COMPARISON OF THE APPLE WATCH, FITBIT SURGE, AND ACTIGRAPHT
GT9X LINK IN MEASURING ENERGY EXPENDITURE, STEPS, DISTANCE, AND
HEART RATE

SARAH KIRK

Bachelor of Science in Exercise Science
University of Mount Union
May, 2014

Submitted in partial fulfillment of requirements for the degree

MASTER OF EDUCATION

at

CLEVELAND STATE UNIVERSITY

May, 2016
This thesis has been approved for

Sarah Elizabeth Kirk

Candidate for the Master of Education in Exercise Degree

Department of Health and Human Performance

And

Cleveland State University’s

College of Graduate Studies by

Thesis Chairperson, Dr. Kenneth Sparks

Thesis Committee Member, Dr. Kathleen Little

Thesis Committee Member, Dr. Emily Kullman

Associate Dean of Student Services, Dr. Kristine Still

Student’s Date of Defense: April 18, 2016
ACKNOWLEDGMENTS

I would like to thank a few people who have helped me throughout this journey of completing my thesis. First, I have such gratitude for my mom and dad for all of their support and guidance in everything that I do. You were there for me through every step of this long process. I love you both so much, and I don’t know how I would have made it through this many years of school or life without you two! I will never be able express how lucky I am to have two amazing role models in my life. I would also like to thank my siblings, Tommy, Katie, and Jenny. The three of you are such a strong support system, and I am forever grateful. I would also like to thank my grandpa for supporting me and always finding a way to put a smile on my face. Thank you to Dr. Sparks for helping and guiding me throughout this entire thesis and for all the hours you spent helping me collect data. It was a privilege to work with you, and I have learned so much from you over the last two years. Finally, thank you to Dr. Little and Dr. Kullman for serving on my committee and for all of your feedback.
COMPARISON OF THE APPLE WATCH, FITBIT SURGE, AND ACTIGRAPH GT9X LINK IN MEASURING ENERGY EXPENDITURE, STEPS, DISTANCE, AND HEART RATE

Sarah Kirk

ABSTRACT

Introduction: Physical activity tracking devices have become a popular tool to help individuals become active and healthy. Purpose: The purpose of this study was to compare the accuracy of the Apple Watch, Fitbit Surge, and Actigraph in estimating energy expenditure, steps, HR, and distance. Methods: An experimental study was used to compare the accuracy of the Apple Watch, Fitbit Surge, and Actigraph to indirect calorimetry (energy expenditure), hand counting (steps), ScottCare Telemetry System (HR), and calibrated treadmill (distance). Twenty subjects (10 males, 10 females), wore the Apple Watch, Fitbit Surge and Actigraph while connected to the Cosmed and ScottCare system. Subjects jogged and walked two miles (on two separate testing days) while wearing all devices simultaneously. Energy expenditure, steps, HR, and distance were compared using Repeated Measures. If there was statistical significance, a protected T-Test was used to assess differences between devices. Results: During the two mile walk, both the Fitbit Surge and Actigraph significantly differed from the Cosmed (p=0.0001); while the Apple Watch did not differ (p=0.238). For the two mile jog, the Fitbit Surge and Actigraph significantly differed from the Cosmed (p=0.0001 and 0.008, respectively); while the Apple Watch did not differ (p=0.231). For steps during the walk and jog, there was no statistical difference between the hand counter and any of the three commercial devices. Compared to ScottCare HR, the Surge, Apple, and Actigraph did not
significantly differ (p=0.916) for the walk. During the jog, the Surge and Actigraph did not significantly differ from ScottCare (p= 0.117 and 0.895 respectively), but the Apple watch did significantly differ (p=0.0001). During the walk and run, the Fitbit significantly differed from the calibrated treadmill (p=0.0001), while the Apple Watch did not significantly differ (p=0.793 and 0.047, respectively). **Conclusion:** The results of this study support previous research that fitness devices are not always as accurate, but still give individuals a “ballpark” estimate of calories burned, steps, HR, and distance. Of the devices tested, the Apple Watch was generally the most accurate.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPERS</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>1.2 PURPOSE</td>
<td>2</td>
</tr>
<tr>
<td>1.3 HYPOTHESES</td>
<td>2</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>2.1 ENERGY EXPENDITURE</td>
<td>4</td>
</tr>
<tr>
<td>2.2 STEPS</td>
<td>12</td>
</tr>
<tr>
<td>2.3 HEART RATE</td>
<td>18</td>
</tr>
<tr>
<td>2.4 DISTANCE</td>
<td>20</td>
</tr>
<tr>
<td>III. METHODS</td>
<td>22</td>
</tr>
<tr>
<td>3.1 RESEARCH DESIGN</td>
<td>22</td>
</tr>
<tr>
<td>3.2 SUBJECTS</td>
<td>22</td>
</tr>
<tr>
<td>3.3 INSTRUMENTS</td>
<td>23</td>
</tr>
<tr>
<td>3.4 PROCEDURES</td>
<td>24</td>
</tr>
<tr>
<td>3.5 DATA ANALYSIS</td>
<td>25</td>
</tr>
<tr>
<td>IV. RESULTS AND DISCUSSION</td>
<td>27</td>
</tr>
</tbody>
</table>
4.1 SUBJECT CHARACTERISTICS ......................................................... 27
4.2 ENERGY EXPENDITURE ............................................................. 28
4.3 STEPS ......................................................................................... 29
4.4 HEART RATE ............................................................................. 30
4.5 DISTANCE .................................................................................. 32
4.6 DISCUSSION ............................................................................. 33

V. SUMMARY AND CONCLUSION ..................................................... 36

5.1 SUMMARY .............................................................................. 36
5.2 CONCLUSION ............................................................................ 37
5.3 LIMITATIONS ........................................................................... 37
5.4 FUTURE RESEARCH RECOMMENDATIONS .............................. 38
5.5 APPLICATION .......................................................................... 38

REFERENCES .................................................................................. 39

APPENDICES .................................................................................. 44

A. RECRUITMENT FLYER ................................................................. 45
B. INFORMED CONSENT ................................................................. 46
C. AHA/ACSM PRESCREENING QUESTIONNAIRE ....................... 49
D. APPLE WACH ............................................................................ 50
E. FITBIT SURGE ............................................................................ 51
F. ACTIGRAPH GT9X LINK .............................................................. 52
G. COSMED K4b2 ........................................................................... 53
H. SCOTTCARE TELEMETRY SYSTEM ........................................... 54
LIST OF TABLES

1. CHARACTERISTICS OF SUBJECTS .....................................................27
2. WALKING ENERGY EXPENDITURE ....................................................28
3. JOGGING ENERGY EXPENDITURE ....................................................29
4. WALKING STEPS .................................................................29
5. JOGGING STEPS .................................................................30
6. WALKING HEART RATE ...........................................................31
7. JOGGING HEART RATE ............................................................31
8. WALKING DISTANCE ..............................................................32
9. JOGGING DISTANCE ..............................................................33
CHAPTER I
INTRODUCTION

1.1 Background Information

Obesity is a major contributing factor to many diseases such as diabetes, hypertension, and coronary artery disease (Neiman, 2011). Physical inactivity, which is a common problem in the world today, is a major contributor to the obesity epidemic (Tudor-Locke, 2000). According to the Centers for Disease Control and Prevention (CDC), adults should accumulate at minimum 150 minutes of moderate- to- vigorous physical activity a week (cdc.org, 2014). Physical activity monitors have gained attention over the past few years. Most accelerometers track energy expenditure (calories burned) and steps. Some also measure distance, stairs climbed, heart rate (HR), and sleep patterns. Much research has been conducted on the accuracy and validity of accelerometers and pedometers for estimating energy expenditure and counting steps.

Having an accurate and reliable device can provide motivation or encouragement and could help promote a decline in the obesity epidemic via an increase in physical activity. Recently, Apple has developed a new device that measures energy expenditure, steps,
HR, and distance. Fitbit has an enhanced device, the Surge, which also measures energy expenditure, steps HR, and distance. The ActiGraph Link pedometer is commonly used by the National Institute of Health (NIH) and measures energy expenditure, steps, and HR. The Apple Watch and Fitbit products have not been validated or compared to each other on measuring energy expenditure, steps, distance, and HR. Also, all three products have not been compared to each other on measuring energy expenditure, steps, and HR.

