5-1-2021

Investigation of Algorithms to Assess Validity of Wearable Technology During Field Testing

Brenna Barrios

INVESTIGATION OF ALGORITHMS TO ASSESS VALIDITY OF WEARABLE TECHNOLOGY DURING FIELD TESTING

By

Brenna Barrios

Bachelors of Science - Kinesiology
California State University Fullerton
2019

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science – Exercise Physiology

Department of Kinesiology and Nutrition Sciences
School of Integrated Health Science
The Graduate College

University of Nevada, Las Vegas
May 2021

Thesis Approval

The Graduate College
The University of Nevada, Las Vegas

April 12, 2021

This thesis prepared by

Brenna Barrios

entitled

Investigation of Algorithms to Assess Validity of Wearable Technology During Field Testing

is approved in partial fulfillment of the requirements for the degree of

Master of Science – Exercise Physiology
Department of Kinesiology and Nutrition Sciences

James Navalta, Ph.D.
Examination Committee Chair

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Dean

John Mercer, Ph.D.
Examination Committee Member

Tedd Girouard, MS
Examination Committee Member

Kai-Yu Ho, Ph.D.
Graduate College Faculty Representative

Abstract

Wearable technology is an emerging fitness trend where the technology which supports it lacks validity verification. Furthermore, heart rate validity of these devices vary greatly when observed in laboratory settings vs. field testing. Secondarily, Consumer Technology Association guidelines require a minimum five minute trial for wearable testing. This study examined heart rate data of previously tested wearable devices (Garmin Fenix 5, Jabra Elite Sport, Motiv Ring, Scosche Rhythm+) in an outdoor setting to further understand their performance, and to examine the relationship of the five minute regulation. Two separate algorithms were applied to the original data set, the first shortened the data to consist of the first five minutes of each trial. The second algorithm eliminated values that were outside of a range from the criterion (10% above the criterion or 10% below the criterion). Results of the first five minutes validity measures showed no change in validity decisions for the Garmin Fenix 5, Jabra Elite Sport, and Motiv Ring, confirming the regulation is sufficient time to determine heart rate validity in longer trials. Results of the 10% data removal revealed at that range the Garmin Fenix 5 had 56% of the data removed, for the Jabra Elite sport 38% was removed, For the Movti Ring 52% was removed, and for the Scosche Rhythm+ 12% of the data was removed. Wearable devices worn on the wrist (Garmin Fenix 5), finger (Motiv Ring), and ear (Jabra Elite Sport) have poor heart rate performance in outdoor settings. While the forearm device (Scosche Rhythm+) had the best heart rate performance, we still express caution as there is still error associated with the device.

Table of Contents

Introduction

The usage of wearable technology has been declared a top trend since 2016 by fitness professionals (38). Wearables have evolved from pedometer capabilities, to now estimating metrics like energy expenditure and lactate threshold (5; 43). With over 10,000 device options (22) and advances in technology, these devices range from analyzing sleeping habits, repetition counting, excess post oxygen consumption (EPOC), and VO2max (3-5).

Heart rate (HR) is a common metric obtained from wearable devices. In laboratory settings, a variety of wearable technology devices have been reported to have agreement with criterion devices (8). Wearable devices that provide heart rate measures such as smart bras (28), earbuds (8; 28), forearm sensors (16; 35), and wrist watches (2; 7) have been reported to provide valid HR data. The protocol to validate different devices is typically done using treadmill running protocols with HR measurements taken concurrently from wearable devices and electrocardiogram (ECG) (16; 17; 37; 41) or a chest strap (8; 12; 34-37) as a criterion measure. Other validation protocols analyzing metrics like energy expenditure use HR in conjunction with formulas and body composition for estimation (15; 23; 32). Though these devices have good HR agreement in laboratory settings, when wearable devices are used in applied settings, the devices have been reported to have poor heart rate agreement with a criterion measure (17; 27; 37).

Laboratory validation trial lengths for wearable technology typically last five minutes (11; 13; 17; 29) due to Consumer Technology Association (CTA) guidelines (1). Though this present study focuses on HR, the inclusion of other wearable validation protocols utilizing the metrics listed above, aims to demonstrate the widely accepted guidelines which are used in varying protocols for wearable testing.

- For example, in laboratory validations of energy expenditure for the Fitbits One, Zip, Flex, and Jawbone UP24 during walking, jogging, cycling, and a stair step exercises, researchers asked thirty participants to perform tasks for five minutes (29). Overall, devices were reported to be inaccurate in estimating energy expenditure (Appendix 1, Table 1, Row 1) since accuracy varied from task-to-task between devices.
- More specifically, when analyzing step count during ambulation and walking activities, thirty participants exercised in five-minute trials wearing the Fitbit Flex, the Garmin Vivofit, and the Jawbone UP (11). It was reported that accuracy depends on movement type, though contrastingly they found that overall, the devices provided accurate data (Appendix 1, Table 1, Row 2).
- Laboratory validations of the Biak Peak and Fitbit Charge HR also included five-minute trials where twenty-four participants performed cycling, walking, jogging, running, resisted arm raises, resisted lunges, and isometric plank exercises (18). In that study, it was reported that both devices provided accurate HR data during lower intensity activities and at rest when compared to a 12-lead electrocardiogram (ECG) monitor (Appendix 1, Table 1, Row 3).
- In a laboratory validation testing the accuracy of both energy expenditure and HR of the Apple Watch Series 4, Polar Vantage V, Garmin Fenix 5, and Fitbit Versa, twenty-five participants exercised in five-minute trials while sitting, walking, and running (13). The authors report that at light to vigorous intensities the Apple Watch Series 4, Polar Vantage V, Garmin Fenix 5, and Fitbit Versa had good HR agreement when compared to the criterion 12-lead ECG, and the Apple Watch Series 4 and Polar Vantage V having high HR agreement at all intensities (Appendix 1, Table 1, Row 4).

It is unknown whether evaluating only the First Five Minutes of exercise bouts extending to longer durations would return a different decision with respect to device agreement.

HR data removal has been used in validation studies to clear ‘noise’ from devices, due to faulty technology, and poor skin connection (9; 16; 40; 42). In this context, ‘noise’ is erroneous data that is not representative of HR. Examples of data removal during wearable validation protocols are described in this section.

- The evaluation of earbud and wristwatch devices during aerobic exercise and resistance training led to a loss of three participants total data and a removal of 6,270 out of 125,400 data points (5%) due to poor connection of devices (9). Twenty-two participants performed three bouts of exercise, 30-minute treadmill, 25-minute HIT training, and 40 minutes of an outdoor walking/running activity while wearing the Jabra Elite Sport and the Mio Alpha wristwatch (Appendix 2, Table 2, Row 3). A description of the data removed from specific devices was not presented with the findings (9).
- In another study which looked at the accuracy of HR monitors during aerobic exercise 15 data points out of 4,000 were removed (0.4%) due to devices failure to capture HR (16). Fifty participants exercised for a total of 24 minutes on a treadmill, stationary bike, and an elliptical trainer while wearing Scosche Rhythm+, Apple Watch, Fitbit Blaze, Garmin Forerunner 235, and TomTom Spark Cardio (Appendix 2, Table 2, Row 1). A description of the data removed from specific devices was not presented.
- Evaluation of the Hexoskin wearable vest lead to four trials removed due to poor HR quality, eight trials removed due to poor accelerometer quality, and four trials lost due to firmware update (40). Twenty participants wore the Hexoskin vest during ambulatory activities such as lying, sitting, standing, and during a walking exercise (Appendix 2, Table 2, Row 4).

- An evaluation of wrist worn devices during a treadmill exercise led to 27 out of 1773 data points removed (1.5%) due to either an inability to complete the treadmill protocol or loss of skin contact with the device (42). Fifty participants completed an 18-minute treadmill protocol while wearing Charge HR (Fitbit), Apple Watch (Apple), Mio Alpha (Mio Global), and Basis Peak (Basis) (Appendix 2, Table 2, Row 2). A description of the data removed from specific devices was not presented with the findings (42).

While data removal is a practice employed by some investigators, the effect of removing data from a data set on device accuracy has not been determined or reported to date in the literature.

The authors of the present study have access to an existing data set that recorded HR during an outdoor trail run (average trial time = 21:56±5:38 min; trial distance 2-miles) using multiple HR wearable devices that recorded data concurrently (27). This data set was originally published, and it was determined that the Garmin Fenix 5 photoplethysmogram (PPG) watch, Jabra Elite Sport, Motiv Ring, and Scosche Rhythm+ were invalid. However, it is not clear if a different analysis using only a subset of data as recommended by the CTA (i.e., First Five Minutes) if the validity of the devices could be determined. Therefore, the purpose of this study was to use the existing data set published by Navalta et al. (2020) to determine if the wearable devices yielded valid HR data following the CTA five-minute guideline for assessing validity. A secondary purpose of this study was then to explore whether a data removal algorithm would change the determination of HR validity.

