

Validity and Reliability of a Tower Climb Test for the Assessment of
Anaerobic Performance in Urban Firefighters

by

Melissa Clarke

B.Sc., University of Victoria, 1998

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the School of Exercise Science, Physical and Health Education

© Melissa Clarke, 2012

University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by
photocopy or other means, without the permission of the author.

Supervisory Committee

Validity and Reliability of a Tower Climb Test for the Assessment of
Anaerobic Performance in Urban Firefighters

by

Melissa Clarke

B.Sc., University of Victoria, 1998

Supervisory Committee

Dr. Catherine Gaul (School of Exercise Science, Physical and Health Education)
Supervisor

Dr. Lynne Stuart-Hill (School of Exercise Science, and Physical and Health
Education)
Departmental Member

ABSTRACT

Supervisory Committee

Dr. Catherine Gaul (School of Exercise Science, Physical and Health Education)
Supervisor

Dr. Lynneth Stuart-Hill (School of Exercise Science, and Physical and Health Education)
Departmental Member

The purpose of this study was to determine the validity and reliability of an 8-flight tower climb test (TCT) to assess anaerobic performance in urban firefighters. Twenty-five professional urban firefighters participated in the validity testing of the TCT versus the Wingate Anaerobic Test (WAT) over 2 randomly sequenced testing sessions. Test-retest reliability was assessed separately in 21 active male and female participants over 2 TCT trials. During both validity and reliability testing for the TCT, participants ascended a firefighting training tower as fast as possible from a 1.7m running start while wearing firefighter protective equipment. Time was measured and power was calculated from the foot of the training tower to the top of the first (height = 1.75m) and eighth (height = 13.89m) flights of stairs. During the other session assessing TCT validity, participants completed a 30-second WAT using a resistance of $85\text{g}\cdot\text{kg}^{-1}$ body weight (BW). Several significant correlations were found including those between TCT power and: 1) mean WAT power generated for the duration equivalent to TCT time ($r = 0.869$), 2) peak power for the first 2 seconds of the WAT ($r = 0.868$), and 3) WAT peak power ($r = 0.864$). TCT test-retest performance in 21 active males and females showed that the test is highly reproducible. The mean time of

completion of the 8-flight TCT was 21.81 ± 5.03 seconds and 21.38 ± 4.86 seconds for Trials 1 and 2, respectively. Intraclass correlations for time and power data from the first and eighth flights ranged from 0.94 to 0.99, and coefficients of variance ranged from 2.0% to 7.5%. These findings provide strong evidence that the TCT is a valid and reliable field-based assessment of occupation-specific anaerobic performance in urban firefighters.

TABLE OF CONTENTS

Supervisory Committee.....	ii
Abstract	iii
Table of Contents.....	v
List of Tables.....	vi
List of Figures	viii
Acknowledgements	ix
Dedication	x
Chapter 1: Introduction	1
Chapter 2: Methods	11
Chapter 3: Results	19
Chapter 4: Discussion	25
References	47
Appendix A: Review of Literature.....	57
Appendix B: Data Collection Sheets	75
Appendix C: Data Collection Protocols	77
Appendix D: Results for TCT Excluding PPE Mass	87
Appendix E: Normative Data for Other Anaerobic Performance Tests.....	88

LIST OF TABLES

<i>Table 1: Mean (SD) Anthropometric Data for Male Firefighters Completing the TCT Validity Study, and Participants Completing the TCT Test-Retest Reliability Study.....</i>	<i>19</i>
<i>Table 2: Mean (SD) Time and Power Values for Flight 1 and Flight 8 For the TCT and WAT During TCT Validation Testing.....</i>	<i>20</i>
<i>Table 3: Pearson Correlation Coefficients for TCT Times and WAT Power During TCT Validation Testing</i>	<i>21</i>
<i>Table 4: Pearson Correlation Coefficients for TCT and WAT Power During TCT Validation Testing</i>	<i>22</i>
<i>Table 5: Mean (SD) Pre-test, 1 and 3 Minute Post-test, and Peak Blood Lactate Values for the WAT and TCT.....</i>	<i>23</i>
<i>Table 6: Mean (SD) Time and Power Output for TCT Reliability Tests (n=21) ...</i>	<i>23</i>
<i>Table 7: Comparison of Absolute and Relative Anaerobic Alactic Power from Present and Previous Research</i>	<i>27</i>
<i>Table 8: Mean (SD) Calculated Flight 1 and Flight 8 Power Output Excluding PPE Mass During TCT Validation Testing.....</i>	<i>87</i>
<i>Table 9: Pearson Correlation Coefficients for TCT and WAT Power Excluding PPE Mass During TCT Validation Testing.....</i>	<i>87</i>