1.2 Purpose

The purpose of this study was to compare the accuracy of the Apple Watch, Fitbit Surge, and ActiGraph in estimating energy expenditure, steps, distance, and HR while walking and jogging on a treadmill.

1.3 Hypotheses

Hypothesis 1) It was hypothesized that the Apple Watch will be accurate in measuring energy expenditure, steps, HR, and distance compared to the Cosmed (energy expenditure), hand counting (steps), ScottCare Telemetry System (HR), and distance (calibrated treadmill).

Hypothesis 2) It was hypothesized that the Fitbit Surge will be accurate in measuring energy expenditure, steps, HR, and distance compared to the Cosmed (energy expenditure), hand counting (steps), ScottCare Telemetry System (HR), and distance (calibrated treadmill).

Hypothesis 3) It was hypothesized that the ActiGraph will be accurate in measuring energy expenditure, steps, HR, and distance compared to the Cosmed (energy
expenditure), hand counting (steps), ScottCare Telemetry System (HR), and distance (calibrated treadmill).

Hypothesis 4) There will be no difference between the Apple Watch and Fitbit Surge for measuring energy expenditure, steps, HR, and distance.

Hypothesis 5) There will be no difference between the Apple Watch, Fitbit Surge, and the ActiGraph for measuring energy expenditure, steps, and HR.
CHAPTER II
LITERATURE REVIEW

2.1 Energy Expenditure

One feature of most physical activity monitors is an estimation of calories burned during exercise or throughout the day. To determine the accuracy of accelerometers, they need to be compared to the “gold standard” of energy expenditure, indirect calorimetry. Dannecker, Sazonova, Lee, Kim, & Welk (2013) examined energy expenditure of several different physical activity trackers in 19 young adults (10 men, 9 women), ages 20-36 years. Subjects wore several different fitness accelerometers; one was a pair of walking shoes that had an accelerometer and an insole that recognized applied pressure. Data from these shoes was uploaded through a smartphone via Bluetooth. In addition to the shoe devices, the subjects also wore five other monitors, three of which are used strictly for research purposes (ActiGraph, Actical, and IDEEA) and two directed towards consumers (Fitbit and DirectLife). Subjects performed different postures and activities including
equilibrium, supine sitting, standing, walking (3.5 mph at 0% grade), walking at 2.5% at 2.5 mph, stepping, sweeping, cycling at 75W, standing, sitting, and free living activities (any of the above at their own pace). Energy expenditure of the devices was compared to indirect calorimetry for three and a half hours. Dannecker et al. (2013) found that the estimated energy expenditure via the shoe device, IDEEA and DirectLife was not significantly different than the measured value. However, calories were underestimated from both the Fitbit (499.0 ± 23.8 vs 362± 19.8) and Actigraph (494.2± 20.0 vs. 375.0 ±20.6) devices. Dannecker et al. (2013) concluded that estimation of energy expenditure is more accurate based on total physical activity.

Noah, Sprierer, Gu, & Bronner (2013) examined energy expenditure of three different fitness accelerometers compared to indirect calorimetry in 23 individuals (13 males, 10 females), ages 18-35 years. The Cosmed K4b² was used to measure energy expenditure while two Fitbits (Tracker and Ultra) and two Actical accelerometers were attached to a waistband to estimate energy expenditure. Subjects were first measured at rest for six minutes. Subjects then walked for six minutes each at 3.5 mph, 0% grade; 3.5 mph, 5% grade; and jogged at 5.5mph, 0% grade. Between each six-minute session, there was a minute of rest. After the walking and jogging portion of the test, subjects then recovered for six minutes. Following rest, subjects completed six minutes on the stair stepper at 96 steps per minute. Noah et al. (2013) found that the Fitbit Tracker accuracy was 90% and 88% accurate for walking and jogging, respectively; the Fitbit Ultra for walking and jogging was 91 and 113% of the Cosmed, respectively; the Actical for walking and jogging was 81% and 91% accurate, respectively. Compared to indirect calorimetry, the Fitbit Tracker, Fitbit Ultra and Actical devices underestimated energy
expenditure during the inclined walk (11.1 ± 3.0 vs. 6.5 ± 1.4, 6.7 ± 2.1, and 5.9 ± 1.3, respectively) and stepping (8.2 ± 2.1 vs. 4.9 ± 1.3, 4.6 ± 1.2, and 4.3 ± 1.2, respectively).

Crouter, Schneider, Karabulut, & Basset (2003) examined the validity of 10 electronic pedometers to measure energy expenditure in 10 subjects (5 males and 5 females), ages 21-45 years. Subjects wore 10 different pedometers on each side of the body as they walked on a treadmill at five different speeds (54, 67, 80, 94, and 107 m/min), each in five minute bouts. Energy expenditure was measured using indirect calorimetry simultaneously during each stage. After each stage, subjects rested for one minute so that the researchers could retrieve data from the devices. Crouter et al. (2003) stated it was not clear to the researchers if the pedometers were estimating net or gross energy expenditure. If they were calculating total energy expenditure, seven out of eight pedometers were within ±30% of total calories burned; however if calculating net calories, the pedometers overestimated. Crouter et al. (2003) concluded that pedometers are not very accurate when estimating energy expenditure.

Lee, Kim, & Welk (2014) examined eight different physical activity monitors for measuring energy expenditure compared to indirect calorimetry in 60 subjects (30 males, 30 females), ages 18-43 years. Subjects wore all eight of the physical activity accelerometers simultaneously while connected to a portable metabolic cart. Each testing session lasted 69 minutes. All activities were 5 minutes long, except for the treadmill activities which were 3 minutes. Activities included: sedentary (sitting in a chair or at a computer), walking on a treadmill (2.5 mph, 3.5 mph, incline walking at own pace with and without a 15 pound weighted backpack), running on a treadmill (5.5 mph and 6.5 mph), and moderate-to-vigorous activities (walking up and down stairs, elliptical,
bike, basketball, Wii tennis). Lee et al. (2014) found that the majority of the physical activity monitors were within 10% of the measured energy expenditure from indirect calorimetry and concluded that most, if not all, of these activity monitors were valid in estimating energy expenditure.

Sasaki et al. (2015) examined the Fitbit Classic wireless activity tracker in estimating energy expenditure in 20 subjects (10 males, 10 females), ages 20-28 years. Participants wore a Fitbit Classic wireless activity tracker in an elastic belt around the waist positioned in line with the anterior axillary line, and an Oxycon Mobile, a portable indirect calorimetry device. Participants attended two testing sessions; during the first testing session they completed the following protocol: walking at 3.0 mph at 5% grade, walking at 4.0 mph at 5% grade, and running at 5.5 mph at 0% grade. During the second session, the subjects were assigned to one of the following activity groups: 1. Office work, driving, carrying a box, carrying groceries, and walking up and down stairs; 2. cycling at 300kgm/min, golf, tennis, and basketball; or 3. dusting, vacuuming, laundry, raking, and gardening. For both sessions, each activity was performed for six-minutes, except for driving for 11 minutes. In between each activity bout was a four-minute rest period. During the rest period, the Fitbit was restarted and the indirect calorimeter was monitored to make sure data returned to baseline values. Sasaki et al. (2015) found that the Fitbit Classic underestimated energy expenditure for 6/15 of the activities including: cycling, laundry, raking, walking on a treadmill at 3.0 mph at 5% grade, walking up/down stairs, and walking on a treadmill at 4.0mph at 5% grade. For the other nine activities: basketball, carrying groceries, dusting, gardening, golf, office work, tennis, and walking on a treadmill at 5.5mph at 0% grade, and vacuuming, there was no difference
between the Fitbit and the Oxycon Mobile. It was concluded that the Fitbit underestimated energy expenditure compared to indirect calorimetry, which could be detrimental for individuals using this device as a weight loss tool.