Our hypothesis was that when analyzing the First Five Minutes of an outdoor trail all devices previously determined to have poor heart rate agreement would be valid. Secondarily, we hypothesized that removing data outside of a 10% range (above and below the criterion) would result in devices returning an acceptable decision with respect to validity. The 10% range was chosen because this is a commonly reported validity threshold for Mean Absolute Percent Error in field testing investigations (9; 17; 30).

Methods

The procedures to collect the data have been previously reported by Navalta et al. (2020). The procedures for data processing using the First Five Minutes and data removal techniques are methods original to this paper and described in detail in this section.

Data set

The data set consisted of HR data recorded during trail running Navalta et al. (2020). Data from twenty-one healthy participants were included in the data set (female n = 10; male n = 11). With the (mean±SD): age = 31±11 years, the mean height = 173.0±6.9 cm, and the mean mass = 75.6±12.9 kg.

The data set consisted of HR data recorded during a self-paced 3.22km (2mile) trail run at: (McCullough Hills Trail, Henderson, NV), (Three Peaks Trail, Cedar City, UT), and (Bristlecone Trail, Mt. Charleston, NV). With elevation changes of 48m, 55m, and 104m respectively. The data set includes HR data that were recorded concurrently using the following devices:

- The Garmin Fenix 5 wristwatch (Garmin Ltd, Olathe, KS) was fitted around the left wrist, and uses the Garmin ElevateTM multi-sensor monitor to measure heart rate, the dimensions are 47 x 47 x 15.5 mm.
- The Jabra Elite Sport Earbuds (Jabra, Copenhagen, Denmark) were fitted in the ear canal, the dimensions are 120 x 45 x 179 mm.
- The Motiv ring (Motiv Inc, San Francisco, CA) was secured to the finger with the best fit (size 10mm), which was self-selected by the participants. The ring is made of a titanium alloy, with a green LED optical heart rate sensor.

- The Scosche Rhythm+ forearm band (Scosche Industries Inc., Oxnard, CA) was placed around the forearm with a neoprene strap, sensor size is 54.4 x 48.8 x 14.7 mm.
- The criterion measure was obtained using the Polar H7 heart rate monitor (Polar Electro, Kempele, Finland), which uses a sensor that is placed at the sternum and secured around the torso. The Polar H7 contains a 1000 Hz sample rate and has high agreements with ECG measurements in treadmill, cycle, and elliptical exercises with a Lins' concordance correlation coefficient = 0.99 each (16).

The data received by these devices were reported on a second-by-second basis for the Garmin Fenix 5, Jabra Elite Sport, and Scosche Rhythm+, and on a minute-by-minute basis for the Motiv Ring. The sampling rates for the devices were not released by their respective companies. The data received by the criterion Polar H7 was reported on a second-by-second basis and the sampling rate was taken at 1000Hz.

<u>Data Processing</u>

The data set was processed using two algorithms: First Five Minutes and Data-Removal. All processing was executed in Microsoft Excel, Microsoft Corporation (2018).

First Five Minutes

The First Five Minutes of data was extracted for each device and used for analysis, this was done by lining the criterion and the device then removing cells after 300 as that is equivalent to five minutes. This process was repeated for each of the four devices and for all twenty-one participants creating 84 new data sets totaling 19,005 data points.

Data-Removal

Obtaining the data set for analyzing 10% data removal are as follows.

- Each cell was lined up with the criterion Polar H7 by their respective reporting rates. To obtain the new data set the values of 10% above and 10% below the criterion had to be attained. This was done by applying a formula to each of the cells, the formula for 10% above was: (first cell containing Polar H7 value*10/100 + same Polar H7 cell), and the formula for 10% below was: (first cell containing Polar H7 value*10/100 – same Polar H7 cell). The values were specific to that data point and were repeated for each data point in that respective set. The new values represent the upper and lower limits of the accepted data range.

- The values outside of that range were found with a conditional format which highlighted cells if the value of the cell was equal to or greater than/less than the calculated 10% value. The conditional format used to find the values equal to or greater than was: =AND(cell containing device>=cell containing values 10% or greater). The conditional format used to find the values equal to or less than was: =AND(cell containing device<=cell containing values 10% or lesser).
- Those values were separated from the data set by a filter which removed the highlighted values. Once the values were filtered, the column was copied to an adjacent column and the filter was removed. That created the new data set which only contained values within 10% above and 10% below the criterion. The new set also contained blank cells where those values were removed.
- The last portion of this process was matching the appropriate corresponding criterion Polar H7 values. That was done by taking a column containing the Polar H7 values and applying a conditional format which reported the cell if the corresponding cell in the device column was blank, =AND(ISBLANK(first cell in device column)). This highlighted the corresponding data points which were removed from the device, and a similar filter was applied to the Polar H7. The highlighted values were removed leaving the values that matched the device, the values were then moved to a separate column.

- The columns now contained the values specified in the range - as well as the blank cells. To remove the blank cells a search function in the editing group was applied. This was done by selecting ‘blanks’ in the ‘Go To Special…’ menu. That selected all the blank values within the sheet allowing the delete function in the ‘cells’ menu to be used. The new data sets only contained values 10% above and 10% below the criterion and the corresponding criterion value for that time point. This extensive process of lining was chosen to ensure that the values for exact time points were accurate.

The algorithm described above was conducted for all devices and associated data points. Prospective device data points were as follows:

- The Garmin Fenix 5 had a total of 11,393 data points on a second-by-second basis.
- The Jabra Elite had a total of 30,663 data points on a second-by-second basis.
- The Motiv Ring had a total of 377 data points on a minute-by-minutes basis.
- The Scosche Rhythm+ had a total of 11,393 data points on a second-by-second basis.

The total amount of data points analyzed between all four devices was 53,826 points. The data removed will be represented as a percent and the remaining values will be analyzed with the traditional validity measures listed below.

Statistical analysis

Statistical analyses followed the procedures described by Navalta et al. (2020). Specifically, tests to determine validity are as follows: Mean Absolute Percent Error (MAPE), Lin’s Concordance Coefficient (r_C), and Intraclass Correlation (ICC). The inclusion of MAPE was used to determine the trending accuracy of the devices, measuring the error, and as a regression analysis. The inclusion of Lin’s Concordance Coefficient (r_C) allowed us to compare a standard (Polar H7) to the tested devices, quantifying the agreement. Our determining values for validity were a MAPE of 10% or lower in conjunction with a r_C of at least 0.9.

These statistical procedures were used to analyze the First Five Minutes data set and then again for the Data-Removal data set.

Results

Figure 1 Is a representation of an individual's data without any removal.

First Five Minutes

As a result of the First Five Minutes process MAPE for the Garmin Fenix 5 increased from 13.5% to 14.1%, and the r_C decreased from 0.316 to 0.304. For the Jabra Elite Sport that process resulted in a decrease in MAPE from 21.3% to 15.1% and increase in r_C from 0.384 to 0.574. For the Motive Ring that process resulted in an increase in MAPE from 15.9% to 18.8%, and a decrease in r_C from 0.293 to 0.197. For the Scosche Rhythm+ that process resulted in a decrease in MAPE from 5.6% to 4%, and an increase in r_C from 0.780 to 0.913.

Tables 1-4 presents the summary of the statistical tests for validity for the First Five Minutes data set and Navalta et al. (2020) published results (i.e., original data set and analysis). Figure 2 Is a representation of an individual's data after the First Five Minutes removal.

Data Removal

For the Garmin Fenix 5 6,360 values were excluded, this converts to 56% of the data taken out. As a result, MAPE decreased from 13.5% to 4.2% and r_C increased from 0.316 to 0.942 (Data is shown in table 1, column 4). For the Jabra Elite Sport 11,575 values were excluded, this converts to 38% of the data taken out. As a result, MAPE decreased from 30% to 3.1% and r_C increased from 0.384 to 0.976 (Data is shown in table 2, column 4). For the Motiv Ring 197 values were excluded, this converts to 52% of the data taken out. As a result, MAPE decreased from 15.9% to 4.7% and r_C increased from 0.293 to 0.907 (Data is shown in) table 3, column 4.

For the Scosche Rhythm+ 1,398 values were excluded, this converts to 12% of the data taken out. MAPE decreased from 7.3% to 2.1% and r_C increased from 0.78 to 0.988 (Data is shown in table 4, column 4).

Tables 1-4 presents the summary of the statistical tests for validity for the Data Removal set and Navalta et al. (2020) published results (i.e., original data set and analysis). Figure 3 is a representation of and individuals' data after the Data Removal.

Table 1. Comparison of validity measures from participants (N=21) who wore a concurrent criterion device during a 2-mile trail run.

Garmin Fenix 5			
	Data Removal	*Navalta et. al, 2020*	*First Five Minutes*
Heart Rate ([SD] bpm)	147 (23)	143 (24)	141 (21.3)
MAPE %	4.2	13.5	14.1
MAE (bpm)	6.2	20.8	21.5
rc (95% CI)	0.942 (0.934 to 0.945)	0.316 (0.305 to 0.326)	0.304 (0.288 to 0.319)
ICC (p-value)	0.357(<0.001)	0.415 (<0.001)	0.409 (<0.001)

Data expressed as mean (SD). MAPE = Mean absolute percent error, MAE = Mean Absolute Error, rC = Lin's Concordance Correlation Coefficient, ICC = Intraclasscorrelation coefficient.