<i>Table 10: Margaria-Kalamen Stair Sprint Test Absolute Power Guidelines for Males</i>	<i>88</i>
<i>Table 11: WAT Peak Power Percentile Norms for Males (n = 62)</i>	<i>88</i>
<i>Table 12: WAT Mean Power Percentile Norms for Males (n = 62)</i>	<i>89</i>

LIST OF FIGURES

Figure 1: Photographs of (a) the 8-flight tower used during the TCT, and (b) a participant completing the TCT while wearing PPE and SCBA..... 13

Figure 2: Reliability of (a) TCT-T1, (b) TCT-T8, (c) absolute TCT-P1, (d) absolute TCT-P8, (e) relative TCT-P1, and (f) relative TCT-P8 during 2 trials of the TCT (n = 21). Intraclass correlations (ICC), typical error (TE), and coefficients of variance (COV) are shown for each test-retest variable24

ACKNOWLEDGEMENTS

The following pages would be mostly blank if it were not for the help, insight, guidance and support of my supervisors, Dr Kathy Gaul and Dr Lynneth Stuart-Hill. Kathy – thank you so much for all of the time you spent talking to me about this research, your attention to detail, your support and compassion toward my schooling and life in general, and your belief in me that I could actually get this done! Lynneth – thank you so much for your guidance and ideas to help pull all of this together, for allowing me to talk through my thought processes, and for being a constant shoulder to lean on through some very challenging times. Also thank you to Greg for your ears and insight, and Lindsay, Jordan, Leanne, and all of the other graduate and undergraduate students who assisted at all of the various testing sessions – it was a dirty job, but someone had to do it!

This study would not have been possible without the support of the Victoria Fire Department. To the guys who stepped up and agreed to donate your body to science, thank you! Special thanks to Paul Bruce and Gary Charlton for going above and beyond to make this happen, and for facilitating an environment that made the coordination of this research run so smoothly. I appreciate your help in gaining the support of the firefighters who participated in this study, and the support of the officers who had to make adjustments to their training schedule to accommodate the research that was being conducted during their shifts and drill times.

Thank you to everyone at CSEE and PISE for always greeting me with a smile and pulling up a chair every time I knocked on your door and wanted to run something past you! I am happy for you that you no longer have to deal with my questions and thought processes!

Last but not least, thank you to all of my family and friends for the unending support that you have shown me over the days, weeks, months, and years to pull this together. Mom, Ang, Sue, Andrea, Jennie, Alyssa, the rest of “my girls”, and the entire Wilson clan, from the bottom of my heart, thank you, thank you, thank you. Your cheering, support, and willingness to lend a hand whenever you could meant and continue to mean more to me than you will ever know. I really truly could not have done this without you!

DEDICATION

Taya – you are the most precious thing I will ever know or have.
Thank you for always helping me to keep things in perspective, and for
continually being able to put a smile on my face. *We did it!*

CHAPTER 1 – INTRODUCTION

Firefighting is a physically demanding profession that incorporates unique challenges and presents a high risk of injury and on-duty death compared to other occupations. Increased physical fitness has been shown to enhance job performance, and decrease injuries and disabilities in firefighters while on-duty (Davis & Dotson, 1987). Research over the past several decades has led to the acceptance and application of minimal physical fitness standards for firefighting, and the development of physical training recommendations by various organizations, including the National Fire Protection Agency (NFPA), International Association of Firefighters (IAFF), and the International Association of Fire Chiefs (IAFC).

Activities of firefighting include lifting and carrying heavy equipment up and down stairs, hoisting ladders, handling and maneuvering charged and uncharged hoses, breaking down doors and walls, fighting fires, auto extrication, property salvage, ceiling overhaul and working with heavy tools overhead, and search and rescue of victims (Hart, 1999; Barr, Gregson, & Reilly, 2010; Gledhill & Jamnick, 1992). Field tests documenting the physiological demands of firefighting tasks have shown that firefighters are required to intermittently work near, at, or above age-predicted maximum heart rate (HR), and at more than 70% maximal oxygen consumption (VO_2 max) for sustained periods of time (Angerer, Kadlez-Gebhardt, Delius, Raluca, & Nowak, 2008; Adams et al., 2009; Smith, Manning, &

Petruzzello, 2001; Williams, Peterson, & Douglas, 1996; Bilzon, Scarpello, Smith, Ravenhill, & Rayson, 2001; Lemon & Hermiston, 1977; Holmer & Gavhed, 2007; Gledhill & Jamnik, 1992; Romet & Frim, 1987).