Groot and Nieuwenhuizen (2013) assessed the validity and reliability of measuring activities while wearing the DynaPort MoveMonitor. This study was conducted in two sessions; one session tested validity and the other tested reliability, in 28 subjects (12 men, 16 women), ages 19-75 years. Subjects performed a variety of activities including laying, sitting, walking at different speeds (2, 3, 5 mph), standing, walking up and down stairs, and biking at different workloads (25, 50, 75W). Subjects performed all of these activities within a 45 minute timeframe, including time to transfer from task to task (transfer time was not included in the data). During the validity trial, subjects were attached to the Cosmed K4b2 to measure energy expenditure. For the reliability session, subjects had to complete 2, 26 minute trials on the same day performing a variety of different activities. These activities included laying (on back, belly, side, with an incline of 20, 40, and 60 degrees), walking (with self-preferred speed, -2 km/h below and +2 km/h above self-preferred speed), biking (50, 75, 100 rpm), standing, and sitting in a chair. Subjects completed both trials while wearing a DynaPort MoveMonitor on their lower back and minute by minute analysis was computed. Groot and Nieuwenhuizen (2013) found that energy expenditure for walking was on average 104 W higher and 170 W lower for cycling, for the MoveMonitor compared to indirect calorimetry. The MoveMonitor software plots showed that energy expenditure was overestimated during walking the intra-class correlation coefficients (intra-class correlation coefficients (ICC) = 0.54), but underestimated during cycling (ICC= 0.03).
Groot and Nieuwenhuizen (2013) concluded that the MoveMonitor is not valid in estimating energy expenditure and has low reliability.

Brazeau et al. (2014) also examined the reliability and validity of two activity monitors for estimating energy expenditure in 20 participants ages 18-45 years old. Participants wore two monitors, the SenseWear Armband (SWA) and Actical (ACT) device, to measure the accuracy and validity of energy expenditure during rest and light-to-moderate exercise (walking, biking). Participants performed a graded exercise test on an ergocycle until voluntary exhaustion to determine their VO2peak. Every three minutes of this session, power output was increased by 25W. An Ergocard (indirect calorimetry) was used for breathe- by- breathe gas analysis. For the out-clinic (free-living) session, subjects wore the two monitors all day for seven days; SWA was worn on the arm and ACT on the hip. Brazeau et al. (2014) compared the estimated energy expenditure of the monitors to measured energy expenditure through the doubly labeled water (DLW) method. Following the seven-day free living sessions, subjects returned to the lab for 3, 10 hour/day in laboratory trials. During these 10 hour days, subjects entered the lab and laid down for 60 minutes then they participated in two, 45-minute exercise sessions (treadmill intensities: 22-41% of their VO2peak and ~50% of their VO2peak for the bike). The activities were repeated except the speeds on the treadmill increased each day to examine energy expended during different speeds. During the exercise sessions, energy expenditure was measured using the Ergocard (indirect calorimetry), which was used to obtain the VO2peak. Brazeau et al. (2014) found that compared to indirect calorimetry, SWA overestimated calories by ~94 calories, while the ACT underestimated by -244 calories. During the in-clinic session, the results showed that energy expenditure
was overestimated during rest (SWA 210 ± 116, ACT 124 ± 133 kcal/d) and on the treadmill (SWA 54 ± 46 to 67 ± 38, ACT -269 ± 111 kcal/d), but was underestimated on the bike (SWA -93 ± 65, ACT -269 ± 111 kcal/d). Brazeau et al. (2014) concluded that these devices were reliable, but not accurate, when estimating energy expenditure.

Ellis et al. (2014), examined the prediction of energy expenditure during different types of activity using accelerometers in 40 adults (19 men, 21 women) ages 35.8 ± 21.3 years. Subjects were connected to a Cosmed K4b2 to measure VO2 and wore three ActiGraph GT3X+ accelerometers, one on each hip and one on their non-dominant wrist while performing activities including stair climbing, brisk and slow walking, and jogging, laundry, window washing, dusting, dishes, and sweeping. Four different routines were created which consisted of three household and three locomotion activities. Subjects were randomly assigned to four different routines groups and performed each activity in that routine for six minute bouts. Ellis et al. (2014) found that the accelerometer on the waist was the most accurate when estimating energy expenditure (92.3%), while the wrist accelerometer was 87.5% accurate. Among all of the activities, the wrist accelerometer was more accurate than the hip (80.2% and 70.2%, respectively). Ellis et al. (2014) concluded that when estimating energy expenditure, the wrist and hip placement produced comparable results.

King, Torres, Potter, Brooks, and Coleman (2004) compared activity monitors in estimating energy cost during treadmill exercise in 21 subjects (10 men, 11 women) ages 21 to 29 years. Subjects performed 10 minute bouts of exercise on a treadmill walking at 54, 80, and 107 m/min, and running at 134,161,188, and 214 m/min. While exercising on the treadmill, subjects wore the CSA, TriTrac-R3D, RT3, and BioTrainer-Pro
accelerometers on their waists, and the SenseWear Armband Monitor on their arm, while connected to an indirect calorimetry system. King et al. (2004) found that the estimated energy expenditure for the SenseWear Armband, CSA, and TriTrac-R3D all were significantly higher during both the running and walking tests. All of the accelerometers overestimated energy expenditure for both running and walking compared to indirect calorimetry. King et al. (2004) concluded that the CSA estimation was similar to indirect calorimetry during walking and jogging; the SenseWear Armband was similar during all speeds, and the TriTrac-R3D was most similar during the running.

Giannakidou et al. (2011) examined the accuracy of two Omron pedometers on energy expenditure in 24 males and 18 females (ages 19 to 25 years). Subjects walked on a treadmill at five different speeds (54, 67, 80, 94, and 107 m min\(^{-1}\)) in a randomized order, wearing a HJ-113 on the left hip and a YAM pedometer on their right hip, 2 cm away from a HJ-720 pedometer for five minute stages. After each five-minute stage, subjects stepped onto the sides of the treadmill and all data was recorded. Giannakidou et al. (2011) found that both Omron pedometers underestimated calories burned during all walking stages by 28% compared to indirect calorimetry. Giannakidou et al. (2011) concluded that these two devices were not accurate in estimation of energy expenditure.

Haymes & Byrnes (1993) estimated energy expenditure using a Caltrac accelerometer during running and walking in 20 subjects (10 men, 10 women). Subjects wore Caltrac pedometers on each side of their waist. One device was programmed with the subject’s height, weight, age, and sex and the other device programmed as age= 99, height=36, weight= 25 and sex=0. This design calibrates the computer to a metabolic rate of zero. Each subject walked (2, 3, 4, & 5 mph) and ran on a treadmill (4, 5, 6, 7, &
8mph) for four minute increments with one minute rest in between each stage. There was a 10-minute rest stage between the walking and running, and the Caltracs were reset and values were recorded before the next stage began. Indirect calorimetry was measured using an Applied Electrochemistry S-3A oxygen analyzer and Beckman LB-1 carbon dioxide analyzer. Haymes & Byrnes (1993) found that Caltrac overestimated energy expenditure by 1.5 kcal/min. Caltrac also overestimated energy expenditure during all five stages of running (4, 5, 6, 7, & 8mph) by 3.6, 3.7, 2.6, and 1.1 kcal/min, respectively. There was a significant correlation between the Caltrac and actual energy expenditure during both the walking and running (r= .91 and .71, respectively). Haymes & Byrnes (1993) concluded that the Caltrac is a valid tracker for walking, however it does not differentiate energy expended between speeds of 5 and 8mph.