Table 2. Comparison of validity measures from participants (N=21) who wore a concurrent criterion device during a 2-mile trail run.

Jabra Elite Sport			
	Data Removal	*Navalta et. al, 2020*	*First Five Minutes*
Heart Rate ([SD] bpm)	149 (31)	126 (58)	144 (37)
MAPE %	3.1	21.3	15.1
MAE (bpm)	4.9	30	21.6
rc (95% CI)	.976 (0.975 to 0.976)	0.384 (0.377 to 0.390)	0.574 (0.560 to 0.588)
ICC (p-value)	0.976 (<0.001)	0.395 (<0.001)	0.526 (<0.001)

Data expressed as mean (SD). MAPE = Mean absolute percent error, MAE = Mean Absolute Error, rC = Lin's Concordance Correlation Coefficient, ICC = Intraclasscorrelation coefficient.

Table 3. Comparison of validity measures from participants (N=21) who wore a concurrent criterion device during a 2-mile trail run.

Motiv Ring			
	Data Removal	*Navalta et. al, 2020*	*First Five Minutes*
Heart Rate ([SD] bpm)	155 (19)	137 (27)	129 (27)
MAPE %	4.7	15.9	18.8
MAE (bpm)	7.1	25.1	30.2
rc (95% CI)	0.907 (0.88 to 0.928)	0.293 (0.225 to 0.358)	0.197 (0.07 to 0.319)
ICC (p-value)	0.201 (<0.001)	0.287 (<0.001)	0.199 (<0.001)

Data expressed as mean (SD). MAPE = Mean absolute percent error, MAE = Mean Absolute Error, rC = Lin's Concordance Correlation Coefficient, ICC = Intraclasscorrelation coefficient.

Table 4. Comparison of validity measures from participants (N=21) who wore a concurrent criterion device during a 2-mile trail run.

Scosche Rhythm+			
	Data Removal	*Navalta et. al, 2020*	*First Five Minutes*
Heart Rate ([SD] bpm)	155 (20)	149 (29)	146 (32)
MAPE %	1.4	5.6	4
MAE (bpm)	2.1	7.3	5.7
rc (95% CI)	0.988 (0.987 to 0.988)	0.780 (0.774 to 0.786)	0.902 (0.9 to 0.906)
ICC (p-value)	0.988 (<0.001)	0.120 (<0.001)	0.913 (<0.001)

Data expressed as mean (SD). MAPE = Mean absolute percent error, MAE = Mean Absolute Error, rC = Lin's Concordance Correlation Coefficient, ICC = Intraclasscorrelation coefficient.

Figure 1. Is a comparison of the devices complete data set to the criterion across time

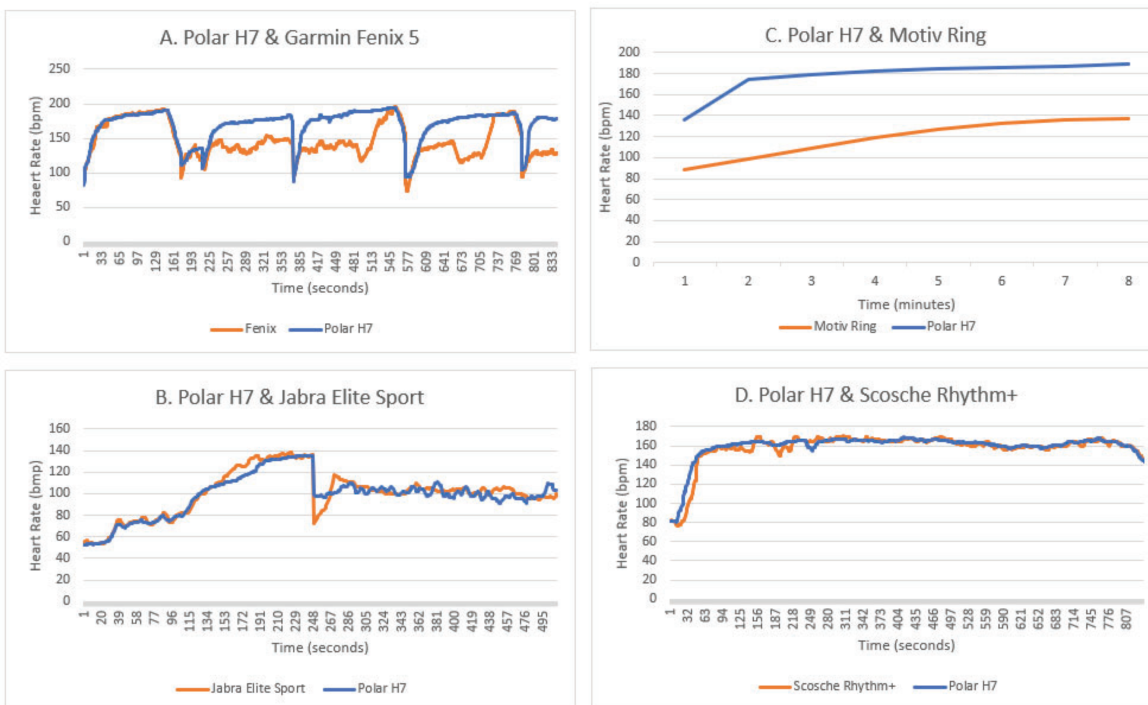


Figure 2. Is a comparison of the devices five minute data to the criterion across time

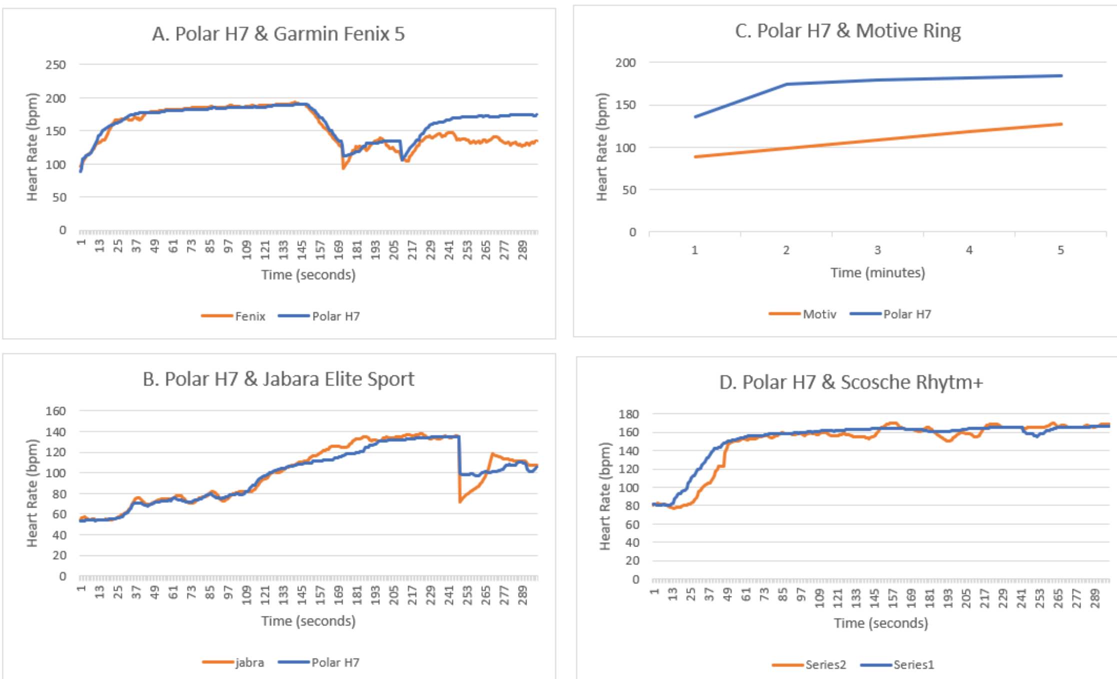
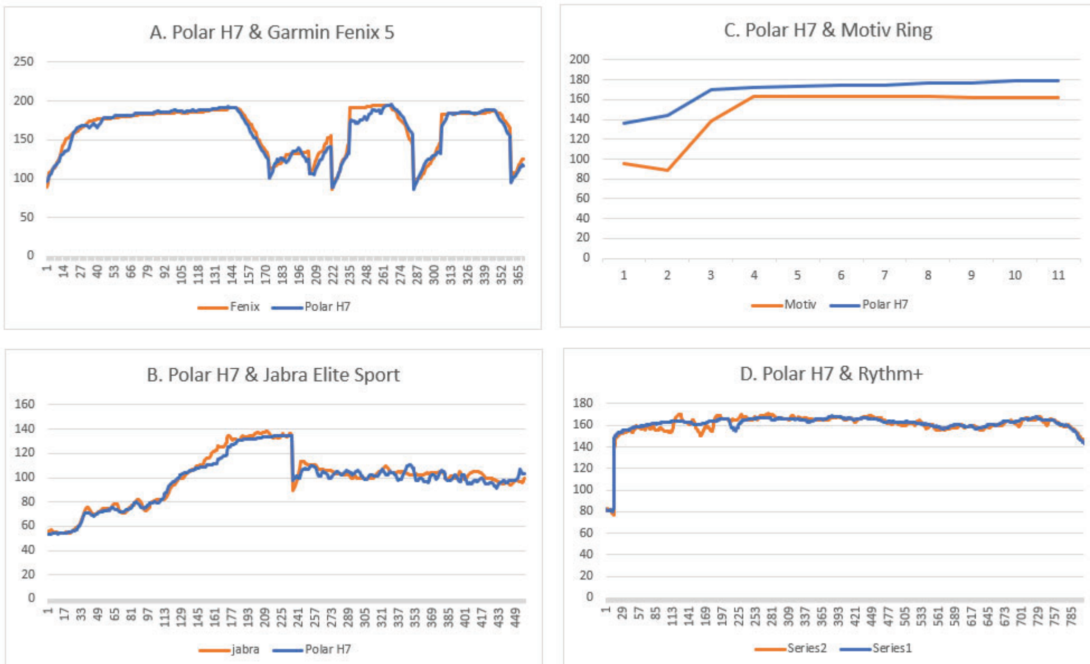