While there has been considerable attention paid to the assessment of aerobic requirements related to firefighting, there has been less focus on the measurement of anaerobic contributions to occupation-specific metabolic demands. The high intensity, intermittent physical exertion, thermal loads, and the use of personal protective equipment (PPE) and a self-contained breathing apparatus (SCBA) inherent to firefighting increase physical demands, and hence, require increased reliance on anaerobic metabolism. This reliance is greater in firefighters with lower levels of aerobic fitness. Lemon and Herminston (1977) reported 59.84% anaerobic contribution during various firefighting tasks in participants with lower aerobic power ($\text{VO}_2 \text{ max}$ less than $40 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$), compared to 52.73% anaerobic contribution in participants with a $\text{VO}_2 \text{ max}$ greater than $40 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$. During various firefighting activities, firefighters were required to work at up to 97% $\text{VO}_2 \text{ max}$ (Davis & Dotson, 1978), and blood lactate values were reported to be as high as $13 \text{ mmol} \cdot \text{L}^{-1}$ (Gledhill & Jamnick, 1992). These findings provide evidence that firefighters regularly work above their anaerobic threshold and that significant contributions from anaerobic sources are required during firefighting activities.

The inclusion of valid and reliable anaerobic performance testing is warranted given the anaerobic nature of many firefighting tasks. Validity refers to the degree to which a test measures what it reports to measure, and can be determined by comparing two or more variables from one group of participants, while reliability measures the reproducibility of results by testing the same participants with the same methods two or more times (Hopkins, 2000). The use of valid and reliable field tests in physically demanding occupations can assist in matching the demands of an occupation with individuals that are most physiologically capable of handling the workload, and can enhance education, awareness, and training amongst workers (Jackson, 1994).

Acceptable validation strategies for employment testing include studies addressing content and criterion-related validity (Jackson, 1994). Content validity relies on subjective decision-making to consider whether or not a certain test is appropriate to represent the given intention (Baumgartner & Jackson, 1991). Content validity has been used successfully in occupational settings to develop work-sample tests specific to firefighting, including climbing a ladder or stairs, dragging a dummy, and pushing, pulling, and carrying equipment (Jackson, 1994). Criterion-related, or ecological, validity depicts how closely a test is predictive of job performance by analyzing data using a Pearson product-moment correlation coefficient (r) to compare field test scores to a criterion measure (Baumgartner & Jackson, 1991).

The most highly utilized anaerobic performance test is the Wingate Anaerobic Test (WAT) (MacDougall, Wenger, & Green, 1991). The WAT is a 30-second maximal effort cycle test that provides measurements of peak power, mean power, and fatigability to assess anaerobic power of the legs (Inbar, Bar-Or, & Skinner, 1996). The WAT has been used previously to measure the energy requirements of specific firefighting tests, and to describe anaerobic performance of firefighters (Williams-Bell, Villar, Sharratt, & Hughson, 2009; Sheaff et al., 2010; Findley, Brown, & Whitehurst, 2002; Misner et al., 1988).

Williams-Bell et al. (2009) demonstrated that male firefighters who successfully passed the criterion for the Firefighter Candidate Physical Abilities Test (CPAT), a content-based, 8-event circuit of simulated firefighting tasks, produced significantly greater peak power on the WAT compared to males who did not complete the CPAT. Conversely, however, leg power, as measured by the WAT, did not distinguish female firefighters from a non-firefighting female control group (Findley et al., 2002). Findley et al. (2002) concluded that incumbent female firefighters should focus on the development of lower body power to meet occupational demands; however, these findings could also indicate that the WAT did not provide an accurate reflection of occupation-specific anaerobic performance for the female firefighters participating in the study.

Although highly popularized as a gold-standard measure for anaerobic power, use of the WAT may not provide the most appropriate or

feasible assessment of anaerobic firefighting performance. According to MacDougall et al. (1991), test selection should be relevant and specific to the muscles and modes of activity typically used by a population. For firefighters, measurement of anaerobic performance during common firefighting tasks could be more useful than measurement on a cycle ergometer. Use of a cycling test limits the ability to consider the specific physical demands of firefighting, which mainly include weight-bearing locomotion, and eliminates the increased physiological strain imposed by the weight, restrictiveness, and insulation of the PPE.