2.2 Steps

According to the Heart Foundation of Australia (2009), 10,000 steps is the recommended number of steps for a healthy individual to try to achieve daily. Accelerometers, such as pedometers and fitness watches, have a feature that keeps a tally of total steps taken throughout the day. Many research studies have examined the accuracy of these accelerometers’ step counting feature. Noah et al. (2013) who examined energy expenditure of three different fitness accelerometers compared to indirect calorimetry in 23 individuals (13 males, 10 females) ages 18-35 years, also examined estimation of steps using the Fitbit Tracker, Fitbit Ultra, and Actical. Noah et al. (2013) found no difference in the number of steps between the Fitbit and Fitbit Ultra, as well as a high relationship (r=0.80-0.99) between all of the devices and concluded that these devices are reliable and valid when it comes to the accuracy of step counting.
Schneider, Crouter, Lukajic, & Basset (2003) examined the accuracy and reliability of measuring steps over 400-m using 10 pedometers in 20 subjects (10 males, 10 females) ages 21-45 years. Each subject wore two of the same pedometers on each hip (one on the left and one on the right) and walked 400-m on a track at their normal walking pace. To count actual steps, a researcher walked behind the subject and recorded their steps. Testing was completed over a span of 1-4 days. Schneider et al. (2003) found that a couple of the pedometers overestimated steps, while a couple of them underestimated steps. The Kenz Lifecoder, New Lifestyles NL-2000, and Yamax Digi-Walker SW-01 were more accurate in their steps, all within ±3% steps of the actual counter. The Sportline 300 and the Omron HJ-105 were not as accurate, varying ± 37% of the actual count. However, the majority of the pedometers were fairly accurate in measuring steps.

Crouter et al. (2003) who examined the validity of 10 electronic pedometers to measure energy expenditure in 10 subjects (5 males and 5 females) ages 21-45 years, also tested the validity of measuring steps. While subjects walked on a treadmill at five different speed stages (54, 67, 80, 94, and 107m/min), a researcher stood next to them counting their steps using a step counter. In between each stage, subjects rested for one minute, while the estimated steps were obtained from the pedometer. Crouter et al. (2003) found that at the beginning stages of the test, when the speed was slower, the pedometers were not as accurate as when the speed increased, and concluded that when it comes to measuring steps, pedometers are better at faster speeds.

Melanson et al. (2007) tested if speed made a difference on accurate step counting using pedometers in a two part study. Part one included 259 subjects (108 males, 151
females) ages 19-85 years, while part two included 32 subjects (16 males, 16 females) ages 19-51 years. During part one of the study, subjects walked on a treadmill at self-selected speeds (one slow and one faster) for 10 minutes each while wearing a pedometer. Since error could have occurred if the researchers counted steps for 10 minutes, the total number of steps were obtained by recording the length of time it took to walk 10 steps and then converting it to steps per minute. Steps per minute were averaged for each of the minute observations (10 observations) and then multiplied by 10, which gave total number of steps. During part two, subjects walked for 12 minutes at 1.0, 1.8, and 2.6 mph while wearing three different pedometers. Subjects were then instructed to walk 72 minutes at a comfortable walking pace. Researchers counted the number of steps between minutes two and 12 for each stage and also during the 72 minutes at the comfortable walking pace, to confirm the validity of the step counter. Similarly to Crouter et al. (2003), Melason et al. (2004) found that as the speed increased, the accuracy of step counting for the pedometers increased and concluded that pedometers may not be as accurate as one ages, due to decreased speed of walking (71% accuracy for below 2mph, 74-91% between 2 and 3mph, and 96% above 3 mph).

Mammen et al. (2012) examined the quality of the FitBit step counter compared to the Yamax SW-200 pedometer in 10 subject (5 males, 5 females) ages 22-24 years. Subjects wore four Fitbit and four Yamax pedometers while completing three different activities: step test, motor vehicle test, and treadmill test. During the step test, subjects wore one Fitbit on one side of the hip and one Yamax on the other side of the hip, and walked 20 steps. Six trials for each device were completed and the data was recorded at the end of each trial. For the motor vehicle test, subjects wore the pedometers to see if
any steps were tracked while moving in a vehicle, but not physically walking. During the treadmill test, the Yamax pedometer was placed on the waist, while the Fitbit was placed in various locations (one on the waist, one in the pocket, and one on the shirt collar (males) and bra (females). For the walking trial, subjects walked on the treadmill at four different speeds (2, 3, 4, 5 and 6 km/h) and at 8, 9, 10, and 11 km/h for the running trials. Subjects ran at each speed for one minute, and a researcher counted the steps for the entire minute. Video was also set up for later confirmation of the steps. Mammen et al. (2012) found that for the step test, error was $\pm 5\%$ for both devices; for the motor vehicle test, the Fitbit recorded zero steps, while the Yamax recorded three. For the treadmill test, there was a statistically significant difference between the Yamax during the 2 and 3 km/h walk, but the Yamax and FitBit were accurate during the 4.5 and 6 km/h walk. For running, there was no significant difference between observed and recorded steps with the exception of the Fitbit device in the pocket. Mammen et al. (2012) concluded that the Fitbit is a reliable and accurate device for counting steps.

Tudor-Locke, Barrieira, & Schuna (2015) examined step counting comparing waist and wrist placement on 15 subjects (10 females, 5 males) ages 25-30 years. Subjects were instructed to walk at speeds ranging from 14 m/min to 188m/min while wearing one Actigraph accelerometer on their wrist and one on their waist. These 14 speed increments were increased after 5 minute bouts of walking. Two minute bouts of rest were given in between each five minute bout of walking or running. Steps were counted during the walk visually by a research and also were confirmed by video. Visually counted steps were obtained by taking total number of steps per bout and dividing by 5, to reflect steps per minute. Following the laboratory session, subjects were
instructed to wear the two accelerometers continuously for 7 days straight. Data was only counted if 1500 steps were achieved in a day. Tudor-Locke et al. (2014) found that for the laboratory session, the waist accelerometer was more accurate than the wrist placement. In regards to the free-living session, wrist placement overestimated steps, on average (2500 to 8700 steps more than waist placement).

Esliger et al. (2007) examined the validity of the Actical accelerometer step count function in 38 subjects, ages 9-59 years. Subjects wore four Actical and four Actigraph devices placed in a pouch that was placed on the subject’s waist. After warm-up, subjects completed 6 minute sessions of: slow walk (50 m/min), normal walk (83m/min), and run (133m/min) at 0% grade. The steps from the devices were compared to a trained hand counter, who counted every step on the treadmill representing the criterion. In addition, eight Actical devices were tested using a mechanical shaker tester. Esliger et al. (2007) found that for the exercise testing compared to the counter, during the slow walk (50m/min) both the Actical and Actigraph underestimated steps (95 vs. 88 and 90, respectively). However there was no statistical difference for the walk (88m/min) for both the Actical and Actigraph (p=.241 and .193, respectively). The run (133m/min) produced the same results for both Actical and Actigraph (p=.970 and .960, respectively). Esliger et al. (2007) concluded that the Actical and Actigraph are practical for the step count feature during faster walking and running paces, but not during a slow walking pace.

Kozy, Staudenmayer, Troiano, and Freedson (2010) compared two accelerometers during self-paced locomotion in 116 participants (49 males, 67 female), ages 18-79 years. Participants wore two Actigraph accelerometers (AM1 and AM2) and
walked at three self-determined speeds (slow, medium, and fast) in a circular hallway (1 lap; 0.47 km = 1 trial). A stopwatch was used to determine the speed of each trials. To gather step count data, repeated measures mix mode was used every 60 seconds to determine steps per minute. Kozy et al. (2010) found that the step count for AM1 was significantly higher than AM2 for all three of the speeds (bias = -19.8%; CI = -23.2, -16.4) and concluded that AM1 and AM2 were not similar to each other during slower speed trials, but were similar to each other during medium and fast speed trials.

Takas et al. (2014) validated steps using the Fitbit One activity monitor during treadmill walking in 30 adults (15 males, 15 females), ages 20-35 years. Subjects walked for five minute stages on a treadmill at five different speeds (.90, 1.12, 1.33, 1.54, and 1.78 m/s), while wearing a Fitbit One on each hip and one in the pocket of the dominant leg. Steps recorded off of the Fitbit devices were compared to two hand counters and an eight-camera motion capture system. Step counting was reviewed if there was a discrepancy between the two step counters. Takas et al. (2014) found that the there was no statistical difference between the Fitbit One and the hand counters, concordance correlation coefficients ranged from .97 to 1.00. It was concluded that the Fitbit One is an accurate and valid device for estimation of steps at different speeds.