Figure 3. Is a comparison of the devices data removal to the criterion across time



Discussion

The purpose of this study was to determine if the CTA five-minute guidelines for wearable technology testing were sufficient to determine HR validity in devices being tested for longer trials. We hypothesized that analyzing the First Five Minutes of an outdoor trail run would result in devices previously determined to have poor HR validity to be determined as valid. We found that this limited time frame was sufficient in determining validity in longer trials, as decisions regarding validity were not affected in three out of the four devices. The second purpose of this study was to observe differences in HR validity when data was removed from the devices at 10% above and 10% below the criterion. We hypothesized that once data outside of 10% above and 10% below the criterion was removed the previously invalidated devices would be seen as valid. Working within 20% of the criterion measure through data removal allowed each of the devices to be considered valid using our predetermined thresholds. Unfortunately, between 12-53% of the dataset for individual devices had to be removed in order to allow for this agreement. The discussion below further analyses the points of the First Five Minutes and Data Removal processes. It will also discuss a comparison of the performance of these devices reported by other studies, the Navalta et al., 2020 original analysis, and the manipulations of the current study.

First Five Minutes

The First Five Minutes algorithm was based on the CTA guideline stating that trials for testing wearable technology need only last five-minutes (1). Studies which used the minimum requirement analyzed activities such as: walking, jogging, running, cycling, resistance training, and household activities (11; 13; 17; 29). To our knowledge this is the first paper to observe how the minimum requirement interacts with HR validity in wearable devices. We hypothesized the time was not sufficient for validity testing in longer trials, proposing that the accumulation of connection issues leading to poor HR data which occur over longer trials were not captured in the time limitation. The results of the present investigation reveal that the five-minute limitation is sufficient to determine HR validity in longer trials as three out of the four devices did not result in a change with respect to decisions regarding HR validity. The devices that were not affected by the First Five Minutes analysis included the Garmin Fenix 5, Jabra Elite Sport, and the Motiv Ring (table 1. column 2, table 2. column 2, table 3. column 2, respectively). The Scosche Rhythm+ was the only device whose validity decision was affected by analyzing the First Five Minutes (table 4. column 2). The validity decision of the Scosche Rhythm+ in the Navalta et al. (2020), paper was influenced by the r_C = 0.780, as the MAPE was well within the accepted values (MAPE = 5.6%). A possible reasoning as to why the First Five Minutes algorithm was able to increase validity in the Scosche Rhythm+ could be because at higher intensities PPG devices become less accurate (16; 18; 19; 21). By removing most of the trial we could have negated the effects of this phenomenon, giving the Scosche Rhythm+ the incremental improvement it needed to meet our predetermined validity testing thresholds. Analyzing trials in five-minute increments could provide insight into this occurrence. Furthermore, a more comprehensive breakdown of devices into respective segments could lead to a better understanding of the performance of these devices throughout different phases of exercise - warm-up, intensity/exhaustion, recovery.

<u>Data Removal</u>

To the best of our knowledge this is the first paper excluding data with the express purpose of evaluating the effect on determining device validity. The exclusion of data allowed us to understand the effects of data removal on the accuracy of HR devices. We hypothesized the elimination of values not within 10% above and 10% below the criterion would allow for a different validity interpretation. Validity was achieved in the Garmin Fenix 5, Jabra Elite Sport, Motiv Ring, and the Scosche Rhythm+ (table 1, column 1; table 2, column 1; table 3, column 1; table 4., column 1, respectively). This study aimed to understand how data removal affected HR validity, in doing so, revealed a way to describe HR performance at levels of acceptance. At a 10% acceptance level the Garmin Fenix 5 required 56% of the data to be removed. For the Jabra Elite Sport at a 10% acceptance level 38% of the data will be removed, and for the Motiv Ring and Scosche Rhythm+ 52% and 12% was removed at a 10% acceptance rate.

While the removed data percentage and the associated validity measures cannot determine overall validity of the device, it can begin to characterize the data. Exploring the ranges of acceptance rate in wearable devices in conjunction with the percent of data removed show the extent of manipulation needed to determine validity. In the case of the Scosche Rhythm+ the accepted range can be lowered as the device showed high HR agreement (MAPE = 1.4%, r_C = 0.988) with a lower percentage of data removed (12%). Acceptance rates of 7% or 5% would show an increase of data removal, but it is unclear as to how much of data would be removed. For the Garmin Fenix 5 and the Motiv Ring over half of the data was removed (56%, 52%) displaying a need to increase the accepted range. The range at which these devices would mirror the performance of the Scosche Rhythm+ is also unclear. To achieve data removal percentages of 12% in those devices will require higher ranges of allowed error. A deeper analysis into how the ranges affect the devices will characterize the performance of these devices. Validity studies measuring wearable devices should include a characterization table which shows the amount of removed data at their acceptance.

While analyzing the data removal we noticed a delay in the capture of HR at varying points. The device displayed a ‘lag’ when compared to the criterion Polar H7 and in some instances did not record sudden jumps in HR. These random occurrences of temporarily elevated HR have more opportunity in field testing than laboratory and could be a factor as to why performance between the settings are different. Temporal factors such as speed, humidity, air resistance, and temperature are monitored and controlled. Perturbations in field environments such as obstacles, grade increases/decreases, or volition may cause sudden changes in HR and intensity. In laboratories HR is steadily increased and is less likely to incur rapid changes in intensity. Evaluating the influence of factors discussed is beyond the scope of this paper and requires further analysis.

Individual Devices

The discussion below compares the performance of the devices in published findings during other research articles, which included metrics outside of HR, to the data published by Navalta et al. (2020), and the application of the algorithms from the current study. This approach was designed to provide relevant information of the tested performance of each of the devices.

Wrist Device – Garmin Fenix 5

The Garmin Fenix 5, despite the popularity of the Garmin brand and the Fenix line has not been tested heavily. One publication of a laboratory-based protocol of the Garmin Fenix 5 was found (13), and two published conference abstracts were found for the Fenix 3 (10; 19).

- Results from a laboratory-based testing of the Garmin Fenix 5 showed a moderate heart rate validity when compared to the criterion during sitting and walking (standardized typical error of the estimate [sTEE] = 0.63, 0.62) (13). They also reported poor heart rate validity when the intensity of the exercise was increased, expressing concern for higher rates of error (9.9 MET level [sTEE] = 1.24, 13.8 MET level = 1.44) (13).

- The first published abstract for the Garmin Fenix 3, tested the estimation of maximal aerobic capacity compared to a laboratory-based metabolic cart, finding that there was no difference in metabolic analysis ($p>0.05$) (10).
- The second abstract published findings on the biomechanical running parameters of the Garmin Fenix 3 (19). They reported that there was no difference in stride length and run cadence when compared to laboratory measurements ($p>0.05$) but reported a significant difference in vertical oscillation and ground contact time ($p<0.05$) (19).

During the Navalta et al. (2020), trail run the Garmin Fenix 5 was reported to have poor heart rate agreement with the Polar H7 (MAPE = 13.5%, bias =15.9 bmp, r_C = 0.316). This finding adds to the collection of validation studies observing poor PPG wrist device HR agreement at varying intensities (18; 26; 35; 37). With the application of the five-minute algorithm we were unable to manipulate the Garmin Fenix 5 into validity. The five-minute protocol suggested by CTA guidelines is sufficient time to determine the overall validity of the device as the application of this guideline did not change our view of validity (MAPE = 14.1%, r_C = 0.304). The application of the data removal algorithm allowed the new Garmin Fenix 5 dataset to be considered valid according to predetermined thresholds (MAPE = 4.2%, r_C = 0.942). The application inherently created an acceptable MAPE as it limits the data to a range within 20%, the r_C or correlation value is what determined the validity of the device within the new data set. The Garmin Fenix 5 met our threshold for validity, though in order to manipulate the data into validity 56% was removed. Out of all the devices the Garmin Fenix 5 had the most data removed, and in both settings, laboratory and field, the Garmin Fenix 5 has shown poor HR agreement with criterions (13; 26). This, in conjunction with the results of the data removal show that HR for the Garmin Fenix 5 is not reliable in field settings.