Anaerobic performance has also previously been assessed through the use of valid and reliable staircase running tests in various populations (see Appendix A). The Margaria and Margaria-Kalamen tests are 2 highly-utilized, short-term anaerobic staircase tests in which participants ascend a single flight of stairs from a running start, 2 or 3 stairs at a time, as fast as possible (Margaria, Aghemo, & Rovelli, 1966; Kalamen, 1968).

Advantages of staircase tests include: inexpensive equipment requirements, minimal training and specialized skill requirements for testers and participants, applicability to a number of populations, and increased compliance from participants due to the simplistic nature of the test (Margaria et al., 1966). Misner et al. (1988) compared a 3-flight stair climb test to the WAT to determine occupation-specific power requirements amongst incumbent female firefighters, but found there were no significant

relationships between WAT peak power, mean power, and fatigue index compared to the stair climb.

Stair climbing has also been incorporated into firefighter fitness testing, such as the CPAT, as well as various firefighter-sanctioned competitions and events, including the Firefit Challenge, the Grind in the City Charitable Stairclimb, and the Stairclimb for Clean Air (*Firefit*, 2010; “Stairclimb for clean air”, 2011). During the CPAT, firefighting applicants wear a weighted vest while stair climbing on a motorized stepmill. While use of a stepmill more closely replicates the demands of firefighting compared to use of a cycle ergometer, it may not accurately reflect the influence of gravity and the power required to move a mass over a vertical distance. In addition, while use of a weighted vest would simulate the increased load that would otherwise be imposed by the PPE and SCBA, it may not provide an accurate representation of the performance limitations caused by the insulative and restrictive properties of the PPE. The weight and bulkiness of the PPE and SCBA are considered contributing factors leading to exhaustion in firefighters and have been shown to increase the physiological demands of a given activity by up to one-third compared to regular clothing (Davis & Dotson, 1987). Mean VO_2 , heart rate, and workload increased by 24% to 35% when measuring stair climbing on a motorized stair-treadmill with gear compared to without gear (Davis & Dotson, 1987). These findings highlight the importance of including PPE during the measurement of firefighter performance. Full PPE is worn

during the Firefit Challenge and other firefighter-specific stair climb events. The Firefit Challenge is an international firefighting competition that, based on the intensity and duration, relies significantly on the anaerobic system (MacDougall et al., 1991). The inclusion of stair climbing in previous firefighting research, and its use in other occupation-specific testing and popular firefighting recreational events illustrates the relevance of stair climbing to firefighter job performance.

Purpose & Rationale

The purpose of this study was to evaluate the use of a Tower Climb Test (TCT) as a valid and reliable field-based measure of anaerobic performance in urban firefighters, and to examine the relationships that exist between TCT time, TCT power, and WAT power. Development of a firefighting occupation-specific test, such as the TCT, that can easily be conducted at fire training facilities will minimize challenges inherent to testing within this population, including accessibility to equipment, ease of testing, and recruitment of research participants. Such a test could provide a more effective and accurate measure to evaluate anaerobic performance in firefighters, contribute to the growing body of research supporting firefighter health and safety, and contribute to the continued development of physical fitness recommendations and guidelines.

Research Questions

1. Is the TCT a valid measure of anaerobic performance?
 - a) What is the relationship between TCT power and WAT power?
 - b) What is the relationship between TCT time and WAT power?
2. Is the TCT a reliable measure of anaerobic performance?

Hypotheses

1. There will be a significant positive correlation between TCT power and WAT power.
2. There will be a significant negative correlation between TCT time and WAT power.
3. The TCT will be a valid and reliable measure of anaerobic performance.

Assumptions

1. Participants adhered to pre-test guidelines before testing sessions.
2. Participants provided maximal effort during the TCT and WAT.
3. Blood lactate concentration reflected anaerobic contributions to performance during the TCT and WAT.

Delimitations

1. Participants for the TCT and WAT validity testing were experienced firefighters from the Victoria Fire Department between 27 and 50 years of age.

2. Participants for the TCT test-retest reliability were active males and females between 21 to 58 years of age.

Limitations

1. Environmental conditions could not be controlled for TCT sessions, but were recorded precisely.
2. Effort during the TCT and WAT validation study was only monitored through lactate analysis and observation of the researchers.

Operational Definitions

Personal Protective Equipment (PPE):

Protective equipment worn by firefighters, including firefighting pants, jacket, bella clava, helmet with a protective face mask, Scott 2.2 harness, gloves, and boots.