Giannakidou et al. (2011) who examined the accuracy of two Omron pedometers on energy expenditure in 24 males and 18 females (ages 19 to 25 years), also examined accuracy of steps. Two hand counters were used to compare the pedometers. Giannakidou et al. (2011) found that the YAM pedometer underestimated steps only at 54 m min-1 (r=0.46), while the two Omron pedometers were accurate at all speeds (±1.5%
of hand counted steps) and concluded that the two Omron pedometers were accurate in estimating number of steps.

Masurier and Tudor-Locke (2003) compared the accuracy of two pedometers in two different controlled studies in 13 males and 7 females (20-55 years). Subjects wore the Yamax SW-200 pedometer and the dual-mode CSA (accelerometers). In study one, subjects walked on a treadmill at five different speeds (54, 67, 80, 94, 107 m min⁻¹) for five minute increments. After each five minute stage, the subject would stand still on the treadmill and the data was recorded from the pedometers. Masurier and Tudor-Locke (2003) found that only the Yamax pedometer underestimated steps, compared to the observer counting and the video recording, at the slowest speed than the accelerometer (75.4% vs. 98.9%), however, for the other four speeds, there was no statistical difference in number of steps between the pedometer and the accelerometer. It was concluded that both the pedometer and accelerometer are both beneficial for estimating steps over a long period of time.

2.3 Heart Rate

Another common feature of fitness activity trackers is HR. Some monitors require a chest strap which records HR via Bluetooth, while others track HR from the wrist. There is not much research on the accuracy of the HR feature on devices compared to a gold standard (i.e. ECG). Since neither a pedometer nor a HR monitor are gold standards for HR, researchers either focused on using HR to estimate energy expenditure or comparing HR monitors to other devices.
Dystad & Hausken (2014) compared accelerometer and HR monitoring during interval running, interval spinning and Zumba in 26 participants (15 women, 11 males) ages 19-24 years. Heart rate was recorded every five seconds using Polar team 2 HR belts and RS 100 monitors. Subjects ran 4 minute interval x 4 times (4x4), which included 12 minute warm-up test, followed by the 4x4 minute runs at 90-95% of HRmax. This sequence occurred four times, for a total time of 28 minutes, followed by a cool down period of 5 minutes. Subjects also completed a 4x4 spinning bout for 45 minutes. The interval spinning was the same protocol as the running and 50 minutes of Zumba, with four small breaks for water. Dystad & Hausken (2014) found that between vector magnitude/min and %HRmax, correlations ranged from .69 to -.42 (Zumba and running, respectively). It was also found that the correlation between accelerometer and mean %HR was 0.69, 0.14, and -0.42 (Zumba, 4x4 spinning, 4x4 running, respectively).

Goodie, Larkin, & Schauss (2000) validated the Polar HR monitor during physical and mental stress in 30 students (53% female, 47% male) ages 18 to 43 years. Subjects wore a HR chest strap and were connected to a three electrode electrocardiogram (ECG) and asked to complete two series of tests; a handgrip test (strength) and mental arithmetic. Subjects rested for four minutes before the strength and mental test while HR was recorded throughout the four minutes. Before the handgrip test, max handgrip was obtained and 30% was used for the test. Subjects held 30% max for three minutes. For the mental arithmetic test, subjects counted backwards by 17 from a four digit number. At the beginning of each minute of the three minute test, a different four digit number test was given. For both the handgrip and mental arithmetic test, HR from the Polar and ECG were recorded throughout the full three minutes. Goodie et al.
(2014) found that the Polar compared to the ECG $r$ ranged from .75 to 1.00, and only three subjects did not exceed $r=.90$. It was concluded that Polar HR monitors are a valid measurement of HR.

Johnstone et al. (2012) examined the reliability and validity of the Bioharness™ monitoring system device in measuring HR in 20 male subjects (ages 18-25 years). The Bioharness™ device is a data logger and telemetry system that obtains five different variables simultaneously and transmits the data through Bluetooth. To validate Bioharness™ devices, 10 of the 20 subjects completed a walk-jog-run shuttle protocol while wearing the Polar T31, which was considered to be the precision measurement device. Subjects wore two Bioharness™ devices, one on the normal chest position and the other directly above the first device. Subjects walked at 4km/h for three minutes, followed by 6km/h for three minutes. HR was obtained during the last minute of each active stage, and one minute of rest was given to the subjects before starting the jog-run stage. During the jog-run stage, subjects performed 20 seconds of 20m shuttles for 20 minutes, followed by intervals starting at 8km/h increasing .5km/h about every minute. Johnstone et al. (2012) found that the Bioharness™ was reliable for estimating HR ($r=0.91$), but suggested caution when using Bioharness™ to record HR at higher speeds.

2.4 Distance

Similarly to HR, there has not been much research on the validity accelerometers to estimate distance. Takas et al. (2014), who examined the validity of steps using the Fitbit One activity monitor during treadmill walking in 30 adults, ages 20-35 years, also examined the validation of distance. Subjects walked five minutes stages on a treadmill at
five different speeds (.90, 1.12, 1.33, 1.54, and 1.78 m/s) while wearing a Fitbit One on each hip and one in the pocket of the dominant leg. Distance recorded on the devices was compared to a calibrated treadmill. The treadmill was restarted before each stage started. Takas et al. (2014) found that the estimated distance was significantly different than the calibrated treadmill at all speeds (5-39.6% error). It was concluded that the Fitbit One is not an accurate device for estimating distance.

Giannakidou et al. (2011) who also examined the accuracy of two Omron pedometers on energy expenditure and steps, also examined the accuracy of distance. Giannakidou et al. (2011) found that in regards to distance, the HJ-720 and HJ-113 overestimated distance when the subject was at 54 and 67 m/min and underestimated at 94 and 107 m/min. It was also found that the two pedometers were most accurate at 80 m/min and the YAM was valid at all speeds except for the 54 m/min (SEM= 12.66). Giannakidou et al. (2012) concluded that the two Omron devices did not accurately estimate distance.
CHAPTER III

METHODS

3.1 Research Design

This study used an experimental research design to determine the accuracy of the Apple Watch, Fitbit Surge, and Actigraph. The dependent variables were distance, steps, HR, and energy expenditure. Each subject wore each device simultaneously.

3.2 Subjects

A convenience sample of 20 subjects (10 males, 10 females) ages 21-27 years, was recruited via flyers (Appendix B) and word of mouth from Cleveland State University and the Cleveland community. Subjects were all recreationally active, exercised at least 3-5 times per week, and could run at least two miles continuously at 6 mph. Prior to beginning the study, subjects signed an informed consent (Appendix C) approved by the Cleveland State University Institutional Review Board (Appendix A). Each participant also filled out an AHA/ACSM pre-screening questionnaire (Appendix D). Only subjects who classified as low risk participated in this study. Subjects came in for two different testing sessions, separated by at least 48 hours.
3.3 Instruments

**Apple Watch (Apple.com) (Appendix D)**

The Apple watch measures distance, steps, calories burned, and HR. Height, weight, gender, and birth date of the subject was entered into the researcher’s iPhone Health app prior to the beginning of each testing session. The type of exercise was also entered into the device (i.e. indoor walking or indoor running). Subjects wore the Apple Watch on their dominant wrist.

**Fitbit Surge (Fitbit.com) (Appendix E)**

The Fitbit Surge measures distance, steps, calories burned, and HR. Height, weight, gender, and date of birth of the subject was entered into the Fitbit app on the researcher’s smartphone prior to the beginning of each testing session. Subjects wore this device on their nondominant wrist, which was suggested by Fitbit.