Earbud Device- Jabra Elite Sport

Research done on earbud devices during various exercises have shown positive results.

- In a study looking at the Bose SoundSport Pulse (BSP) headphones during resistance training and during a graded cycle exercises test they found positive HR agreement during resistance training (MAPE = 6.24%) (6). When tested during the cycle exercise graded test, they found that when intensity increased heart rate validity decreased (MAPE at 50W = 6.4%, MAPE at 200W = 15.42%) (6).
- During testing of the Jabra Elite Sport in a graded treadmill exercise test they found that the Jabra Elite Sport had good correlation with the criterion during both aerobic (bias = 0.8 bpm, MAPE = 2.48%, r_C = 0.943) and anaerobic exercise (bias = -3.6, MAPE = 3.53%, r_C = 0.861) (9).
- Testing on the PerformTek earbud sensor during a graded treadmill exercise test showed positive HR agreement with the criterion (bias = -0.2%, R2 = 0.98) (21).
- In a study where the Jabra Sport Pulse was tested on cardiac patients, they found that HR agreement declined when HR rose above 100 bpm (average difference to true heart rate = 20.3 bpm, r2 = 0.434) (20).

Navalta et al. (2020), found that during an outdoor trail run the Jabra Elite Sport had poor heart rate validity (bias = 19.2, MAPE = 30%, r_C = 0.384) (26). In our findings HR validity for the Jabra Elite Sport was not affected when the data was limited to the First Five Minutes (MAPE = 15.1%, r_C = 0.574). Therefore, the five-minute protocol suggested by CTA guidelines is sufficient time to determine the overall validity of the device as the application of this guideline did not change our view of validity. The application of the data removal algorithm allowed the new Jabra Elite Sport data set to be considered valid according to predetermined thresholds (MAPE = 4.9%, r_C = 0.976). Though the Jabra Elite Sport met our threshold of validity 38% of the data was removed. The literature suggests that when heart rate variability is high and at higher intensities earbud devices have poor heart rate agreement (6; 20; 26). This in conjunction with our data manipulation suggests that HR measurements for the Jabra Elite Sport are not reliable in field settings.

Ring Device – Motiv Ring

Research done on ring devices has mainly been seen as conference abstracts (4; 24; 33) with one published paper measuring the performance of the Motiv Ring during exercise (21).

- One abstract utilized a ring prototype as a mode for maneuvering through varying running applications, where differing motions of swiping or tapping functioned as lap counting, distance counting and music control (4).
- In another abstract looking at relative weight of watch, bracelet, and ring trackers found that the ring was the most suitable for everyday life as it was less cumbersome (24).
- The last abstract tested the development of a wearable ring-shaped biosensor during rest, mental stress testing, and physical stress testing which measured galvanic skin response and HR variability (N=4) (33). They found that the parameters for galvanic skin response and HR variability were within 10% of the criterion.

Navalta et al., 2020 found that out of all the devices tested during an outdoor trail run the Motiv Ring had the poorest heart rate agreement (MAPE = 15.9%, bias =21.7 bpm, r_C = 0.293). They do note some limitations. Record of the finger to which the device was fitted was not taken, as well as the size of ring being limited to 10mm for all participants.

In our findings HR validity for the Motiv Ring was not affected when the data was limited to the First Five Minutes (MAPE = 18.8%, r_C = 0.199). Therefore, the five-minute protocol suggested by CTA guidelines is sufficient time to determine the overall validity of the device as the application of this guideline did not change our view of validity. The application of the data removal algorithm allowed the new Motiv Ring data set to be considered valid according to predetermined thresholds (MAPE = 4.2%, r_C = 0.942). Though the Motiv Ring met our thresholds for validity 52% of the data was removed. Little is known about the performance of ring devices as there has only been one published paper which tested the Motiv Ring in a field setting, currently there are no published papers looking at the performance of ring devices in laboratory settings. From current investigations and manipulation of the data HR measurements of the Motiv Ring are not reliable in field settings.

Forearm Device – Scosche Rhythm+

Research on forearm wearables have been tested in activities such as virtual reality ship handling simulators (3), tag (25), driving in differing weather conditions (44), and triathlon training (31). Only three studies testing the Scosche Rhythm+ during exercise tests were found (16; 26; 41).

- In a study looking at varying exercises the Scosche Rhythm+ agreed with the criterion ECG at rest (r_C = 0.93), during a cycle and treadmill exercise (r_C = 0.84, 0.92), but not during an elliptical exercise (r_C = 0.41 with arms, 0.27 without arm movement) (16).
- In another study they found no difference when comparing the Scosche Rhythm+ to the Polar H7 in a 5-minute treadmill walking and running test ($p>0.05$) (41).

Navalta et al. (2020) determined that during a trial run the Scosche Rhythm+ performed better while participants were running downhill (MAPE = 3.8%, bias = 1.9 bpm, r_C = 0.885) than when running at an incline (MAPE = 6.2%, bais = 3.9 bpm, r_C = 0.699). In our findings HR validity for the Scosche Rhythm+ was affected when the data was limited to the First Five Minutes (MAPE = 5.7%, r_C = 0.902). Therefore, the five-minute protocol suggested by CTA guidelines was not sufficient time to determine the overall validity of the device as the application of this guideline did change our view of validity. The application of the data removal algorithm allowed the new Scosche Rhythm+ data set to be considered valid according to predetermined thresholds (MAPE = 1.4%, r_C = 0.988).The Scosche Rhythm+ met our threshold for validity, though 12% of the data was removed. The literature once again suggests that devices, now including forearm, have poor HR agreement when intensity begins to rise. Literature on the Scosche Rhythm+ and other forearm wearable devices suggest that they can be seen as reliable in all settings but have issues with intensities (16; 26; 41). The Scosche Rhythm+ had the least amount of data removed from the set at 12% though we express caution of reliance on HR measures in field settings as there is still error associated with the device.

Device Placement

Each of the devices are paired with a detection photoplethysmography system and an algorithm which takes raw measurements compiling them into a format that is designed to aid in device accuracy and provide user friendly information. We took the raw data files and created our own algorithm to understand why these devices when tested outdoors were considered less valid. Analyzing the raw data reported by the devices allowed the true performance of the sensors to be seen. All tested devices use similar PPG sensors, though we observed extreme differences in performance. Since the data removal algorithm resulted in all devices meeting our threshold for validity, a greater performance is based on the amount of data removed, where the less data removed equates to better performance. Since the five-minute protocol aimed to determine the usefulness of the CTA guideline it will not be a factor in discussing the change in performance of the devices. In our findings we observed better performance in the forearm device when compared to the placement at the wrist, ear, and finger. Factors that influence this could be due to the shape of the forearm aligning with the convex shape of the sensor housing unit. Forearm devices can be subject to less arm movement during activities such as running, as the elbow can experience less movement than when compared to the wrist or finger. Studies have reported poor performance due to artifact movement (14), loss of data through poor device skin connection (9; 16; 40; 42), and through light interaction of the sensors (39). Thus, the shape of the device in comparison to the placement on the body is of importance. This cannot be the only factor leading to better HR agreement as ring and earbud devices mold to the surface of the body. The data removal algorithm resulted in identifying 3,881 zero values for the Jabra Elite Sport earbud, equating to about 34%. During activity movement of the head can shift rapidly and the impact of movements like running can cause moments of disconnection which could explain the device's registration of a zero HR. In the case of the Motiv Ring, the placement of the device on specific fingers could have influenced the performance as well as the overall fit of the ring. The device is constructed from an inflexible material which limits the user to adjusting for knuckle size over phalanx diameter. Poor circulation to the fingers could also be a confounding factor for HR performance in ring devices. The literature discussed previously surrounding forearm devices, and the outperformance within this study, suggests that the placement of wearable PPG devices on the forearm yield the most accurate HR measurements when compared to the wrist, ear, and finger (16; 26; 41).

