

# A Multitrait-Multimethod (MTMM) Study of Fitness Assessments in College Students

Peter D Hart<sup>1,2,3\*</sup>, Gabriel Benavidez<sup>1,2</sup>, Nickie Detomasi<sup>1</sup>, Andrew Potter<sup>1</sup>, Kilby Rech<sup>1</sup>, Cory Michael Budak<sup>1</sup>, Natalee Faupel<sup>1</sup>, Jacy Thompson<sup>1</sup>, Laramie Schwenke<sup>1</sup>, Garrett Jericoff<sup>1</sup>, Malcolm Manuel<sup>1</sup>, Trevail Lee<sup>1</sup>, Warren Edmonson<sup>1</sup>, Cydney Auzenne<sup>1</sup>, Taruha Kirkaldie<sup>1</sup>, Michelle Lonebear<sup>1</sup> and Linda Miller<sup>1</sup>

<sup>1</sup>Health Promotion Program, Montana State University, USA

<sup>2</sup>Kinesmetrics Lab, Montana State University, USA

<sup>3</sup>Health Demographics, USA

## Article Information

Received date: Oct 31, 2017

Accepted date: Nov 21, 2017

Published date: Nov 22, 2017

## Corresponding author(S)

Peter D Hart, Health Promotion, College of Education, Arts & Sciences and Nursing, Montana State University - Northern, USA, Tel: 406.265.3719; Email: peter.hart@msun.edu

**Distributed under** Creative Commons CC-BY 4.0

**Keywords** MTMM; Physical fitness; Validity; Reliability; Measurement

## Abstract

**Purpose:** In the fitness industry, several different traits are measured using many different tests. Therefore, a need exists to study and evaluate the convergent and divergent validity of different fitness traits across their tests. The purpose of this study was to examine the measurement properties of different fitness tests designed to assess the five components of health-related physical fitness.

**Methods:** A total of N=131 college students attending a rural public institution participated in this cross-sectional study. Four different fitness tests were administered for each of the five fitness traits: cardiorespiratory, muscular strength, muscular endurance, body composition, and flexibility. A modified MultiTrait-MultiMethod (MTMM) matrix was used to simultaneously examine the measurement properties of the fitness assessments, which included internal consistency reliability, convergent validity, and divergent validity.

**Results:** The overall modified MTMM matrix indicated strong internal consistency ( $\alpha=0.89$ ) across the twenty fitness tests. Each fitness component showed at least moderate reliability ( $\alpha$ s=0.68-0.88) with the exception of flexibility ( $\alpha=0.38$ ). Same trait convergent validity coefficients ( $CV_{ST}$ ) were significant ( $p<0.05$ ) for all traits with exception of flexibility. Majority of different trait convergent validity coefficients ( $CV_{DT}$ ) were significant for all traits with exception of flexibility.

**Conclusions:** Results from this study provide moderate to strong validity evidence for fitness assessments in college students. However, several tests appear to lack strong convergence with their same trait counterpart tests. Furthermore, flexibility appears to lack convergence with its same trait tests as well as other fitness trait tests.

## Introduction

Health-related Physical Fitness (PF) is a set of traits related to an ability to carry out specific tasks with a certain performance level and that are also related to good health [1]. There are five (5) components of health-related PF [2]. These components consist of (1) cardiorespiratory endurance, (2) muscular strength, (3) muscular endurance, (4) body composition, and (5) flexibility. Many different assessments exist to measure these components of health-related PF, both lab-based and field-based, with the latter being much more common in school-based PF assessments [3]. Separately, most field-based PF assessments have shown to be valid and reliable tests [4]. However, in a norm-referenced context, less is known regarding our ability to rank individuals equally across multiple same trait PF tests.

The Multitrait-Multimethod (MTMM) matrix is a powerful multivariate technique that allows for the simultaneous evaluation of different measurement properties of multiple tests and multiple constructs [5,6]. In the fitness industry, different PF traits are measured using many different techniques. Therefore, a need exists to study and evaluate the convergent and divergent validity of different fitness traits across their tests. The purpose of this study was to examine the measurement properties of different fitness tests designed to assess the five components of health-related PF.