**Actigraph (actigraphcorp.com) (Appendix F)**

The Actigraph measures steps, calories burned, and HR. Height, weight, date of birth, ID number, gender, and placement of Actigraph device (i.e. Hip) were all entered into the Actigraph computer software prior to each testing session. The device was attached to the subject’s right hip. Subjects also wore a Polar HR monitor around their chest which obtained HR data on the Actigraph device. Once each testing session was completed, the device was removed from the subject and synced with the software.
**Indirect Calorimetry (Cosmed K4b2) (Appendix G)**

The Cosmed K4b2 was used to measure energy expenditure from the amount of oxygen (O₂) utilized and carbon dioxide (CO₂) produced. Measurement of expired gases was obtained in minute intervals. Energy expenditure was calculated using the caloric equivalents for oxygen.

**ECG Telemetry (Appendix H)**

Heart rate was measured using the ScottCare VersaCare ECG telemetry system. The subject was prepped using alcohol wipes and Nuprep. Electrodes were placed on the subject’s right arm (RA), right leg (RL), and left leg (LL) to monitor a lead II ECG. ECG monitoring measured each heart beat and displayed HR throughout the entire session. HR was determined by counting each beat during the exercise session and obtaining the average.

**3.4 Procedures**

This study was completed at Cleveland State University in the Human Performance Laboratory on a treadmill. The devices used for this study included the Apple Watch and Fitbit Surge which measure distance, steps, HR, and energy expenditure, and the Actigraph which measures steps, HR, and energy expenditure.

Subjects entered the lab for the first testing session and filled out the informed consent and the pre-screening questionnaire. Date of birth, height, weight, and gender were recorded and entered into each of the devices. For the Fitbit Surge and Apple Watch, the information was entered into the devices app on the researcher’s
smartphone. For the Actigraph, the information was entered into the Actigraph Software on a desktop computer.

At the start of each testing session, after the calibration was completed, the researcher attached the Apple Watch, Fitbit Surge, Actigraph, Polar HR monitor, Cosmed K4b2, and ScottCare Telemetry system to the subject. All devices were started at the same time and were worn in the same location; Apple and Fitbit on each wrist and the Actigraph on the right hip, for both testing sessions. During testing session one, the subjects jogged on the treadmill at 6.0mph at 0% grade for two miles. Once two miles were complete, all devices were stopped and all data was recorded. The number of steps were obtained by the researcher using a clicker counting every time the subject’s right foot contacted the treadmill. This number was then doubled to obtained total number of steps. Subjects waited at least 48-hours before returning back to the laboratory for day two of testing. For test session two, the same calibration and set up of devices was completed. Subjects walked two miles at 3.5mph at 0% grade. Once two miles was completed, all devices were stopped and data was recorded. Similar to session one, total number of steps were counted by the researcher. Test sessions one and two were performed in this specific order to ensure that the subjects could complete the two mile jog.

3.5 Data Analysis

Descriptive statistics were obtained. A repeated measures ANOVA was used to assess differences due to the independent variable (device type) on the dependent variables (HR, distance, energy expenditure and steps) for walking and jogging. SPSS (version 22.0) was used for all analyses with .05 used as the level of significance. If
significant difference were found, protected t-test were used with 0.01 used as the level of significance for HR, steps, and energy expenditure (i.e. four devices) and 0.02 used as the level of significance for distance (i.e. three devices).
CHAPTER IV

RESULTS & DISCUSSION

4.1 Subject Characteristics

There were a total of 20 subjects (10 males, 10 females). Subjects’ age, height, and weight were obtained and entered into the Fitbit Surge, Apple Watch and Actigraph devices before each testing session (Table 1).

Table 1. Characteristics of Subjects.

<table>
<thead>
<tr>
<th></th>
<th>Total (N=20)</th>
<th>Males (N=10)</th>
<th>Females (N=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>24±3.4</td>
<td>24±3.9</td>
<td>23±2.8</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>172.7±10.2</td>
<td>180.3±4.8</td>
<td>163±4.3</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>69.5±13.4</td>
<td>80.9±7.6</td>
<td>58.2±6.1</td>
</tr>
</tbody>
</table>
4.2 Energy Expenditure

The Apple Watch, Fitbit Surge and Actigraph all estimated energy expenditure (kcal) and were compared to indirect calorimetry (Cosmed K4b2) to determine their accuracy. During the two mile walk, both the Fitbit Surge and Actigraph significantly overestimated kcals; while the Apple Watch did not significantly differ from the Cosmed (Table 2.) All three of the commercial devices significantly differed from each other (p<0.01).

Table 2. Walking Energy Expenditure (kcal).

<table>
<thead>
<tr>
<th></th>
<th>Mean+SD</th>
<th>Sig. (Two Tailed) Compared to Cosmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosmed</td>
<td>161.2±34.5</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surgeab</td>
<td>266.5±75.3</td>
<td>.000(1)</td>
</tr>
<tr>
<td>Apple Watchb</td>
<td>149.7±44.8</td>
<td>.238</td>
</tr>
<tr>
<td>Actigraphab</td>
<td>217.1±62.1</td>
<td>.000(1)</td>
</tr>
</tbody>
</table>

a. Significant difference from Cosmed (p<0.01)
b. Significant difference among the three commercial devices (p<.01).

During the two mile jog, both the Fitbit Surge and Actigraph significantly overestimated kcals; while the Apple Watch did not significantly differ from the Cosmed (Table 3). All three of the commercial devices significantly differed from each other (p<0.01).
Table 3. Jogging Energy Expenditure (kcal).

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD</th>
<th>Sig. (Two Tailed) Compared to Cosmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosmed</td>
<td>171.1±37.8</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge(^a)(^b)</td>
<td>215.7±39.6</td>
<td>.000(1)</td>
</tr>
<tr>
<td>Apple Watch(^b)</td>
<td>181.4±40.3</td>
<td>.231</td>
</tr>
<tr>
<td>Actigraph(^a)(^b)</td>
<td>203.8±58.6</td>
<td>.008</td>
</tr>
</tbody>
</table>

a. Significant difference from Cosmed (p<0.01).

b. Significant difference among the three commercial devices (p<0.01).

4.3 Steps

The Apple Watch, Fitbit Surge and Actigraph estimated total steps and were compared to a hand counter to determine their accuracy. During the two mile walk, the Fitbit Surge, Apple, and Actigraph did not significantly differ from the Counter (Table 4). The Fitbit and Actigraph did not significantly differ from each other (p=.502), but Apple significantly differed from the Fitbit and Actigraph (p=.055 and .968, respectively).

Table 4. Walking Steps (total steps).

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD</th>
<th>Sig. (Two- Tailed) Compared to Counter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counter</td>
<td>4135.0±230.1</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge(^a)</td>
<td>3855.8±373.9</td>
<td>.002</td>
</tr>
<tr>
<td>Apple Watch(^a)</td>
<td>3988.4±372.2</td>
<td>.101</td>
</tr>
<tr>
<td>Actigraph</td>
<td>3996.4±875.7</td>
<td>.490</td>
</tr>
</tbody>
</table>

a. Significant difference between commercial devices (p<0.01).
During the two mile run, the Fitbit Surge, Apple, and Actigraph did not significantly differ from the Counter (Table 5). The Fitbit and Actigraph and the Apple and Actigraph did not significantly differ from each other (p=.678 and .467; respectively), while the Fitbit and Apple significantly differed from each other (p=.001).

**Table 5. Jogging Steps (total steps).**

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD</th>
<th>Sig. (Two-Tailed) Compared to Counter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counter</td>
<td>3219.9±215.5</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3215.6±196.9</td>
<td>.851</td>
</tr>
<tr>
<td>Apple Watch&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3283.6±225.6</td>
<td>.056</td>
</tr>
<tr>
<td>Actigraph</td>
<td>3148.3±206.5</td>
<td>.163</td>
</tr>
</tbody>
</table>

<sup>a</sup>. Significant difference between commercial devices (p<0.01).

**4.4 Heart Rate**

The Apple Watch, Fitbit Surge and Actigraph all estimated HR and were compared to the ScottCare Telemetry System, to determine their accuracy. During the walk, the Fitbit Surge, Apple, and Actigraph did not significantly differ from the ScottCare system (Table 6). The Fitbit vs Apple, Fitbit vs Actigraph, and Apple vs Actigraph also did not significantly differ from each other (p=.728, .562, and .880; respectively).
### Table 6. Walking Heart Rate (Avg. HR in bpm).