Conclusion

We found that CTA guidelines for trial time in wearable technology are sufficient in determining validity in longer trials, with three out of the four devices having no change in validity decisions. We also noted the placement of devices on the body possibly having an influence on HR agreement, with the forearm device having the best overall performance. We showed that HR data removal of the Fenix Garmin 5, Jabra Elite Sport, Motiv Ring, and Scosche Rhythm+ at 20% range will cause validity according to our thresholds, though the amount of data removed was substantial in the cases of the Garmin Fenix 5 (56%), Jabra Elite Sport (38%), and the Motiv Ring (52%). The rates of acceptable data removal are undetermined, in the cases listed above achieving less data removal would require increasing the percentage of allowed error, this would likely lead to poor validity measures. This process was extensive and is unrealistic for users who wish to have accurate information about their daily performance. These devices show poor performance in field testing and are not reliable for HR measurement

Appendix 1. Articles Containing Protocols Related to Five Minute Trials

Location	Authors	Title	Subjects	Devices	Activities	Time	Data Removed	Stats done	Results
LABORATORY	Gillinov, S., et. al (2017)	Variable Accuracy of Wearable Heart Rate Monitors during Aerobic Exercise	N = 50 (54% Female) Age 38 ± 12	C: Polar H7 Scosche Rhythm+, Apple Watch, Fitbit Blaze, Garmin Forerunner 235, and TomTom Spark Cardio	treadmill, a stationary bicycle, Elliptical	total trial length: 24 min exercise time: 18 min break-dow: 4.5min work/2min rest	Total Points: 4,000 Removed: 15 **does not specify device**	rc, Bland Altmand	Apple Watch rc = 0.92 TomTom Spark Cardio rc = 0.83 Garmin Forerunner 235 rc = 0.81 The Scosche Rhythm+ rc = 0.75 Fitbit Blaze rc = 0.67
LABORATORY	Wang, R., et. al (2017)	Accuracy of Wrist-Worn Heart	N= 50 (58 % Female) Age 37 ± 11.3	C: Polar H7 Charge HR (Fitbit), AppleWatch (Apple), Mio Alpha (Mio Glob al), and Basis Peak (Basis)	Treadmill	tota trial length: 18 min exercise time: 15 min break-down: 3min at each speed	Total Points: 1773 Removed: 27 **does not specify device**	rc	Polar H7 = .99 (.987-.991) Apple Watch = .91 (.884-.929) Mio Fuse = .91 (.882-.929) Fitbit Charge HR = .84 (.791-.872) Basis Peak = .83 (.779-.865)
LABORATORY	Bunn, J., et. al (2019)	Evaluation of Earbud and Wristwatch Heart Rate Monitors during Aerobic and Resistance Training	N= 22 (16 female, 9 male) Age 25.4 ± 6.9 Height 171 ± 11 cm Mass 73.9 ± 3.1 kg Body Fat 25.2 ± 9.2%	C: Polar H7 Jabra Elite Sport, Mio Alpha	Treadmill, HIT, Outdoor Walking/Running	total trial length: 95 min exercise time: 30 min (treadmill), 25 min (HIT), 40 (outdoor acitivty)	Loss of 3 participants total data Total Points: 125,400 Removed: 6,270	rc, MAPE	** Data set too large for table**
LABORATORY	Villar, R., et. al (2015)	Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities	N = 20 (11 females, 9 male) Age 26.3 ± 5.9 years Height 171.6 ± 11.5 cm Mass 71.2 ± 12.6 kg Body mass index 24.0 ± 1.9 kg/m2	C: ECG Module, Finometer Hexoskin	Lying, Sitting, Standing, Walking	*Day one* Total Trial Length: 78 min Exercise Time: 18min *Day Two* Total Trial Length: 19 min Exercise Time: 19 min	Lost Data Sets: Poor singnal quality HR = 4 tests Accelerometer quality = 8 tests Firmware update = 4 tests	CV, Difference, ICC	1km/h = 0.52±0.55, −0.36±1.11, 0.99 3km/h = 0.69±0.63 , −0.06±1.88, 0.96 4.5km/h = 0.79±0.77, 0.25±2.17, 0.99 80%VT 0.37±0.50, 0.10±1.78, 1.00 80% HRmax 0.27±0.17, 0.72±0.68, 0.99 **CV, Difference, ICC**

***Coefficient of Variation (CV), Lin's Concordance Correlation Coefficient (rc), Mean Absolute Percent Error (MAPE), Root Mean Square Error (Difference)**

Appendix 2. Articles Analyzing Data Removal

Location	Authors	Title	Subjects	Devices	Activities	Stats done	Results
LABORATORY	Benjamin, N., et. al, (2016)	Validity of Consumer-Based Physical ActivityMonitors for Specific Activity Types	N = 30 (15 female, 15 male) Age 48.9 ± 19.4 years BMI 26.3 ± 5.2 kg/m	Fitbits One, Zip, Flex, Jawbone UP24	Sedentary Activitiy, Household Acitivity, Walking, Jogging, Cycling, Stairs	Friedman, Dunn's, MAPE, MAE, RMSE	** Data set too large for table**
LABORATORY	Chen, M., et. al, (2016)	Accuracy of Wristband Activity Monitors during Ambulation and Activities	N = 30 (15 female, 15 male) Age 21.5 ± 2.0 years Height 167.6 ± 8.9 cm Mass 60.7 ± 8.3 kg BMI 21.5 ± 1.9	Fitbit Flex, the Garmin Vivofit, and the Jawbone UP	Daily activities, Walking	APE	*Treadmill* Fitbit Flex 2.5% to 8.2% Garmin Vivofit 1.5% to 4.2% Jawbone UP 2.4% to 9.6%
LABORATORY	Edward, J., et. al, (2016)	Validation of Biofeedback Wearables for Photoplethysmographic Heart Rate Tracking	N = 24 (12 female, 12 male) Age = 24.8 ± 2.1 years Mass = 71.3 ± 9.9 kg Height = 1.66 ± 0.07 m	C: 12-lead ECG Basis Peak, Fitbit Charge HR	Cycling, Walking, Jogging, Running, Resisted Arm Raises, Resisted Lunges, Isometric Plank	Pearson Product-Moment, Bland-Altman, Mean absolute differential	Bias Peak $r = 0.92$, $p < 0.0001$, mean absolute difference 5.3 (8.3)% Fitbit $r = .83$, $p < 0.0001$, mean absolute difference 9.8 (14.0)%
LABORATORY	Düking, P., et. al, (2020)	Wrist-Worn Wearables for Monitoring Heart Rate and Energy Expenditure While Sitting or Performing Light-to-Vigorous Physical Activity: Validation Study	N = 25 (14 female, 11 male) Age 26 years Height 174 ± 10 cm Mass 70.1 ± 12 kg	C: Polar H7 Apple Watch Series 4, Polar Vantage V, Garmin Fenix 5, and Fitbit Versa	Sitting, Walking, Running	sTEE, CV, Pearsons r	Apple Watch Series = 4 0.09-0.62 Polar Vantage V = 0.13-0.88 Garmin Fenix 5 = 0.62-1.24 Fitbit = 0.47-1.94 **sTEEe*

***Absolute Percent Error (APE), Standardized Typicall Error of the Estimate (sTEE), Coefficient of Variation (CV), Mean Absolute Percent Error (MAPE), Mean Absolute Error (MAE), Root Mean Square Error**

References

1. Association, C. (2018). Physical activity monitoring for heart rate., 1-12.
2. Bai, Y., Hibbing, P., Mantis, C., & Welk, G. J. (2018). Comparative evaluation of heart rate-based monitors: Apple watch vs fitbit charge HR. *Null, 36*(15), 1734-1741. doi:10.1080/02640414.2017.1412235
3. Bassano, C., Chessa, M., Fengone, L., Isgrò, L., Solari, F., Spallarossa, G., et al. (2019). *Evaluation of a virtual reality system for ship handling simulations* SCITEPRESS - Science and Technology Publications. doi:10.5220/0007578900620073
4. Boldu, R., Dancu, A., Matthies, D., Cascón, P., Ransir, S., & Nanayakkara, S. (Oct 13, 2018). (Oct 13, 2018). Thumb-in-motion. Paper presented at the pp. 150-157. doi:10.1145/3267782.3267796
5. Borges, N. R., & Driller, M. W. (2016). Wearable lactate threshold predicting device is valid and reliable in runners. *Journal of Strength and Conditioning Research, 30*(8), 2212-2218. doi:10.1519/JSC.0000000000001307
6. BOUDREAUX, B. D., HEBERT, E. P., HOLLANDER, D. B., WILLIAMS, B. M., CORMIER, C. L., NAQUIN, M. R., et al. (2018). Validity of wearable activity monitors during cycling and resistance exercise. *Medicine and Science in Sports and Exercise, 50*(3), 624-633. doi:10.1249/mss.0000000000001471
7. Brazendale, K., Decker, L., Hunt, E. T., Perry, M. W., Brazendale, A. B., Weaver, R. G., et al. (2019). Validity and wearability of consumer-based fitness trackers in free-living children. *International Journal of Exercise Science, 12*(5), 471-482. Retrieved from https://explore.openaire.eu/search/publication?articleId=od_______267::9c1ad567d65b940b561106081f621840
8. 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 PubMed database. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/29541338

9. Bunn, J., Wells, E., Manor, J., & Webster, M. (2019). Evaluation of earbud and wristwatch heart rate monitors during aerobic and resistance training. *International Journal of Exercise Science, 12*(4), 374-384. Retrieved from PubMed database. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30899350
10. Carrier, B., Creer, A., Williams, L. R., Holmes, T. M., Jolley, B. D., Dahl, S., Weber, E., & Standifird, T. (2021). Validation of garmin fenix 3 HR fitness tracker biomechanics and metabolics (VO2max). *Journal for the Measurement of Physical Behaviour, 3*(4), 331-337. Retrieved from http://journals.humankinetics.com/view/journals/jmpb/3/4/article-p331.xml
11. CHEN, M., KUO, C., PELLEGRINI, C., & HSU, M. (2016). Accuracy of wristband activity monitors during ambulation and activities. *Medicine and Science in Sports and Exercise, 48*(10), 1942-1949. doi:10.1249/MSS.0000000000000984
12. Dooley, E. E., Golaszewski, N. M., & Bartholomew, J. B. (2017). Estimating accuracy at exercise intensities: A comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR mHealth and uHealth, 5*(3), e34. doi:10.2196/mhealth.7043
13. Düking, P., Giessing, L., Frenkel, M. O., Koehler, K., Holmberg, H., & 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 Publications Inc. doi:10.2196/16716
14. Estepp j., Ethan B. Blackford, Christopher M. Meier. (2014). Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging

photoplethysmography. *IEEE International Conference on Systems, Man, and Cybernetics (SMC),* , 1462-1469.