## Methods

### Participants

Participants were recruited via campus flyers and word-of-mouth. Inclusion criteria included being an enrolled college student, acceptable responses to the PAR-Q screening tool [7], and a willingness to undergo testing spanning twenty different assessments. After reading a study flyer, participants signed an IRB approved consent form and were scheduled for their battery of fitness tests.

## Research design

This was a cross-sectional assessment study. Trained exercise testing research assistants administered all fitness assessments. Each assistant administered the complete test battery to at least one individual prior to data collection, which served as a pilot run for each tester. Each study participant completed their testing within a three day period and followed a logical order that would always increase each test's performance and score reliability.

## Physical fitness test battery

Four different fitness tests were administered for the Cardiorespiratory (CR) trait. These included (1) Multi-stage fitness (Beep) test [8], (2) Queens College step (Step) test [9], (3) Ebbeling VO2max TreadMill (TM) test [10], and (4) George Non-Exercise (NE) test [11]. All CR test variables were measured in units of ml/kg/min. Four different fitness tests were administered for the Muscular Strength (MS) trait. These tests included (1) Hand grip (Grip) test [12], (2) 1RM Bench Press (1BP) test [13], (3) 1RM Leg Press (1LP) test [13], and (4) Vertical Jump (VJ) test [14]. The two 1RM MS test variables were measured in pounds (lb), the grip test was measured in kilograms (kg), and vertical jump was measured in inches (in). The four Muscular Endurance (ME) tests included (1) Curl-Up (CU) test [13], (2) Push-Up (PU) test [13], (3) Flexed Arm Hang (FAH) (modified) test [15], and (4) YMCA Bench Press (YBP) test [14]. All ME variables were measured in number (#) of completed repetitions except for flexed arm hang which was measured in seconds (s). Flexibility (FL) tests included (1) Sit-and-Reach (SnR) test [13], (2) Back Scratch (BS) test [16], (3) Side Bend (SB) test [17], and (4) Trunk Lift (TL) test [15]. All FL variables were measured in centimeters (cm). Finally, the four Body Composition (BC) tests included (1) Skinfold-

determined percent body fat (PBF) (SF) test [13], (2) Circumference-determined PBF (CM) test [18,19], (3) Body mass index (BMI) test [13], and (4) Handheld bioelectrical impedance-determined PBF (HH) test [20]. All BC variables were measured in percent (%) except BMI which was measured in kg/m<sup>2</sup>.

## Statistical analyses

A modified MTMM matrix was used in this study to account for its cross-sectional design. Table 1 shows that the different testing methods are grouped by trait and can be evaluated in blocks according to the trait they measure. For example, the upper left corner block contains measurement properties of Trait 1 only. Therefore, the validity coefficients in this block represent same trait convergent validity (CV<sub>ST</sub>). For example, the correlation between two cardiorespiratory tests scores from two different cardiorespiratory tests. In blocks with two traits that are different yet similar in their relationship, different trait convergent validity coefficients (CV<sub>DT</sub>) are evaluated. For blocks with two opposing traits, traits known to have no relationship, Divergent Validity (DV) coefficients are evaluated.

All reliability coefficients in the modified MTMM matrix (see Table 1) are internal consistency (standardized) reliability coefficients (Cronbach alpha coefficients). The alpha coefficient in the title is the overall combined alpha ( $\alpha$ ) of all tests and traits. Each trait in the matrix will also have its own alpha ( $\alpha_t$ ). Finally, within each block, the diagonal represents trait-specific alpha with that method (test) deleted ( $\alpha_D$ ).

Pearson correlation coefficients were used for all convergent and divergent validity coefficients. Spearman correlation coefficients were also computed and determined to be no different from the Pearson coefficients. Therefore, the more traditional coefficients were

**Table 1:** Theoretical Model for the Modified Multitrait-Multimethod (MTMM) Correlation Matrix (Population N = xx,  $\alpha$  = xx).