Self-Contained Breathing Apparatus (SCBA):

Breathing apparatus utilized by firefighters to provide breathable air in dangerous environments.

Anaerobic Power:

An indicator of anaerobic fitness assessed by the WAT and TCT to represent the amount of work done in a given unit of time.

Peak Power (WAT-PP):

The highest 5-second average power achieved during the WAT. This typically occurs during the first 5 seconds of the test.

2-Second Peak Power (WAT-PP2):

The average power measured during the first 2 seconds of the WAT.

Mean Power (WAT-MP):

The average power measured over the entire 30-second WAT.

Corrected Mean Power (WAT-MP_{TCT}):

The average power output calculated during the WAT for the time equivalent to participant TCT completion time.

Tower Climb Time (TCT-T1 and TCT-T8):

The time to reach the top of the first (-T1) and eighth (-T8) flights of stairs from the foot of the training tower.

Tower Climb Power (TCT-P1 and TCT-P8):

The power calculated at the top of the first (-P1) and eighth (-P8) flights of stairs from the foot of the training tower.

CHAPTER 2 – METHODS

This study was conducted in two major segments to assess the use of a Tower Climb Test (TCT) as a measure of anaerobic fitness in urban firefighters:

- 1) Validation testing of the TCT versus a Wingate Anaerobic Test (WAT)
- 2) Reliability of the TCT by way of test-retest TCT performance

Participants and Recruitment

Twenty-five male professional urban firefighters were recruited for the validation component of this study from the Victoria Fire Department. Twenty-one active male ($n = 12$) and female ($n = 9$) participants, including 2 of the 25 firefighters who participated in TCT validity testing, were recruited for the reliability component of this study. An information session was conducted that addressed all aspects of the study, including rationale, research procedures and measurements, risks and benefits of participation, and answered questions posed by potential participants. Participants provided written informed consent prior to participation. This study was conducted with approval from the University of Victoria Human Research Ethics Board and Biosafety Committee.

Procedures

Data were collected for the validation component of this study between November 2010 and January 2011 (see Appendix B and C).

Within this timeframe, participants completed 2 randomly sequenced, anaerobic testing sessions at the Victoria Fire Department, which included an occupation-specific TCT and a 30-second WAT. Sessions were separated by a minimum of 24 hours. Blood lactate was collected via a finger prick using an auto lancet and analyzed using a blood lactate test meter (Lactate Pro, Arkray Inc., Japan). Resting blood lactate was measured prior to participants donning their PPE (for the TCT) and commencing their circulatory warm-up (for both the TCT and WAT). Sampling also occurred at 1 and 3 minutes post-test. During data analysis, peak lactate was determined as the highest 1 or 3 minute post-test measurement for both the TCT and WAT.

Data were collected for the reliability component of this study between July and September 2011. Within this timeframe, participants completed two TCT trials at the Victoria Fire Department with a minimum of 1 hour rest between trials.

Participants for both the validity and reliability components were given pretest guidelines, which included being well rested, avoiding exercise and alcohol for 24 hours, and attending each testing session properly nourished and hydrated.

Tower Climb Test

Participants completed the TCT using an 8-flight (4-storey) training tower at the Victoria Fire Department while wearing firefighter PPE (see

Figures 1a and b). Participants were weighed before the test in PPE, which included firefighting pants, jacket, boots, bella clava, helmet, gloves, and Scott harness and pack. Participants wore, but did not breathe through, their SCBA.