15. First Beat Technologies Inc. (2014). *Automated fitness level (VO2max) estimation with heart rate and speed data*

16. Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A. M., et al. (2017). Variable accuracy of wearable heart rate monitors during aerobic exercise. *Medicine and Science in Sports and Exercise, 49*(8), 1697-1703. doi:10.1249/MSS.0000000000001284

17. Jo, E., Lewis, K., Directo, D., Kim, M. J., & Dolezal, B. A. (2016). Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *Journal of Sports Science & Medicine, 15*(3), 540-547. Retrieved from PubMed database. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27803634

18. Jo, E., Lewis, K., Directo, D., Kim, M. J., & Dolezal, B. A. (2016). Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *Journal of Sports Science & Medicine, 15*(3), 540-547. Retrieved from PubMed database. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27803634

19. Jolley, B., Carrier, B., Standifird, T., & Creer, A. (2019). Validation of garmin fitness tracker metabolic data (VO2max): 191 board #29 may 29 9:30 AM - 11:00 AM. *Medicine and Science in Sports and Exercise, 51*(37)

20. Leader N, Dorian P, Lam J, Lee C, Woo A, Chow C. (2018). Evaluation of heart rate trackers in patients with
atrial fibrillation. *Canadian Journal of Cardiology., 34*((10)), :S152–S3.

21. Leboeuf, S. F., Aumer, M. E., Kraus, W. E., Johnson, J. L., & Duscha, B. (2014). Earbud-based sensor for the assessment of energy expenditure, HR, and VO2max. *Medicine and Science in Sports and Exercise, 46*(5), 1046-1052.

doi:10.1249/MSS.0000000000000183 [doi]

22. Liguori, G., Kennedy, D. J., & Navalta, J. W. (2018). Fitness wearables. *ACSM's Health & Fitness Journal, 22*(6), 6-8. doi:10.1249/FIT.0000000000000426
23. Lu, K., Yang, L., Seoane, F., Abtahi, F., Forsman, M., & Lindecrantz, K. (2018). *Fusion of heart rate, respiration and motion measurements from a wearable sensor system to enhance energy expenditure estimation* MDPI AG. doi:10.3390/s18093092
24. Marinescu, R. (2018). Multi-criteria analysis of wearable activityfitness trackers and 3D concept models of smart jewellery. *MATEC Web of Conferences, 178* doi:10.1051/matecconf/201817805017
25. Moreno, A., Poppe, R., Gibson, J., & Heylen, D. (2019). Automated and unobtrusive measurement of physical activity in an interactive playground. doi:10.17863/CAM.40470
26. 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
27. Navalta, J. W., Ramirez, G. G., Maxwell, C., Radzak, K. N., & McGinnis, G. R. (2020). Validity and reliability of three commercially available smart sports bras during treadmill walking and running. *Scientific Reports, 10*(1), 7397. doi:10.1038/s41598-020-64185-z
28. NELSON, M., KAMINSKY, L., DICKIN, D., & MONTOYE, A. H. (2016). Validity of consumer-based physical activity monitors for specific activity types. *Medicine and Science in Sports and Exercise, 48*(8), 1619-1628. doi:10.1249/MSS.0000000000000933
29. Neufeld, E. V., Wadowski, J., Boland, D. M., Dolezal, B. A., & Cooper, C. B. (2019). *Heart rate acquisition and threshold-based training increases oxygen uptake at metabolic threshold in triathletes: A pilot study*
30. O'Driscoll, R., Turicchi, J., Hopkins, M., Horgan, G. W., Finlayson, G., & Stubbs, J. R.

(2020). Improving energy expenditure estimates from wearable devices: A machine learning approach. *Journal of Sports Sciences, 38*(13), 1496-1505. doi:10.1080/02640414.2020.1746088

31. Santarelli, L., Diyakonova, O., Betti, S., Esposito, D., Castro, E., & Cavallo, F. (2018). Development of a novel wearable ring-shaped biosensor. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual International Conference, 2018*, 3750-3753. doi:10.1109/EMBC.2018.8513330 [doi]
32. Spierer, D. K., Rosen, Z., Litman, L. L., & Fujii, K. (2015). Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *Journal of Medical Engineering & Technology, 39*(5), 264-271. doi:10.3109/03091902.2015.1047536
33. Stahl, S. E., An, H., Dinkel, D. M., Noble, J. M., & Lee, J. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? are they accurate enough? *BMJ Open Sport & Exercise Medicine, 2*(1), e000106. doi:10.1136/bmjsem-2015-000106
34. Støve, M. P., Haucke, E., Nymann, M. L., Sigurdsson, T., & Larsen, B. T. (2018). Accuracy of the wearable activity tracker garmin forerunner 235 for the assessment of heart rate during rest and activity. *Journal of Sports Sciences, 37*(8), 895-901. doi:10.1080/02640414.2018.1535563
35. Thiebaud, R. S., Funk, M. D., Patton, J. C., Massey, B. L., Shay, T. E., Schmidt, M. G., et al. (2018). Validity of wrist-worn consumer products to measure heart rate and energy expenditure. *Digital Health, 4*, 205520761877032-2055207618770322. doi:10.1177/2055207618770322
36. Thompson, W. R. (2019). Worldwide survey of fitness trends for 2020. *ACSM's Health &*

Fitness Journal, 23(6), 10-18. doi:10.1249/fit.0000000000000526

37. Trivedi, N. S., Ghouri, A. F., Shah, N. K., Lai, E., & Barker, S. J. (1997). Effects of motion, ambient light, and hypoperfusion on pulse oximeter function. *Journal of Clinical Anesthesia, 9*(3), 179-183. doi:S0952-8180(97)00039-1 [pii]
38. Villar, R., Beltrame, T., & Hughson, R. L. (2015). Validation of the hexoskin wearable vest during lying, sitting, standing, and walking activities. *Applied Physiology, Nutrition, and Metabolism, 40*(10), 1019-1024. doi:10.1139/apnm-2015-0140
39. Wallen, M. P., Gomersall, S. R., Keating, S. E., Wisløff, U., & Coombes, J. S. (2016). Accuracy of heart rate watches: Implications for weight management. *PloS One, 11*(5), e0154420. doi:10.1371/journal.pone.0154420
40. Wang, R., Blackburn, G., Desai, M., Phelan, D., Gillinov, L., Houghtaling, P., et al. (2017). Accuracy of wrist-worn heart rate monitors. *JAMA Cardiology, 2*(1), 104-106. doi:10.1001/jamacardio.2016.3340
41. WOODMAN, J. A., CROUTER, S. E., BASSETT, D. R., FITZHUGH, E. C., & BOYER, W. R. (2017). Accuracy of consumer monitors for estimating energy expenditure and activity type. *Medicine and Science in Sports and Exercise, 49*(2), 371-377. doi:10.1249/mss.0000000000001090
42. Wu, B., Chu, Y., Huang, P., & Chung, M. (2019). Neural network based luminance variation resistant remote-photoplethysmography for driver's heart rate monitoring. *IEEE Access, 7*, 57210-57225. doi:10.1109/ACCESS.2019.2913664

Curriculum Vitae

Brenna Barrios
brennaneske@gmail.com
Las Vegas, NV

EDUCATION

M.S.	University of Nevada, Las Vegas, Las Vegas NV, expected graduation date, May 2021; Exercise Physiology (concentration)
B.S.	California State University, Fullerton, Fullerton CA, 2019; Exercise Science (concentration)

CERTIFICATIONS

USA Weightlifting Coach Level 1

American Red Cross Adult and Pediatric First Aid/CPR/AED Certification

FIELD EXPERIENCE

2020 - Current	University Nevada Las Vegas, Las Vegas NV – Part Time Instructor. Taught Physical Education Classes: Womens Weight Training and Womens Conditioning.
2019 - 2021	University Nevada Las Vegas, Las Vegas NV – Graduate Assistant. Taught anatomy and physiology lab classes. Created and taught anatomy/physiology content.
2012 – 2019	Tomacelli Academy, Costa Mesa CA – Head of Strength & Conditioning. Programmed for professional fighters, amature Fighters, and for personal fitness. Programmed pre-habilitation. Lead classes in Olympic Weightlifting, HITT, and self-defense.
2018 – 2019	Crossfit Upgrade, Costa Mesa CA – Coach. Programmed for Personal fitness and daily classes. Lead classes in Crossfit,

Pre-habilitation, and yoga.

