (15.4)	36.5 (12.4)	14.3	0.75	0.45 to 0.90
	Polar Vantage M2	18	37.9 (10.1)	36.7 (8.7)	6.0	0.89	0.74 to 0.96
Jogging	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
	COSMED K5	36	54.6 (12.2)				
	Garmin Instinct	36	64.6 (14.8)	26.0	-10.0	15.4 to -35.4	0.42
	Polar Vantage M2	36	61.1 (13.1)	18.1	-6.6	12.0 to -25.2	0.63
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	18	68.2 (14.3)	64.0 (12.6)	10.9	0.59	0.18 to 0.82
Polar Vantage M2	18	64.0 (12.6)	60.3 (12.8)	4.8	0.90	0.76 to 0.96	
Skipping	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
	COSMED K5	36	59.6 (14.3)				
	Garmin Instinct	36	57.4 (19.6)	28.1	2.2	42.6 to -38.3	0.27
	Polar Vantage M2	36	64.5 (17.4)	18.6	-4.9	23.1 to -32.9	0.57
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	18	60.9 (19.3)	56.1 (20.2)	19.1	0.61	0.21 to 0.83
Polar Vantage M2	18	71.8 (17.5)	59.7 (14.7)	13.1	0.83	0.60 to 0.93	

Average is the arithmetic mean (standard deviation) measurement of kilocalories expended. MAPE = mean absolute percent error, CCC = Lin's Concordance Correlation Coefficient, CV = coefficient of variation, ICC = Intraclass correlation coefficient, CI = Confidence interval of the ICC. Values noted in bold and italics meet the predetermined threshold for validity or reliability.

**Heart Rate**

Heart rate measures were not obtained from the Polar OH1 for one participant during the walking trial. There was no missing heart rate data from the criterion device or any of the other wearable activity trackers. All devices met the pres-

established validity thresholds during the self-paced walking trial (see Table 3). One wrist-worn device (Garmin Instinct) and one biceps-worn device (Polar Verity Sense) met the thresholds for reliability during walking (see Table 3) and can be considered both valid and reliable during this type of exercise.

**Table 3.** Heart rate during self-paced walking, jogging, and skipping bouts. Criterion (H10), and validity and reliability measures of wrist-worn (Garmin Instinct, Polar Vantage M2) and biceps-worn (Polar Verity Sense, Polar OH1) devices.

	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
Walking	Polar H10	36	104.5 (17.1)				
	Garmin Instinct	36	105.5 (16.6)	2.8	-1.0	7.0 to -9.1	0.97
	Polar Vantage M2	36	104.7 (14.2)	3.3	-0.2	13.0 to -13.4	0.91
	Polar Verity Sense	36	104.3 (16.7)	0.7	0.3	2.2 to -1.7	0.99
	Polar OH1	35	104.5 (18.0)	2.1	0.5	14.9 to -13.9	0.91
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	18	107.3 (16.5)	106.6 (15.9)	2.2	0.96	0.91 to 0.99
	Polar Vantage M2	18	105.0 (11.1)	106.0 (16.4)	3.3	0.78	0.50 to 0.91
	Polar Verity Sense	18	105.9 (16.3)	105.6 (16.0)	0.3	0.99	0.99 to 1.0
	Polar OH1	17	106.3 (18.7)	107.0 (16.2)	2.7	0.842	0.62 to 0.94
Jogging	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
	Polar H10	36	154.9 (13.9)				
	Garmin Instinct	36	147.2 (15.8)	5.9	7.7	32.8 to -17.5	0.55
	Polar Vantage M2	36	148.2 (15.1)	4.4	6.7	26.1 to -12.7	0.69
Polar Verity Sense	36	151.7 (13.7)	2.0	3.1	10.3 to -4.0	0.94	

(continued on next page)

(continued from previous page)

**Table 3.** Heart rate during self-paced walking, jogging, and skipping bouts. Criterion (H10), and validity and reliability measures of wrist-worn (Garmin Instinct, Polar Vantage M2) and biceps-worn (Polar Verity Sense, Polar OH1) devices.