		Trait 1 ( $\alpha_{T1} = xx$ )				Trait 2 ( $\alpha_{T2} = xx$ )				Trait 3 ( $\alpha_{T3} = xx$ )			
		Method 1	Method 2	Method 3	Method 4	Method 1	Method 2	Method 3	Method 4	Method 1	Method 2	Method 3	Method 4
Trait 1													
	Method 1	$\alpha_D$											
	Method 2	CV <sub>ST</sub>	$\alpha_D$										
	Method 3	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$									
Trait 2	Method 4	CV <sub>ST</sub>	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$								
	Method 1	DV	DV	DV	DV	$\alpha_D$							
	Method 2	DV	DV	DV	DV	CV <sub>ST</sub>	$\alpha_D$						
Trait 3	Method 3	DV	DV	DV	DV	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$					
	Method 4	DV	DV	DV	DV	CV <sub>ST</sub>	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$				
	Method 1	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	DV	DV	DV	DV	$\alpha_D$			
Trait 3	Method 2	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	DV	DV	DV	DV	CV <sub>ST</sub>	$\alpha_D$		
	Method 3	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	DV	DV	DV	DV	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$	
	Method 4	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	CV <sub>DT</sub>	DV	DV	DV	DV	CV <sub>ST</sub>	CV <sub>ST</sub>	CV <sub>ST</sub>	$\alpha_D$

**Note:**  $\alpha$  in title is the overall Cronbach alpha for all methods across all traits.  $\alpha$  in column heading is Cronbach alpha for all methods on that trait only.  $\alpha_D$  on diagonal is Cronbach alpha with that test deleted. CV<sub>ST</sub> is Pearson correlation coefficient representing convergent validity for same trait. CV<sub>DT</sub> is Pearson correlation coefficient representing convergent validity for different trait. DV is Pearson correlation coefficient representing divergent validity for different trait.

reported. Fitness variables were first T-score transformed by sex prior to the MTMM matrix analyses. Student's T statistics for mean comparisons were also used for descriptive purposes. All analyses were performed using SAS version 9.4 [21,22].

## Results

A total of N=131 (Mean age=21.8, SD age=5.1 years) college students participated in this study and completed all twenty fitness tests. Table 2 shows fitness test score descriptive statistics for all participants combined and by sex. Significant differences ( $p < 0.05$ ) between sex were noticed and expected across many fitness tests. For example, it was anticipated that males would test with higher CR, MS, and ME scores. With the exception of the MEYBP test, where males and females used sex-specific loads. It was also anticipated that males would test with lower BC PBF scores. Females, however, had a lower mean BMI as compared to males ( $p < 0.05$ ). Females also outperformed males in two FL tests, including SnR and BS ( $p < 0.05$ ).

Table 3 contains results for the MTMM analyses. The overall Cronbach alpha for all twenty fitness tests was strong ( $\alpha = 0.89$ ), indicating high consistency across the test battery. Internal consistency reliability was moderate to strong for CR ( $\alpha = 0.78$ ), MS ( $\alpha = 0.88$ ), ME ( $\alpha = 0.68$ ), and BC ( $\alpha = 0.87$ ) traits. Reliability was, however, poor for FL ( $\alpha = 0.38$ ). The alpha with test deleted coefficients (the diagonal) did not indicate the need to eliminate any one test from its trait test group. Although the BC trait reliability could modestly improve ( $\alpha$  from 0.87 to 0.93), BMI were removed from its trait test group.

Same trait convergent validity coefficients ( $CV_{ST}$ ) (the triangles in the matrix) were significant ( $p < 0.05$ ) within each fitness trait except FL. As well, different trait convergent validity coefficients ( $CV_{DT}$ ) (the squares in the matrix) were significant for majority of tests across fitness traits, except for FL. CR tests were directly related to most MS and ME tests and indirectly related to most BC tests. MS was directly related to ME and indirectly related to most BC tests, except BMI.