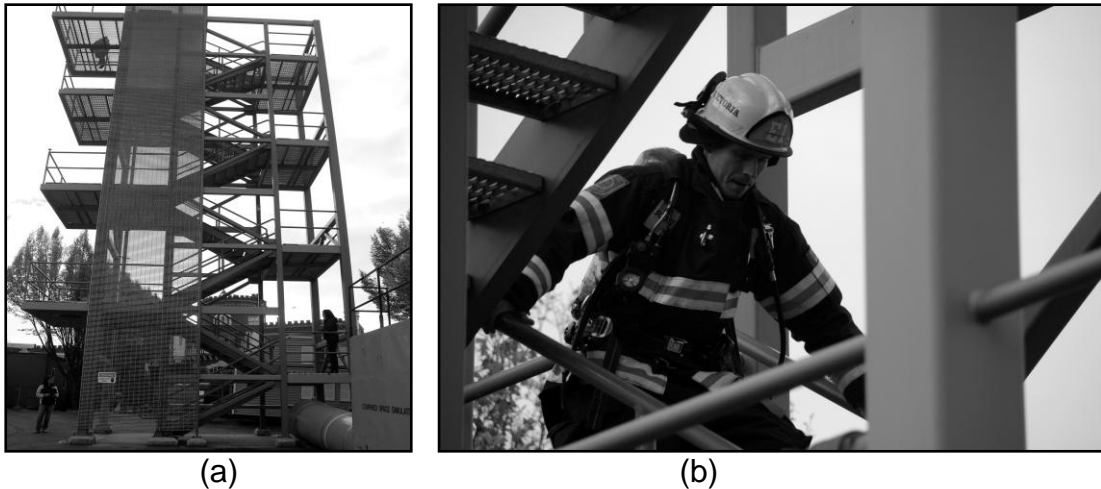


Figure 1. Photographs of (a) the 8-flight tower used during the TCT, and (b) a participant completing the TCT while wearing PPE and SCBA.

After completing a standardized 2 minute warm-up on a stationary cycle, participants were instructed to complete the TCT in the fastest time possible. The start line was positioned 1.7 metres from the foot of the training tower. Timing lights (Brower Timing Systems, Draper, UT) were positioned at the foot of the training tower and at the top of the first (height = 1.75m) and eighth (height = 13.89m) flights of stairs. During ascent, participants were instructed to climb single or multiple stairs at a time but were prohibited from using their hands on the railings to increase upward momentum and pull themselves up the tower. During validation testing, participants touched the handrails at the top of the training tower before

turning around for descent. For safety reasons, participants were instructed to use the handrails for support and their feet were required to touch each stair during the descent.

Wingate Anaerobic Test

Participants in the validity component of the study completed the 30-second WAT at the Victoria Fire Department using a cycle ergometer with a weighted basket loading system (894E Monark, Vansbro, Sweden). WAT testing was completed in exercise clothing. Participants were weighed before the test and a resistance workload was calculated at $85 \text{ g}\cdot\text{kg}^{-1}$ body weight (BW). After completing a 2 minute warm-up between 60 to 80 rpm using zero resistance, participants were instructed to increase their pedal frequency to as fast as possible. The test began when pedal frequency reached 120 rpm, triggering resistance to be applied automatically via computer software (Monark Anaerobic Test Software, Vansbro, Sweden). Participants were instructed and encouraged to perform at maximal intensity for the entire 30-second test duration.

VO₂ Peak Test

Participants who completed the TCT and WAT validation component of this study also completed a walking VO₂ peak test in PPE and carrying a SCBA on a motorized treadmill (Woodway, USA). Data were collected

throughout the test for HR and VO_2 , and were measured as 30 second averages.

Participants completed a 5-minute warm-up at 3.5mph and 0% grade for the first 3 minutes, and 2% grade for the next 2 minutes. Speed remained at 3.5mph throughout the test and grade was increased by 2% every 2 minutes until ventilatory threshold (VT) was reached. VT was considered the point where the V_E/VO_2 ratio increased while V_E/VCO_2 remained relatively constant (Wasserman, 1987). At this point, workload increased by way of elevation every minute. If the participant reached a maximum grade of 16%, the speed was increased by 0.5mph each minute until volitional fatigue when the test was terminated. VO_2 peak was identified as the highest VO_2 averaged over 30 seconds when at least 2 of the following criteria were met:

1. Plateau in VO_2 despite an increase in work rate;
2. Respiratory exchange ratio (RER) > 1.15;
3. HR \geq age-predicted maximum; or
4. Volitional fatigue.

Upon completion of the test, the PPE and SCBA were removed and the participant completed a 5-minute cool down at 2.5mph and 0% grade.

Data Collected

Tower Climb Test

Times (in seconds) were recorded from the foot of the training tower to the:

- 1) top of the first flight (TCT-T1); and
- 2) top of the eighth flight (TCT-T8).

Times were also recorded to return to the foot of the tower after the descent, but were not used in data analysis.

Absolute power (watts) was calculated at the top of the first (TCT-P1) and eighth (TCT-P8) flights of stairs using the formula:

$$P = \frac{\text{Mass (kg)} \times 9.81 \times \text{vertical height of flight (m)}}{\text{Time (sec)}} \quad (\text{Foss \& Keteyian, 1998})$$

Relative power ($\text{W} \cdot \text{kg}^{-1}$) was calculated by dividing absolute power by body weight (kg), including all PPE and SCBA gear.

Blood lactate was measured during the validity testing pre-test, and at 1 and 3 minutes post-test. Peak lactate was determined during data analysis as the highest lactate value taken at either 1 or 3 minutes post-test.