	Polar OH1	38	152.6 (13.7)	1.5	2.3	4.9 to -0.3	0.98
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
Jogging	Garmin Instinct	18	148.5 (14.3)	150.5 (11.4)	3.3	0.71	0.38 to 0.88
	Polar Vantage M2	18	153.5 (14.1)	145.8 (12.4)	4.1	0.67	0.30 to 0.86
	Polar Verity Sense	18	153.1 (13.9)	151.6 (14.1)	0.9	0.96	0.89 to 0.98
	Polar OH1	18	153.3 (14.1)	153.2 (14.0)	0.1	0.99	0.99 to 1.0
	Validity	Data Points	Average	MAPE (%)	Bias	Limits of Agreement	CCC
Skipping	Polar H10	36	170.6 (17.5)				
	Garmin Instinct	36	133.3 (25.5)	21.7	37.3	84.7 to -10.1	0.16
	Polar Vantage M2	36	156.3 (24.0)	8.4	14.3	48.4 to -19.8	0.53
	Polar Verity Sense	36	168.7 (17.1)	1.1	2.0	4.7 to -0.8	0.99
	Polar OH1	36	167.3 (18.6)	1.9	3.3	16.1 to -9.5	0.92
	Reliability	Data Points	Device 1	Device 2	CV (%)	ICC	95% CI
	Garmin Instinct	18	136.8 (22.8)	132.5 (28.6)	9.5	0.49	0.04 to 0.77
	Polar Vantage M2	18	170.1 (17.2)	147.2 (21.2)	10.6	0.67	0.30 to 0.87
	Polar Verity Sense	18	170.1 (17.2)	169.6 (16.9)	0.4	0.99	0.98 to 0.99
	Polar OH1	18	168.8 (18.3)	168.2 (19.1)	0.4	0.99	0.98 to 0.99

MAPE = mean absolute percent error, CCC = Lin's Concordance, CV = coefficient of variation, ICC = Intraclass correlation coefficient, CI = Confidence interval. Values noted in bold and italics meet the predetermined threshold for validity or reliability.

During both self-paced jogging and self-paced skipping, the biceps-worn devices (Polar OH1, Polar Verity Sense) met all preestablished thresholds for validity and reliability (see Table 3). Thus, the Polar OH1 and the Polar Verity Sense can be considered valid and reliable during jogging as well as overground skipping.

**Discussion**

The purpose of this investigation was to determine the validity and reliability of wearable technology devices to return step count, energy expenditure, and heart rate during walking, jogging, and overground skipping. It was hypothesized that, during overground skipping, step count would be reliable but not valid, energy expenditure would be neither valid nor reliable, and heart rate would be both valid and reliable. First, the hypothesis for step count was partially supported because step count during self-paced skipping was neither valid nor reliable. Second, the hypothesis for energy expenditure was supported because estimated energy expenditure during self-paced skipping was neither valid nor reliable. Third, the hypothesis for heart rate was supported for biceps-worn devices (valid and reliable) but not wrist-worn technology (neither valid nor reliable).

The conclusions drawn from step count obtained during skipping are obtained from a single wrist-worn wearable device and should be considered with caution. The device did not meet the validity thresholds, likely because of how a step was defined in the current investigation (as a complete unilateral stride before transferring motion to the contralateral side of the body). The definition did not account for the foot tap down during a stride that many participants performed during the skipping motion, and likely accounted for the very large bias that was observed. It may be useful for future researchers interested in step count during overground skipping

to account for the foot tap, as it appears at least one wearable device registers this action. It is proposed that users of consumer activity trackers may utilize this knowledge to their advantage, to increase step count during wellness program challenges by performing overground skipping as their preferred method of transportation, which would conceivably double the step count compared to the more mundane method of walking. The practical implications of this finding should not be skipped over.

Heart rate was observed to be both valid and reliable in all devices during the walking bout. When wrist-worn devices (Garmin Instinct, Polar Vantage M2) are considered, the finding did not extend to activities of greater relative intensity, being considered neither valid nor reliable during jogging and overground skipping. This finding aligns with a study using a different wrist-based device, the Garmin fenix 5, which reported moderate validity measures during sitting and walking, but poor validity measures during increased intensities of exercise (Duking et al., 2020). However, an earlier Polar wrist-worn model, the Polar Vantage V, was reported to have acceptable MAPE values for heart rate during running and cycling at low and high self-selected intensities (Hajj-Boutros, Landry-Duval, Comtois, Gouspillou, & Karelis, 2023). The specific wrist-based devices used in the current investigation have little published literature outside of conference abstracts, many of which are from our laboratory group. The same holds true for one of the biceps-worn devices, the Polar Verity Sense, in that published validation literature currently exists only in abstract form (Bodell et al., 2021). The other biceps-worn device, the Polar OH1, is reported to have acceptable validity during treadmill (MAPE between 0.2 and 1.9%) and cycle exercise (MAPE between 0.6 and 3.9%) (Muggeridge et al., 2021), spin bike activities (mean bias less than 1 bpm) (Hettiarachchi, Hanoun, Nahavandi, & Nahavandi, 2019),

front crawl swimming at various intensities (ICC between 0.72 and 0.96) (Olstad & Zinner, 2020), and endurance sports (difference from criterion < 5%), as well as acceptable reliability during endurance sports (ICC = 0.99) (Hermand, Cassirame, Ennequin, & Hue, 2019). Based on the results of the current investigation, biceps-worn devices may be preferable for individuals who are interested in obtaining both valid and reliable heart rate measures in non-traditional activities such as overground skipping.