<table>
<thead>
<tr>
<th>Device</th>
<th>Mean±SD</th>
<th>Significance (Two-Tailed) Compared to ScottCare</th>
</tr>
</thead>
<tbody>
<tr>
<td>ScottCare</td>
<td>109.1±15.2</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge</td>
<td>110.6±18.0</td>
<td>.649</td>
</tr>
<tr>
<td>Apple Watch</td>
<td>109.1±17.4</td>
<td>1.000</td>
</tr>
<tr>
<td>Actigraph</td>
<td>108.8±16.0</td>
<td>.625</td>
</tr>
</tbody>
</table>

During the jog, the Fitbit Surge and Actigraph did not significantly differ from ScottCare, while the Apple Watch significantly overestimated HR (Table 7). There was a significant difference between the Fitbit and Apple Watch (p= .0001) and the Apple Watch and Actigraph (p= .006), while there was no significant difference between the Fitbit and Actigraph (p= .203).

### Table 7. Jogging Heart Rate (Avg. HR in bpm).

<table>
<thead>
<tr>
<th>Device</th>
<th>Mean±SD</th>
<th>Significance (Two-Tailed) Compared to ScottCare</th>
</tr>
</thead>
<tbody>
<tr>
<td>ScottCare</td>
<td>157.4±14.0</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge</td>
<td>152.8±14.2</td>
<td>.117</td>
</tr>
<tr>
<td>Apple Watch</td>
<td>165.2±16.0</td>
<td>.000(1)</td>
</tr>
<tr>
<td>Actigraph</td>
<td>157.8±22.1</td>
<td>.895</td>
</tr>
</tbody>
</table>

a. Significant difference from ScottCare (p<0.01)
b. Significant difference from the other two commercial devices (p<0.01).
4.5 Distance

The Apple Watch and Fitbit Surge both measured distance and were compared to a calibrated treadmill to determine their accuracy. During the walk, the Apple Watch did not significantly differ from the treadmill distance, while the Fitbit Surge did significantly underestimated distance (Table 8). The Fitbit also significantly underestimated distance compared to the Apple Watch (p=.0001).

Table 8. Walking Distance (mph).

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD</th>
<th>Significance (Two-Tailed) Compared to Treadmill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill</td>
<td>2.00±0.0</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surgeab</td>
<td>1.76±0.2</td>
<td>.000(1)</td>
</tr>
<tr>
<td>Apple Watchb</td>
<td>2.01±0.1</td>
<td>.793</td>
</tr>
</tbody>
</table>

a. Significant difference from Treadmill (p<0.02).
b. Significant difference between the two commercial devices (p< 0.02).

During the run, the Apple Watch did not significantly differ from the treadmill, while the Fitbit Surge significantly underestimated distance (Table 9). The Fitbit also significantly underestimated distance compared to the Apple Watch (p=.0001).
Table 9. Jogging Distance (mph).

<table>
<thead>
<tr>
<th></th>
<th>Mean±SD</th>
<th>Significance (Two-Tailed) Compared to Treadmill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill</td>
<td>2.00±0.0</td>
<td></td>
</tr>
<tr>
<td>Fitbit Surge$^{ab}$</td>
<td>1.79±0.1</td>
<td>.000(1)</td>
</tr>
<tr>
<td>Apple Watch$^b$</td>
<td>2.09±0.2</td>
<td>.047</td>
</tr>
</tbody>
</table>

a. Significant difference from Treadmill (p<0.02)
b. Significant difference between the two commercial devices (p<0.02).

4.6 Discussion

This study examined the accuracy of energy expenditure, steps, heart rate, and distance estimated by the Fitbit Surge, Apple Watch, and Actigraph. Since the Fitbit Surge and Apple Watch were recently released, at the beginning of this study there were no studies comparing their accuracy. There were few studies examining HR and distance estimation of any fitness accelerometers.

During the two mile walk, both the Fitbit Surge and Actigraph overestimated calories burned (106 and 56 kcal; respectively). Similarly, during the two mile jog, both the Fitbit Surge and Actigraph overestimated calories burned (45 and 33 kcals; respectively). These results of energy expenditure supported the findings of Crouter et al. (2003), Groot & Nieuwenhuizen (2013), and Noah et al. (2013) who found that fitness devices were not accurate and tended to overestimated. In this study, there was no significant difference in estimated energy expenditure between the Cosmed and Apple Watch for either walking or jogging. This finding supports Lee, Kim, & Elk (2014) who
found no significant difference in estimating energy expenditure using fitness activity devices.

Many people who wear physical activity trackers primarily use them as a step counter to obtain the ultimate goal of 10,000 steps/day. During the two mile walk and jog, there was no significant difference between any of the devices, and the hand counter. All three of the devices were accurate in estimating steps. This supports the findings of Giannakidou et al. (2011), Mammen et al. (2012), Masurier & Tudor-Locke (2003), Noah et al. (2013), Schneider et al. (2003), and Takas et al. (2014) who all found that fitness devices were accurate in estimating steps. However, Crouter et al. (2003), Eslger et al. (2007), and Melanson et al. (2007) found that the fitness devices were accurate in estimating steps, only at faster speeds.

For the walk, there was no significant difference in estimated HR in any of the devices, compared to the ScottCare system. However, during the jog, the Apple Watch overestimated HR by 8 beats per minute (bpm). There are few studies assessing the accuracy of HR estimated by fitness devices. However, Goodie et al. (2000) and Johnstone et al. (2012) examined Polar HR monitors and found that they were valid in estimating HR. The findings of this study support this because the Actigraph used a Polar HR strap and the results were valid in estimating HR.

Many athletes and recreationally active individuals use fitness devices to track distance. The Fitbit Surge underestimated the two mile distance compared to the calibrated treadmill by .25 miles for the walk and .21 miles during the jog. There are few studies assessing the accuracy of distance estimated by fitness devices. The Fitbit results
support the findings of Gianakidou et al. (2011) and Takas et al. (2014) who found that the devices are not valid in estimating distance.

The Fitbit Surge and Apple Watch were recently released to the public, and there is no research comparing the two devices to each other. The Fitbit Surge estimated 117 more calories compared to the Apple Watch during the walk, and 34 more during the jog. There was not much difference between the Fitbit and Apple in estimating steps during the walk and jog. During the walk, there was no difference in HR between the two devices. However, during the jog, the Apple Watch was 12 bpm higher. The Fitbit underestimated distance (~.30 miles) compared to the Apple Watch. When comparing the two devices to each other, the Fitbit Surge overestimated calories and underestimated distance compared to the Apple Watch. There was no significant difference between the two devices, for steps and walking HR. The Apple Watch overestimated HR during the jog.
CHAPTER V

SUMMARY AND CONCLUSION

5.1 Summary

Indirect calorimetry, hand counting, ScottCare telemetry ECG system, and a calibrated treadmill are considered to be the gold standards in measuring energy expenditure, steps, HR, and distance, respectively. Fitness tracking devices are a great way for an individual to have a “ballpark” estimate of these variables. Fitness devices have become very popular over the years and have motivated people to become more active and aware of their health. Even though they are not always accurate compared to the gold standards, they still provide a good estimate and can be beneficial for tracking activity variables. These devices are geared towards all age groups and are pretty simple to use.
5.2 Conclusion

The results of this study indicate that the Apple Watch was the most accurate device for most variables. The Apple Watch was accurate in estimating energy expenditure, steps, and distance during the walk and jog, and was accurate in estimating HR during the walk, but not during the jog. The Fitbit Surge and ActiGraph had similar results to each other. The Fitbit Surge was accurate in estimating steps and HR during the walk and jog, but not in estimating energy expenditure or distance during the walk or jog. The Actigraph was also not accurate in estimating energy expenditure during the walk or run, but was accurate in estimating steps and HR. The results support previous research that fitness devices are not always accurate. However, they still provide an estimation of calories, steps, HR, and distance, and serves as a reminder to move throughout the day and be aware of one’s activity levels.