Wingate Anaerobic Test

Peak and mean power, WAT-PP and WAT-MP respectively, were calculated during the WAT using the equation:

$$\text{Power (W)} = \frac{\text{Force (N)} \times \text{rpm} \times 6\text{m} \cdot \text{rev}^{-1}}{\text{Time (sec)}} \quad (\text{Adams \& Beam, 2008})$$

WAT data were also used to calculate:

- 1) mean absolute and relative power for the first 2 seconds (WAT-PP2)
- 2) mean absolute and relative power for the time equivalent to each participant completing the 8-flight TCT (WAT-MP_{TCT})

Blood lactate was measured pre-test (prior to the warm-up), and at 1 and 3 minutes post-test. Peak lactate was determined the same way as for the TCT.

Statistical Analysis

Statistical analysis for validity of the TCT was conducted using SPSS (version 19.0, SPSS Inc., Chicago, IL). Prior to analysis, all data were assessed for normality using the Shapiro-Wilk test ($p < 0.05$). Means and standard deviations (SD) are reported for all variables. T-tests were used to determine differences in means for TCT and WAT power, as well as lactate measurements. Pearson's product moment correlation, r , was used to identify relationships between TCT times of completion and power output from the TCT and WAT tests. Statistical significance was set at $p < 0.05$.

Statistical analysis for the reliability of the TCT was conducted using an Excel Reliability Spreadsheet developed by Hopkins (2000) (Microsoft Excel, 2008, Version 12.3.2). Reliability data were assessed for normality using the Shapiro-Wilk test ($p < 0.05$). Means and standard deviations (SD) are reported for all variables. Intraclass correlation (ICC), typical

error, and coefficient of variance (COV) were used to determine test-retest reliability (Hopkins, 2000). Statistical significance was set for all procedures at $p < 0.05$.

CHAPTER 3 – RESULTS

Participant Characteristics

Twenty-five male professional urban firefighters completed the TCT validation component of this study, while 21 active male (n=12) and female (n=9) participants completed the TCT test-retest reliability testing. Means and standard deviations for anthropometric characteristics of all participants in both components of this study are reported in Table 1. Participant weight for the TCT during both the validity and reliability testing includes the weight of the PPE and SCBA worn during sessions. All TCT data excluding the weight of the PPE and SCBA are included in Appendix D.

Table 1

Mean (SD) Anthropometric Data for Male Firefighters Completing the TCT Validity Study, and Participants Completing the TCT Test-Retest Reliability Study

	N	Age	Weight (kg)	VO ₂ peak
Validity Study	25	39.4 (7.81)	88.3 (10.7) in exercise clothing 110.9 (10.6) in PPE	4.31 (0.69) L·min ⁻¹ 49.0 (5.67) ml·kg ⁻¹ ·min ⁻¹ 40.66 (4.85) ml·kg ⁻¹ ·min ⁻¹ (when mass of PPE is considered)
Reliability Study	21	29.7 (9.6)	93.2 (12.4) in PPE Males (n = 12): 99.5 (12.2) Females (n = 9): 84.6 (6.5)	n/a

TCT Time, TCT Power, and WAT Power

Mean time and power data for the TCT and WAT during the validation component of this study are summarized in Table 2. Absolute and relative power data obtained during the WAT and TCT were significantly different ($p < 0.05$) with the exception of when TCT-P8 was compared to:

1) WAT-MP ($t = -0.972$, $p = 0.341$), and

2) WAT-MP_{TCT} ($t = 1.555$, $p = 0.133$).

Relative TCT-P1 was significantly greater than any power output during the WAT. Relative TCT-P8, on the other hand, was significantly less than any power output during the WAT.

Table 2

Mean (SD) Time and Power Values for Flight 1 and Flight 8 For the TCT and WAT During TCT Validation Testing

	N	Completion Time (sec)	Absolute Power (W)	Relative Power ($W \cdot kg^{-1}$)
TCT Flight 1	24 ^a	1.44 (0.33)	1384.72 (311.22) *	12.49 (2.59) ^
TCT Flight 8	25	21.16 (3.40)	730.21 (127.38)	6.59 (0.99) ^
WAT-PP2	25	n/a	828.34 (161.35) **	9.40 (1.53)
WAT-PP	25	n/a	980.17 (167.14) **	11.13 (1.55)
WAT-MP	25	n/a	710.45 (115.44)	8.06 (1.03)
WAT-MP _{TCT}	25	n/a	750.05 (121.06)	8.52 (1.10)

^a Due to equipment malfunction, one TCT Flight 1 data set was excluded from analysis

* Significantly different from all other absolute power variables, $p < 0.05$

** Significantly different from absolute TCT Flight 8 power, $p < 0.05$

^ Significantly different from all WAT relative power variables, $p < 0.05$

Relationships Between TCT Times and WAT Power Output Data

Correlations between TCT times and WAT power outputs are summarized in Table 3. TCT-T8 had a greater number of significant relationships with WAT power compared to TCT-T1. Apart from WAT-MP, all absolute WAT power data showed significant and moderately strong relationships with TCT-T8, and relative WAT power data showed significant and strong relationships.