It has been noted that estimated energy expenditure in wearable devices during exercise validation studies is generally poor (Bunn et al., 2018). As described with heart rate above, the specific devices used in the current investigation have not been reported in the published literature, and the same holds true for energy expenditure. The existing literature details opportunities for improvement when energy expenditure is considered. Physical activity energy expenditure in wearable devices including the Garmin vívofit, Jawbone Up24, and Fitbit Flex, compared to doubly labeled water were deemed unacceptable (MAPE range for 12 devices 19.4% to 100.2%) (Murakami et al., 2019). Estimated energy expenditure returned from the Garmin vivo HR+ and Fitbit Charge 2 were evaluated against COSMED units (K4, K5) during progressive exercise tests to volitional fatigue (treadmill, cycle ergometer) and MAPE did not meet validity threshold (18.9% to 43.5%) (Reddy et al., 2018). During trail running, the Hexoskin biometric shirt did not return correlated energy expenditure measures compared to the COSMED K4 ( $r = -0.058$ ) (Tanner et al., 2016). The Apple watch 6 (14.9% to 47.8% MAPE), Polar Vantage V (15.6% to 34.6% MAPE), and Fitbit Sense (17.8% to 45.1% MAPE) all displayed unacceptable MAPE during walking, running, resistance exercises, and cycling (Hajj-Boutros et al., 2023). The current investigation adds to the volume of literature, in that estimated energy expenditure obtained from Garmin Instinct and Polar Vantage M2 wrist-worn devices during self-paced walking, jogging, and overground skipping are similarly poor when validity measures are considered. While the current investigation reports the Polar Vantage M2 to satisfy reliability thresholds during walking and jogging, this is of little consolation if the validity assumptions are violated.

The current investigation is not without limitations or

hurdles, which are inherent in wearable technology validation studies (Navalta & Bunn, 2023). As mentioned previously, the predetermined definition for what constitutes a step during the overground skipping motion likely led to the device returning poor validation measures. Future studies employing overground skipping while obtaining step count measures should take this factor into account. Another limitation is that the order of the exercise bouts was constant, with walking first, then jogging, and ending with overground skipping. Because the intent of the study was to obtain concurrent validity and reliability measures, it is believed that this approach is acceptable. However, it is unknown whether scheduling the overground skipping trial last affected mechanics of the skill, and whether potential fatigue may have induced extraneous movements that decreased validity and or reliability measurements in certain devices. Future studies employing various exercise modalities should not skip over randomizing the order of exercise bouts.

In conclusion, the modality of overground skipping is an area of research that represents boundless possibilities. In the context of the current investigation, several findings are reported on specific wearable devices for the first time. Regarding heart rate, the Garmin Instinct, Polar Vantage M2, Polar Verity Sense, and Polar OH1 met all thresholds to be considered both valid and reliable during self-paced walking. Only the biceps-worn devices (Polar OH1, Polar Verity Sense) met the heart rate thresholds for self-paced jogging and overground skipping. The Garmin Instinct was the only device to return step count, and it was neither valid nor reliable during overground skipping according to our predetermined definition of a step. While the Polar Vantage M2 was reliable for energy expenditure during walking and jogging, wrist-worn devices (Garmin Instinct, Polar Vantage M2) were neither valid nor reliable in returning estimated energy expenditure during overground skipping. From a wider perspective, wearable device algorithms for estimating energy expenditure should continue to be refined until they return the same level of accuracy as what is currently observed for heart rate, and to a lesser extent step count. Overground skipping may be an excellent unusual exercise for the continued testing of these measures as leaps and bounds are made in the development wearable activity devices.

#### Acknowledgements

There are no acknowledgements.

#### Conflict of Interest

The authors declare that there is no conflict of interest.