**Table 2:** Descriptive Statistics of All Study Variables by Sex.

Fitness Trait	Overall (N=131)		Males (N=87)		Females (N=44)		t test
	Mean	SD	Mean	SD	Mean	SD	p
<b>Cardiorespiratory (CR)</b>							
Beep (ml/kg/min)	34.8	8.3	36.9	8.1	30.6	7.4	<.001
Step (ml/kg/min)	49	10.7	53.8	9.7	39.3	4	<.001
TM (ml/kg/min)	50.8	10	54.7	8.5	43.3	8.4	<.001
NE (ml/kg/min)	47.7	7.4	50	6.6	43.2	6.9	<.001
<b>Muscular Strength (MS)</b>							
Grip (kg)	47.9	12.7	54.7	8.7	34.4	7.2	<.001
1BP (lb)	186.2	85.9	232	67.4	95.6	23.6	<.001
1LP (lb)	502.1	197.3	593.5	166.2	321.3	110.2	<.001
VJ (in)	20.7	5.7	23.5	4.4	15.2	3.7	<.001
<b>Muscular Endurance (ME)</b>							
PU (#)	32.3	15.1	35.4	16	26.2	11.1	<.001
CU (#)	51	24.2	53.8	22.3	45.5	27.1	0.064
FAH (s)	30.8	19.5	33.7	18.9	25	19.7	0.015
YBP (#)	33.8	15.5	33.1	13.5	35.1	19	0.49
<b>Body Composition (BC)</b>							
SF (%)	17	7.7	13.9	6.7	23.2	5.2	<.001
CM (%)	20.9	8.5	17.3	6.9	28	6.7	<.001
HH (%)	20.1	7.4	17.9	6.9	24.6	6.4	<.001
BMI (kg/m <sup>2</sup> )	26.6	4.6	27.7	4.6	24.4	3.5	<.001
<b>Flexibility (FL)</b>							
SnR (cm)	30.8	8.9	29	8.9	34.4	8	<.001
BS (cm)	-0.3	8.3	-2.4	8.7	4	5.5	<.001
SB (cm)	25.5	5.8	25.7	6.2	25.2	5	0.681
TL (cm)	30	9	29.3	7.6	31.3	11.2	0.227

**Note:** t test represents differences between sex. Beep is the beep test. Step is the Queens College step test. TM is the Ebeling treadmill test. NE is the George non-exercise test. Grip is the dynamometer hand grip test. 1BP is the 1RM bench press test. 1LP is the 1RM leg press test. VJ is the vertical jump test. PU is the push-up test. CU is the curl-up test. FAH is the flexed arm hang test. YBP is the YMCA bench press test. SF is percent body fat (PBF) by skinfold method. CM is PBF by circumference method. HH is PBF by handheld bioelectrical impedance method. BMI is body mass index. SnR is the sit and reach test. BS is the back scratch test. SB is the side bend test. TL is the trunk lift test.

**Table 3:** Modified Multitrait-Multimethod (MTMM) Correlation Matrix for Four Fitness Test Methods Across Five Fitness Traits (Overall  $\alpha = .89$ ,  $N=131$ ).