TEACHING EXPERIENCE

University of Nevada, Las Vegas	KIN 223 - Anatomy Lab PEX 152 - Conditioning for Women PEX 152B - Weight Training for Women

RELEVANT EXPERIENCE

June, 2020	International Journal of Exercise Science – Student Reviewer
Aug. 2019 – Present	University of Nevada, Las Vegas – Graduate student researcher in Dr.Navalta exercise physiology Lab
Aug. 2019	University of Nevada, Las Vegas – Conducted physiological testing on the Las Vegas Lights amature soccer team.
Aug 2015 – May 2017	California State University, Fullerton – Conducted physiological testing on the Anaheim Ducks, National Hockey League
Aug 2012 – May 2015	Orange Coast College – Student athletic trainer (football), worked closely with the head athletic trainer for treatment plans, rehabilitation, and pre-game preparation

MANUSCRIPTS

Published Manuscripts

1. Carrier, B.; **Barrios, B**.; Jolley, B.D.; Navalta, J.W. Validity and Reliability of Physiological Data in Applied Settings Measured by Wearable Technology: A Rapid Systematic Review. Technologies 2020, 8, 70. https://doi.org/10.3390/technologies8040070

2. **Barrios, Brenna**; Carrier, Bryson; Jolley, Brayden; Davis, Dustin W.; Sertic, Jacquelyn; and Navalta, James W. (2020); Establishing a Methodology for Conducting a Rapid Review on Wearable Technology Reliability and Validity in Applied Settings," Topics in Exercise Science and Kinesiology: Vol. 1 : Iss. 2 , Article 8.

3. Salatto, R W.; Davis, Dustin W.; Carrier, Bryson; **Barrios, Brenna**; Sertic, Jacquelyn; Cater, Peyton; and Navalta, James W. (2020); Efficient Method of Delivery for Powdered Supplement or Placebo for an Outdoor Exercise Investigation," Scholarship in Kinesiology: Vol. 1 : Iss. 2 , Article 5.

Manuscripts Submitted (in-review)

1. James W. Navalta; Graham R. McGinnis; Jacob W. Manning; Robert W. Salatto; Bryson Carrier; Dustin W. Davis; **Brenna Barrios**; Mark DeBeliso, (Sep. 2020) “Acute Beta-alanine Supplementation on the Perception of Pain and Exertion Before and after Hiking”. Submitted to Amino Acids.

SCIENTIFIC PRESENTATIONS

Oral

1. Symposium: Wearable Activity Monitors. Introduction of student presenters, Navalta, J.W.; The evolution of wearable devices, Salatto, R.W.; The current state of technology devices in applied settings, **Barrios, B**.; The needed considerations in current testing models, Jolley, B.D.; The future of wearable exercise testing, Carrier, B., Virtual Annual Meeting of the Southwest American College of Sports Medicine, 2020.

2. Carrier, B., Salatto, R.W., Manning, J.W., **Barrios, B**., Sertic, J.V.L., Davis, D.W., Cater, P.C., McGinnis, G., DeBeliso, M., Navalta, J.W., “Does Acute Beta-Alanine Supplementation Improve Performance, Rating of Perceived Exertion and Heart Rate During Hiking?”, American College of Sports Medicine, May 2020, San Francisco, CA, USA. Presented as a thematic poster presentation.

Poster Presentation

1. **Barrios, B**., Sertic, J.V.L., Cater, P.C., Davis, D.W., Carrier, B., Salatto, R.W., Montes, J., Bodell, N., Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the Validity of Heart Rate Measured by the Jabra Elite During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

2. Cater, P.C., Sertic, J.V.L., Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Montes, J., Bodell, N.,Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the

Validity of Heart Rate Measured bythe Rhythm During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

3. Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Sertic, J.V.L., Cater, P.C.1, Montes, J., Bodell,N., Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the Validity of Heart Rate Measured by the Garmin Fenix 5 During Trail Running. Southwest American College of Sports Medicine,Oct. 2019, Costa Mesa, CA, USA

4. Salatto, R.W., Navalta, J.W., Montes, J., Bodell, N., Carrier, B., Sertic, J.V.L, **Barrios, B**., Cater, P.C., Davis, D.W., Manning, J.W., DeBeliso, M., Evaluating the Validity of Heart Rate Measured by the Suunto Spartan Sport Watch During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

5. Sertic, J.V. L., Cater, P.C., Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Montes, J., Bodell,N., Manning, J.W., DeBeliso, M., Navalta, J.W., Validating the Heart Rate Feature of the Motiv Ring on Outside Graded Terrain. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

6. Navalta, J.W., Salatto, R.W., Montes, J., Bodell, N., Carrier, B., Sertic, J.V.L., **Barrios, B**., Cater, P., Davis, D., Manning, J.W., DeBeliso, M., Wearable Device Price is Correlated with the Limits ofAgreement Range as a Measure of Heart Rate Validity during Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

ABSTRACTS

Peer Reviewed Abstracts

1. Kyle Cruz; RW Salatto; Dustin W. Davis; Bryson Carrier; **Brenna Barrios**; Peyton Cater; Heather Farmer; James W. Navalta, “Evaluation of Rating of Perceived Exertion During Mountain Biking”, Southwest American College of Sports Medicine, Oct. 2020, Costa Mesa, CA, USA.

2. Heather Farmer; Heather Farmer, RW Salatto, Dustin Davis, Bryson Carrier, **Brenna Barrios**, Peyton Cater, Kyle Cruz, James Navalta, FACSM, “Felt Arousal Scale is Not Reliable for Use in Repeated Mountain Biking Trial Application”, Southwest American College of Sports Medicine, Oct. 2020, Costa Mesa, CA, USA.

3. Carrier, B.; Salatto, R.W.; Manning, J.W.; **Barrios, B.**; Sertic, J.V.L.; Davis, D.W.; Cater, P.C.; McGinnis, G.; DeBeliso, M.; Navalta, J.W., “Does Acute Beta-Alanine Supplementation Improve Performance, Rating of Perceived Exertion and Heart Rate During Hiking?”, American College of Sports Medicine, May 2020, San Francisco, CA, USA.

4. **Barrios, B**.; Carrier, B.; Cater, P.C.; Sertic, J.V.L.; Salatto, R.W.; Navalta, J.W., “Validation of Heart Rate Monitoring of Fenix 5 During Mountain Biking”, American College of Sports Medicine, May 2020, San Francisco, CA, USA.

5. Sertic, J.V.L.; Carrier, B.; Cater, P. C.; **Barrios, B**.; Salatto, R. W.; Navalta, J. W., “Validation of Two Wearable Chest Straps for Heart Rate Monitoring During Mountain Biking”, American College of Sports Medicine, May 2020, San Francisco, CA, USA.

6. Salatto, R.W.; Navalta, J.W.; Montes, J.; Bodell, N.; Carrier, B.; Sertic, J.V.L; **Barrios, B**.; Cater, P.C.; Davis, D.W.; Manning, J.W.; DeBeliso, M., “Evaluating the Validity of Heart Rate Measured by the Suunto Spartan Sport Watch During Trail Running”, American College of Sports Medicine, May 2020, San Francisco, CA, USA.

7. **Barrios, B**., Sertic, J.V.L., Cater, P.C., Davis, D.W., Carrier, B., Salatto, R.W., Montes, J., Bodell, N., Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the Validity of Heart Rate Measured by the Jabra Elite During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

8. Cater, P.C., Sertic, J.V.L., Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Montes, J., Bodell, N., Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the Validity of Heart Rate Measured by the Rhythm During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

9. Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Sertic, J.V.L., Cater, P.C.1, Montes, J., Bodell, N., Manning, J.W., DeBeliso, M., Navalta, J.W., Evaluating the Validity of Heart Rate Measured by the Garmin Fenix 5 During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

10. Salatto, R.W., Navalta, J.W., Montes, J., Bodell, N., Carrier, B., Sertic, J.V.L, **Barrios, B**., Cater, P.C., Davis, D.W., Manning, J.W., DeBeliso, M., Evaluating the Validity of Heart Rate Measured by the Suunto Spartan Sport Watch During Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

11. Sertic, J.V. L., Cater, P.C., Davis, D.W., **Barrios, B**., Carrier, B., Salatto, R.W., Montes, J., Bodell, N., Manning, J.W., DeBeliso, M., Navalta, J.W., Validating the Heart Rate Feature of the Motiv Ring on Outside Graded Terrain. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

12. Navalta, J.W., Salatto, R.W., Montes, J., Bodell, N., Carrier, B., Sertic, J.V.L., **Barrios, B**., Cater, P.,Davis, D., Manning, J.W., DeBeliso, M., Wearable Device Price is Correlated with the Limits of Agreement Range as a Measure of Heart Rate Validity during Trail Running. Southwest American College of Sports Medicine, Oct. 2019, Costa Mesa, CA, USA

FUNDING

University of Nevada, Las Vegas: Graduate & Professional Student Association Research Grant, $200, 2019

VOLUNTEER

June, 2020	International Journal of Exercise Science – Student Reviewer
To be determined	Or, CA – Guest speaker. Focusing On attending college, and exercise science/personal Training as careers after incarceration