**Received:** 30 July 2023 | **Accepted:** 19 September 2023 | **Published:** 01 October 2023

#### References

- Bodell, N., Carrier, B., Gil, D., Fullmer, W., Cruz, K., Aguilar, C. D., ... Manning, J. W. (2021). *Validity of Average Heart Rate and Energy Expenditure in Polar OH1 and Verity Sense While Self-Paced Walking*. Paper presented at the International Journal of Exercise Science: Conference Proceedings.
- Bunn, J. A., Navalta, J. W., Fountaine, C. J., & Reece, J. D. (2018). Current State of Commercial Wearable Technology in Physical Activity Monitoring 2015-2017. *International Journal of Exercise Science*, 11(7), 503-515. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/29541338>
- Cadmus-Bertram, L. A., Marcus, B. H., Patterson, R. E., Parker, B. A., & Morey, B. L. (2015). Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. *American Journal of Preventative Medicine*, 49(3), 414-418. doi:10.1016/j.amepre.2015.01.020
- Cao, M., Xie, T., & Chen, Z. (2019). Wearable sensors and equipment in VR games: a review. *Transactions on Edutainment*, XV, 3-12.
- Carrier, B., Barrios, B., Jolley, B. D., & Navalta, J. W. (2020). Validity and Reliability of Physiological Data in Applied Settings Measured by Wearable Technology: A Rapid Systematic Review. *Technologies*, 8(4). doi:ARTN 7010.3390/technologies8040070
- Duking, P., Giessing, L., Frenkel, M. O., Koehler, K., Holmberg, H. C., & Sperlich, B. (2020). Wrist-Worn Wearables for Monitoring Heart Rate and Energy Expenditure While Sitting or Performing Light-to-Vigorous Physical Activity: Validation Study. *JMIR Mhealth Uhealth*, 8(5), e16716. doi:10.2196/16716
- Gilgen-Ammann, R., Schweizer, T., & Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *European Journal of Applied Physiology*, 119(7), 1525-1532. doi:10.1007/s00421-019-04142-5
- Gottschall, J. S., & Kram, R. (2005). Energy cost and muscular activity required for leg swing during walking. *Journal of Applied Physiology*, 99(1), 23-30.
- Hajj-Boutros, G., Landry-Duval, M. A., Comtois, A. S., Gouspillou, G., & Karelis, A. D. (2023). Wrist-worn devices for the measurement of heart rate and energy expenditure: A validation study for the Apple Watch 6, Polar Vantage V and Fitbit Sense. *European Journal of Sport Science*, 23(2), 165-177. doi:10.1080/17461391.2021.2023656
- Harrell, J. S., McMurray, R. G., Baggett, C. D., Pennell, M. L., Pearce, P. F., & Bangdiwala, S. I. (2005). Energy Costs of Physical Activities in Children and Adolescents. *Medicine & Science in Sports & Exercise*, 37(2), 329-336. doi:10.1249/01.Mss.0000153115.33762.3f