		CR ( $\alpha_{CR} = .78$ )				MS ( $\alpha_{MS} = .88$ )				ME ( $\alpha_{ME} = .68$ )				BC ( $\alpha_{BC} = .87$ )				FL ( $\alpha_{FL} = .38$ )			
		Beep	Step	TM	NE	Grip	1BP	1LP	VJ	PU	CU	FAH	YBP	SF	CM	HH	BMI	SnR	BS	SB	TL
CR																					
	Beep	0.66																			
	Step	<b>0.49</b>	0.77																		
	TM	<b>0.4</b>	<b>0.32</b>	0.77																	
MS	NE	<b>0.71</b>	<b>0.38</b>	<b>0.48</b>	0.69																
	Grip	<b>0.25</b>	<b>0.38</b>	<b>0.57</b>	<b>0.38</b>	0.84															
	1BP	<b>0.34</b>	<b>0.52</b>	<b>0.49</b>	<b>0.33</b>	<b>0.69</b>	0.8														
	1LP	0.16	<b>0.35</b>	<b>0.43</b>	0.14	<b>0.68</b>	<b>0.77</b>	0.86													
ME	VJ	<b>0.58</b>	<b>0.51</b>	<b>0.5</b>	<b>0.61</b>	<b>0.58</b>	<b>0.71</b>	<b>0.44</b>	0.88												
	PU	<b>0.45</b>	<b>0.35</b>	<b>0.28</b>	<b>0.52</b>	<b>0.35</b>	<b>0.49</b>	<b>0.31</b>	<b>0.51</b>	0.55											
	CU	<b>0.43</b>	<b>0.22</b>	<b>0.28</b>	<b>0.36</b>	<b>0.18</b>	<b>0.19</b>	<b>0.22</b>	<b>0.29</b>	<b>0.31</b>	0.66										
	FAH	<b>0.49</b>	<b>0.22</b>	<b>0.27</b>	<b>0.61</b>	<b>0.25</b>	<b>0.2</b>	0.07	<b>0.49</b>	<b>0.49</b>	<b>0.33</b>	0.62									
BC	YBP	<b>0.2</b>	-0.01	0.15	0.16	0.12	<b>0.31</b>	<b>0.28</b>	<b>0.22</b>	<b>0.45</b>	<b>0.3</b>	<b>0.23</b>	0.64								
	SF	<b>-0.65</b>	<b>-0.42</b>	<b>-0.54</b>	<b>-0.77</b>	<b>-0.55</b>	<b>-0.42</b>	<b>-0.26</b>	<b>-0.68</b>	<b>-0.5</b>	<b>-0.34</b>	<b>-0.64</b>	-0.12	0.82							
	CM	<b>-0.56</b>	<b>-0.47</b>	<b>-0.39</b>	<b>-0.68</b>	<b>-0.47</b>	<b>-0.38</b>	<b>-0.14</b>	<b>-0.68</b>	<b>-0.42</b>	<b>-0.27</b>	<b>-0.63</b>	0.06	<b>0.8</b>	0.81						
	HH	<b>-0.63</b>	<b>-0.32</b>	<b>-0.39</b>	<b>-0.77</b>	<b>-0.35</b>	<b>-0.28</b>	-0.03	<b>-0.63</b>	<b>-0.48</b>	<b>-0.28</b>	<b>-0.69</b>	-0.08	<b>0.81</b>	<b>0.84</b>	0.76					
FL	BMI	<b>-0.32</b>	<b>0.19</b>	0.04	<b>-0.48</b>	<b>0.28</b>	<b>0.45</b>	<b>0.56</b>	-0.03	-0.11	-0.12	<b>-0.48</b>	0.06	<b>0.38</b>	<b>0.38</b>	<b>0.56</b>	0.93				
	SnR	0.13	-0.05	0.04	0.08	-0.09	-0.02	-0.05	-0.07	<b>0.26</b>	0.14	0.16	<b>0.39</b>	0.01	0.1	-0.01	-0.1	0.29			
	BS	-0.06	<b>-0.2</b>	<b>-0.24</b>	-0.01	<b>-0.34</b>	<b>-0.33</b>	<b>-0.46</b>	-0.13	-0.13	0.05	0.14	0.05	0.03	-0.03	-0.05	<b>-0.41</b>	<b>0.18</b>	0.34		
	SB	0.11	0.04	0.09	<b>0.2</b>	0.05	-0.03	-0.02	0	0.12	<b>0.22</b>	0.08	0.04	-0.12	-0.12	-0.11	-0.06	0.09	0.04	0.36	
TL	TL	-0.09	<b>-0.17</b>	0.05	-0.04	0.02	-0.02	<b>0.19</b>	-0.16	-0.1	<b>0.2</b>	0.05	0.14	0.02	0.1	0.1	0.04	<b>0.18</b>	0.13	<b>0.19</b>	0.25

**Note:**  $\alpha$  values are Cronbach alpha reliability coefficients. Diagonal values are Cronbach alpha coefficients with test deleted. All other values are bivariate Pearson correlation coefficients. Bold values are significant ( $p < .05$ ). Beep is the beep test. Step is the Queens College step test. TM is the Ebbeling treadmill test. NE is the George non-exercise test. Grip is the dynamometer hand grip test. 1BP is the 1RM bench press test. 1LP is the 1RM leg press test. VJ is the vertical jump test. PU is the push-up test. CU is the curl-up test. FAH is the flexed arm hang test. YBP is the YMCA bench press test. SF is percent body fat (PBF) by skinfold method. CM is PBF by circumference method. HH is PBF by handheld bioelectrical impedance method. BMI is body mass index. SnR is the sit and reach test. BS is the back scratch test. SB is the side bend test. TL is the trunk lift test.