5.3 Limitations

There were a few limitations noted during this study. The wristband on the Fitbit Surge and Apple Watch were too large for some of the female subjects. The researcher had to tape down the band, but the tape fell off due to sweating, especially during the running. Another limitation of this study was that some subjects would held the sides of the treadmill for brief periods of time, which could have altered steps being counted on the devices. Also, with using any type of technology, there is always a chance for error of the computer/system, and there could also have been human error with counting steps.
5.4 Future Research Recommendations

The Fitbit Surge and Apple Watch both have a GPS feature. Future research should test subjects outdoors to see if the distance is more accurate than on a treadmill. Future studies should allow subjects to self-select their speed. Some of the subjects were elite runners and the speeds were too slow for them. For future research, purchasing a small and large band for the Fitbit Surge and Apple Watch would be better, especially when testing female subjects. When counting steps, having two counters and/or a video camera to double check the number of steps is recommended. Future studies should also test the different exercise modes (i.e. elliptical, weights, etc.) of the devices to determine their accuracy. Future studies should test a variety of fitness levels as well.

5.5 Application

Many individuals wear fitness devices to track calories burned, steps, HR, and distance. The results indicate that the Apple Watch is generally the most accurate device although it is also the most expensive of the three devices.
REFERENCES


APPENDICES
Research Study Participants Needed!

Cleveland State University graduate thesis student needs volunteers (18-35 years old) who can participate in a research study on the accuracy of the Apple Watch, Fitbit Surge, and ActiGraph on estimating energy expenditure, steps, distance, and heart rate.

This study involves jogging (6.0mph at 0% grade) and walking (3.5 mph at 0%) on a treadmill for 2 miles.

If interested please contact Sarah Kirk at s.e.kirk@vikes.csuohio.edu
INFORMED CONSENT

Comparison and Accuracy of Apple Watch, Actigraph, and Fitbit Surge

Introduction

You have volunteered to participate in a study conducted by student and faculty researchers from the Cleveland State University’s Health and Human Performance Laboratory. This research is being conducted by an Exercise Science graduate student, Sarah Kirk, and Dr. Kenneth Sparks, Associate Professor in the Department of Health and Human Performance.

The purpose of this project is to compare the accuracy of the Apple Watch, Actigraph exercise monitor, and Fitbit Surge in estimating energy expenditure, steps, distance, and heart rate while walking and running.

This project has four main goals: 1) measure the accuracy of the Apple Watch on measuring energy expenditure, steps, distance, and heart rate. 2) measure the accuracy of the Fitbit Surge on measuring energy expenditure, steps, distance, and heart rate. 3) measure the accuracy of the Actigraph on measuring energy expenditure, steps, and heart rate. 4) compare the Apple Watch, Fitbit Surge, and Actigraph on measuring energy expenditure, steps, distance, and heart rate.
Risks

I understand that the risks associated with this study include the possibility of muscle soreness, shortness of breath, fatigue, heart attack, overtraining or acute injuries resulting from the exercise. This risk would be the same experienced from my normal training regimen. I understand that every effort will be made to minimize these risks. I also understand that the laboratory is equipped with an AED and all lab personnel are trained in CPR and First Aid. Emergency procedures include calling EMS (x911) stating to the dispatcher: “We have a medical emergency in the Human Performance Laboratory PE Building- Room B60”. CPR/First aid will be administered until EMS arrives. I also understand that I can voluntarily stop if I experience any problems during the protocol. I know every effort will be made to minimize any risks.

Benefits

I understand that there are no direct benefits for participating in the study. The results of this study will be beneficial to the general population who are seeking an active lifestyle. The results of the accuracy on these devices could be utilized by young and old populations who are looking to live a healthier life.

Procedures

I understand will I complete a pre-screening questionnaire to determine my medical history. The information you submit and contained therein will be used in the determination of your eligibility for participation in the study. I understand that I will participate in two sessions that require approximately one hour. At the start of each testing session, after all the calibration is completed, the researcher will attach the Apple Watch, Actigraph, and the Fitbit Surge to me; as well as the Cosmed (k4b2) and Scott Care telemetry system. All devices will be started at the same time and will be worn in the same location for testing session one and two. During testing session one, I will walk on the treadmill at 3.5mph at 0% grade for two miles. Once two miles are complete, all devices will be stopped and all data will be recorded. Based on stride length and speed, the actual number of steps will be determined and recorded. I will wait at least 48-hours before returning back to the laboratory for day two of testing. For testing session two, I will jog two miles at 6.0mph at 0% grade. Once two miles is completed, all devices will be stopped and data will be recorded. Based on stride length and speed, the actual number of steps will determined and recorded.

Confidentiality

I understand that any data and information obtained during my participation will be confidential and will not be disclosed to anyone without my consent. However, I agree
that my data may be used for research purposes in aggregate form, and that I will not be identified.

**Participation**

I understand that participation in this project is voluntary and that I have the right to withdraw at any time with no consequences.

I understand that if I have any questions about my rights as a participant, I can contact Cleveland State University’s Review Board at (216) 687-3630.
If I have any questions about the procedures I can contact Dr. Sparks at 216-687-4831 or graduate student Sarah Kirk at 216-403-6788.
I attest that I am 18 years of age and that I have no known health problems that could prevent me from successfully participating in the study.

**Patient Acknowledgement**

The procedures, purpose, known discomforts and risks, possible benefits to me and to others have been explained to me. I have read the consent form or it has been read to me, and I understand it. I also understand that all data, even data collected to determine eligibility for the study will be stored in a secured file cabinet in the Human Performance Laboratory for at least 5 years then shredded.

Signature: ________________________________  Date: __________
Witness: ________________________________  Date: __________
Appendix C

AHA/ACSM Health/Fitness Facility Preparticipation Screening Questionnaire

Assess your health needs by marking all true statements.

History
You have had:
― A heart attack
― Heart surgery
― Cardiac catheterization
― Coronary angioplasty (PTCA)
― Pacemaker/implantable cardiac defibrillator/rhythm disturbance
― Heart valve disease
― Heart failure
― Heart transplantation
― Congenital heart disease

If you marked any of the statements in this section, consult your physician or other appropriate healthcare provider before engaging in exercise. You may need to use a facility with a medically qualified staff.

Other health issues
― You have diabetes
― You have or asthma other lung disease.
― You have burning or cramping in your lower legs when walking short distances.
― You have musculoskeletal problems that limit your physical activity.
― You have concerns about the safety of exercise.
― You take prescription medication(s).
― You are pregnant.

Symptoms
― You experience chest discomfort with exertion.
― You experience unreasonable breathlessness.
― You experience dizziness, fainting, blackouts.
― You take heart medications.

Cardiovascular risk factors
― You are a man older than 45 years.
― You are a woman older than 55 years, you have had a hysterectomy, or you are postmenopausal.
― You smoke, or quite within the previous 6 mo.
― Your BP is greater than 140/90.
― You don’t know your BP.
― You take BP medication.
― Your blood cholesterol level is >200 mg/dL.
― You don’t know your cholesterol level.
― You have a close blood relative who had a heart attack before age 55 (father or mother) or age 65 (mother or sister).
― You are physically inactive (i.e., you get less than 30 min. of physical activity on at least 3 days per week).
― You are more than 20 pounds overweight.

If you marked two or more of the statements in this section, you should consult your physician or other appropriate healthcare provider before engaging in exercise. You might benefit by using a facility with a professionally qualified exercise staff to guide your exercise program.

None of the above is true.

You should be able to exercise safely without consulting your physician or other healthcare provider in a self-guided program or almost any facility that meets your exercise program needs.


Lippincott Williams and Wilkins http://www.lww.com

www.acsm-mss.org/pnt-core/template-journal/mssemedia/0588e.htm
Appendix D

Apple Watch
Appendix E

Fitbit Surge
Appendix F

Actigraph GT9X Link
Appendix G

Cosmed K4b2
Appendix H

ScottCare Telemetry System