- Hermand, E., Cassirame, J., Ennequin, G., & Hue, O. (2019). Validation of a Photoplethysmographic Heart Rate Monitor: Polar OH1. *International Journal of Sports Medicine*, *40*(7), 462-467. doi:10.1055/a-0875-4033
- Hettiarachchi, I. T., Hanoun, S., Nahavandi, D., & Nahavandi, S. (2019). Validation of Polar OH1 optical heart rate sensor for moderate and high intensity physical activities. *PLoS One*, *14*(5), e0217288. doi:10.1371/journal.pone.0217288
- Jakicic, J. M., Davis, K. K., Rogers, R. J., King, W. C., Marcus, M. D., Helsen, D., . . . Belle, S. H. (2016). Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. *JAMA*, *316*(11), 1161-1171. doi:10.1001/jama.2016.12858
- Liu, L., Wang, H. H., Qiu, S., Zhang, Y. C., & Hao, Z. D. (2021). Paddle Stroke Analysis for Kayakers Using Wearable Technologies. *Sensors (Basel)*, *21*(3). doi:10.3390/s21030914
- Montes, J., & Navalta, J. W. (2019). Reliability of the Polar T31 Uncoded Heart Rate Monitor in Free Motion and Treadmill Activities. *International Journal of Exercise Science*, *12*(4), 69-76.
- Montes, J., Tandy, R. D., Young, J. C., Lee, S.-P., & Navalta, J. W. (2022). Is Body Composition or Body Mass Index Associated with the Step Count Accuracy of a Wearable Technology Device? *Topics in Exercise Science and Kinesiology*, *3*(1), 5.
- Muggeridge, D. J., Hickson, K., Davies, A. V., Giggins, O. M., Megson, I. L., Gorely, T., & Crabtree, D. R. (2021). Measurement of Heart Rate Using the Polar OH1 and Fitbit Charge 3 Wearable Devices in Healthy Adults During Light, Moderate, Vigorous, and Sprint-Based Exercise: Validation Study. *JMIR Mhealth Uhealth*, *9*(3), e25313. doi:10.2196/25313
- Murakami, H., Kawakami, R., Nakae, S., Yamada, Y., Nakata, Y., Ohkawara, K., . . . Miyachi, M. (2019). Accuracy of 12 Wearable Devices for Estimating Physical Activity Energy Expenditure Using a Metabolic Chamber and the Doubly Labeled Water Method: Validation Study. *JMIR Mhealth Uhealth*, *7*(8), e13938. doi:10.2196/13938
- Navalta, J. W., & Bunn, J. A. (2023). Editorial: Continued opportunities in wearable technologies and physiological assessment. *Frontiers in Digital Health*, *5*, 1166088. doi:10.3389/fdgth.2023.1166088
- Navalta, J. W., Davis, D. W., Carrier, B., Sertic, J. V., & Cater, P. (2021). Teaching Applied Exercise Physiology Using a Prototype Energy Expenditure Measurement Device. *Interdisciplinary Journal of Problem-based Learning*, *15*(2), n2.
- Navalta, J. W., Montes, J., Bodell, N. G., Aguilar, C. D., Radzak, K., Manning, J. W., & DeBeliso, M. (2019). Reliability of Trail Walking and Running Tasks Using the Stryd Power Meter. *International Journal of Sports Medicine*, *40*(8), 498-502. doi:10.1055/a-0875-4068
- Navalta, J. W., Montes, J., Bodell, N. G., Salatto, R. W., Manning, J. W., & DeBeliso, M. (2020). Concurrent heart rate validity of wearable technology devices during trail running. *PLoS One*, *15*(8), e0238569. doi:10.1371/journal.pone.0238569
- Olstad, B. H., & Zinner, C. (2020). Validation of the Polar OH1 and M600 optical heart rate sensors during front crawl swim training. *PLoS One*, *15*(4), e0231522. doi:10.1371/journal.pone.0231522
- Pinedo-Jauregi, A., Garcia-Tabar, I., Carrier, B., Navalta, J. W., & Camara, J. (2022). Reliability and validity of the Stryd Power Meter during different walking conditions. *Gait Posture*, *92*, 277-283. doi:10.1016/j.gaitpost.2021.11.041
- Reddy, R. K., Pooni, R., Zaharieva, D. P., Senf, B., El Youssef, J., Dassau, E., . . . Jacobs, P. G. (2018). Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study. *JMIR Mhealth Uhealth*, *6*(12), e10338. doi:10.2196/10338
- Speer, K. E., Semple, S., Naumovski, N., & McKune, A. J. (2020). Measuring Heart Rate Variability Using Commercially Available Devices in Healthy Children: A Validity and Reliability Study. *European Journal of Investigation in Health, Psychology and Education*, *10*(1), 390-404. doi:10.3390/ejihpe10010029
- Stančin, S., & Tomažič, S. (2022). Recognizing Solo Jazz Dance Moves Using a Single Leg-Attached Inertial Wearable Device. *Sensors*, *22*(7), 2446. Retrieved from <https://www.mdpi.com/1424-8220/22/7/2446>
- Tanner, E., Manning, J. W., Taylor, J., Montes, J., McCune, D., Koschel, T., . . . Navalta, J. W. (2016). Validation of Hexoskin Biometric Shirt to Cosmed k4b2 Metabolic Unit in Adults During Trail Running. *Medicine and Science in Sports and Exercise*, *48*(5), 461-461. doi:10.1249/01.mss.0000486387.43740.5e
- The jamovi project. (2023). Jamovi (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>
- Umek, A., & Kos, A. (2018). Wearable sensors and smart equipment for feedback in watersports. *Procedia Computer Science*, *129*, 496-502.
- Verdel, N., Drobnič, M., Maslik, J., Hellman, O., Gumiero, A., Dellatorre, L., . . . Supej, M. (2022). The Validity of Measuring Skipping Cadence with a Novel Wearable Sensor - SINTEC Smart Patch. *ISBS Proceedings Archive*, *40*(1), 728.
- Yen, H. Y., & Chiu, H. L. (2019). The effectiveness of wearable technologies as physical activity interventions in weight control: A systematic review and meta-analysis of randomized controlled trials. *Obesity Reviews*, *20*(10), 1485-1493. doi:10.1111/obr.12909
- Yongmao, Z., & Yuxin, L. (2023). Application of wearable devices based on deep learning algorithm in rope skipping data monitoring. *Soft Computing*, 1-11.