Finally, ME was indirectly related to most BC tests. Although FL tests lacked convergent validity as a whole BS was however significantly and indirectly related to three MS tests (Grip, 1BP, and 1LP).

## Discussion

The aim of this study was to examine the measurement properties of different fitness tests designed to assess the five components of health-related physical fitness. Overall, the results demonstrated many expected relationships within and between measured fitness traits. For example, the CR trait tests exhibited moderately strong internal consistency reliability, which was robust to any single test being removed from the CR battery. As well, all CR tests showed significant  $CV_{ST}$  coefficients, supporting their ability to validly measure the CR trait. These findings were consistent across the MS, ME, and BC traits.

However, and more noteworthy, was the relationship between

the fitness tests and different trait tests. For example, the CR tests significantly converged ( $CV_{DT}$ ) with most MS tests, most ME tests, and most BC tests. The fitness trait tests converging with other trait tests does in fact make sense. One study with similarly aged men showed significant correlations between maximal bench press (MS) and both push-up and sit-up scores (ME) [23]. This study also showed that repeated squats (ME) significantly correlated with cycle ergometer test values (CR). Finally, this study showed that percent body fat (BC) was significantly correlated with tests of ME. Therefore, it would be appropriate to consider tests from the CR, MS, ME, and BC traits as tests that significantly converge with each other and hence appropriate to term their validity coefficients as  $CV_{DT}$ .

Equally noteworthy, and not expected, was the lack of psychometric evidence regarding the FL trait. That is, the FL tests lacked internal consistency reliability, same trait convergent validity, and different trait convergent validity. With only a few exceptions (i.e., BS related to MS tests and the BC BMI test), FL would appear to have no relationship with the other four PF traits. Having shown this,

the lack of reliability and same trait validity can be justified, due to the physiological fact that flexibility is joint specific [2]. Therefore, an individual that ranks high in terms of lower-body flexibility may not necessarily rank high in terms of upper-body flexibility.

The lack of different trait convergent validity is, however, less understood. This may simply reflect the fact that flexibility training is promoted less on college campuses than other forms of training (i.e., exercise facilities with CR, MS, and ME equipment but no designated stretching space). Another possible explanation may be that individuals who do focus on flexibility are less interested in improving the other fitness traits (and vice versa). In any event, more research is suggested here to address the lack of convergent validity of FL tests with other trait tests. Given these results, it would be appropriate to consider FL as a multidimensional trait, where each joint-specific test measures its own construct. Additionally, tests from the FL trait may be considered as tests that diverge from other fitness trait tests and hence appropriate to term their validity coefficients as DV.

The results of this study should be interpreted while considering some limiting factors. Firstly, the psychometric properties assessed in this study are population and situation specific [24]. That is, the measurement results from the MTMM matrix should be considered only for college students who attend a rural public university. Secondly, each of the twenty PF tests administered in this study was field-based techniques and not lab-based techniques. This fact may have limited the results because lab-based techniques serve as “gold-standard” methods in assessing PF traits. Given this fact, our study was based on norm-referenced standards [25] and therefore more concerned about each test’s ability to rank individuals accurately on same trait tests and less concerned about criterion-based standards where test results may be used to diagnose patients. A final limitation in this study was the amount of time and effort required to complete twenty fitness tests by each college student participant. This fact may have presented a limitation by introducing measurement error into the test scores [24]. To reduce measurement error on each of the performance tests, participants would have needed to give maximum effort. And in any testing scenario, the more demand placed on the test taker, the more likely it is that he or she will fatigue. Given this fact, trained research assistants were instructed to schedule the tests across a maximum span of three days and according to principles that would produce the most reliable scores [13].

## Conclusions

Results from this study provide overall moderate to strong validity evidence for fitness assessments in college students. However, several tests appear to lack strong convergence with their same trait counterpart tests. Furthermore, flexibility appears to lack convergence with its own trait fitness tests as well as with different trait tests. This may suggest that the flexibility trait diverges from the other four components of health-related PF.

## References

- American College of Sports Medicine, editor. ACSM's health-related physical fitness assessment manual. Lippincott Williams & Wilkins. 2013.
- American College of Sports Medicine. ACSM's Resources for the Exercise Physiologist, 2nd. Philadelphia, Md.: Lippincott Williams & Wilkins. 2017.
- Raven P, Wasserman D, Squires W, Murray T. Exercise Physiology. Nelson Education. 2012.
- Garber CE, Glass, SC. ACSM's resource manual for guidelines for exercise testing and prescription. Kaminsky LA, Bonzheim KA, editors. Baltimore: Lippincott Williams & Wilkins. 2006.
- Campbell DT, Fiske DW. Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. Psychol bull. 1959; 56: 81-105.
- Bryant FB, Grimm LG, Yarnold PR. Reading and understanding more multivariate statistics. Washington DC: American Psychological Association. 2000; 99-146.
- Shephard RJ. PAR-Q, Canadian Home Fitness Test and Exercise Screening Alternatives. Sports Med. 1988; 5:185-195.
- Ramsbottom R, Brewer J, Williams C. A Progressive Shuttle Run Test to Estimate Maximal Oxygen Uptake. Br Sports Med. 1988; 22: 141- 144.
- McArdle WD, Katch FI, Katch VL. Exercise physiology: nutrition, energy, and human performance. Lippincott Williams & Wilkins. 2010.
- Ebbeling CB, Ward A, Puleo EM, Widrick J, Rippe JM. Development of A Single-Stage Submaximal Treadmill Walking Test. Med Sci Sports Exerc. 199; 23: 966- 973.
- George JD, Stone WJ, Burkett LN. Non-Exercise VO<sub>2</sub>max Estimation for Physically Active College Students. Med Sci Sports Exerc. 1997; 29: 415- 423.
- Suni J, Husu P, Rinne M. Fitness for Health: The ALPHA-FIT Test Battery for Adults Aged 18–69. UKK Institute for Health Promotion Research. 2009; 1- 29.
- Ferguson B. ACSM's Guidelines For Exercise Testing And Prescription 9<sup>th</sup> Edn 2014. Lippincott Williams & Wilkins. JCan Chiropr Assoc. 2014; 58: 328.
- Haff G, Triplett NT. Essentials of strength training and conditioning. 4<sup>th</sup> edition. Human Kinetics. 2016.
- Welk GJ, Meredith MD. Fitnessgram/Activitygram reference guide. The Cooper Institute, Dallas, TX. 2008; 1- 206.
- Wood R. Back Scratch Flexibility test. 2017.
- Oja P, Tuxworth B. Eurofit for adults: A test battery for the assessment of the health-related fitness of adults. Strassbourg: Council of Europe, Committee for the Development of Sport. 1995.
- Hodgdon JA, Beckett MB. Prediction of Percent Body Fat for US Navy Women from Body Circumferences and Height. Naval Health Research Centre. 1984.
- Hodgdon JA, Beckett MB. Prediction of Percent Body Fat for US Navy Men From Body Circumferences and Height. Naval Health Research Center. 1984.
- Omron Fat Loss Monitor. Model HBF-306C. Omron Healthcare Co., Ltd., 2012.
- Tabachnick BG, Fidell LS, Osterlind SJ. Using Multivariate Statistics. 5<sup>th</sup> Edition. Pearson. 2007.
- Cody RP, & Smith, K. Applied statistics and the SAS programming language. 5<sup>th</sup> Edition. Pearson. 2006.
- Vaara JP, Kyröläinen H, Niemi J, Ohrankämmen O, Häkkinen A, Kocay S, Häkkinen K. Associations of Maximal Strength and Muscular Endurance Test Scores with Cardiorespiratory Fitness and Body Composition. J Strength Cond Res. 2012; 26: 2078- 2086.
- Morrow JR, Mood D, Disch J, Kang M. Measurement and Evaluation in Human Performance, 5<sup>th</sup> Edn. Human Kinetics. 2015.
- Wood TM, Zhu W. Measurement theory and practice in kinesiology. Human Kinetics. 2006.