Table 3

Pearson Correlation Coefficients for TCT Times and WAT Power During TCT Validation Testing

	TCT-T1	TCT-T8
WAT-PP2 (W)	-0.526 **	-0.587 **
WAT-PP2 (W·kg ⁻¹)	-0.682 **	-0.826 **
WAT-PP (W)	-0.456 *	-0.545 **
WAT-PP (W·kg ⁻¹)	-0.629 **	-0.827 **
WAT-MP (W)	-0.096	-0.285
WAT-MP (W·kg ⁻¹)	-0.195	-0.544 **
WAT-MP _{TCT} (W)	-0.374	-0.559 **
WAT-MP _{TCT} (W·kg ⁻¹)	-0.534 **	-0.860 **

* p < 0.05; ** p < 0.01

Relationships Between TCT Power and WAT Power

Several statistically significant relationships were found between TCT power and WAT power. As demonstrated in Table 4, TCT-P8 was

significantly correlated with all WAT-related variables, and had the highest *r*-values with the WAT compared to the other TCT variables.

Table 4

Pearson Correlation Coefficients for TCT and WAT Power During TCT

Validation Testing

	TCT-P1 (W)	TCT-P1 (W·kg ⁻¹)	TCT-P8 (W)	TCT-P8 (W·kg ⁻¹)
WAT-PP2 (W)	0.752 **	0.541 **	0.868 **	0.642 **
WAT-PP2 (W·kg ⁻¹)	0.644 **	0.688 **	0.754 **	0.851 **
WAT-PP (W)	0.703 **	0.461 *	0.864 **	0.600 **
WAT-PP (W·kg ⁻¹)	0.563 **	0.620 **	0.728 **	0.845 **
WAT-MP (W)	0.335	0.034	0.654 **	0.326
WAT-MP (W·kg ⁻¹)	0.090	0.097	0.468 *	0.541 **
WAT-MP _{TCT} (W)	0.635 **	0.383	0.869 **	0.596 **
WAT-MP _{TCT} (W·kg ⁻¹)	0.486 *	0.527 **	0.725 **	0.854 **

* *p* < 0.05; ** *p* < 0.01

TCT and WAT Blood Lactate

Mean TCT and WAT pre-test, post-test, and peak blood lactate values are shown in Table 5. While pre-test values were consistent across both exercise protocols, the WAT elicited significantly higher post-test blood lactate values compared to the TCT.

Table 5

Mean (SD) Pre-test, 1 and 3 Minute Post-test, and Peak Blood Lactate

Values for the WAT and TCT

	WAT	TCT
Pre-test	1.68 (0.51)	1.78 (0.85)
Post-1	11.31 (2.93)	9.96 (2.79) *
Post-3	12.83 (2.54)	10.28 (2.66) **
Peak	13.08 (2.38)	11.02 (2.60) **

* Significantly different from corresponding WAT lactate value, $p < 0.05$

** Significantly different from corresponding WAT lactate value, $p < 0.01$

TCT Test-Retest Reliability

Mean time and power data for the TCT reliability testing sessions are summarized in Table 6. Figure 2 contains individual performance data, intraclass correlations (R), typical error (TE), and coefficients of variance (COV) for test-retest reliability.

Table 6

Mean (SD) Time and Power Output for TCT Reliability Tests (n = 21)

	Trial 1	Trial 2
TCT-T1 (sec)	1.56 (0.49)	1.53 (0.50)
TCT-P1 (W)	1138.28 (405.60)	1161.26 (445.04)
TCT-P1 ($W \cdot kg^{-1}$)	11.99 (3.30)	12.22 (3.57)
TCT-T8 (sec)	21.81 (5.03)	21.38 (4.86)
TCT-P8 (W)	617.56 (178.35)	626.10 (175.84)
TCT-P8 ($W \cdot kg^{-1}$)	6.53 (1.32)	6.65 (1.31)

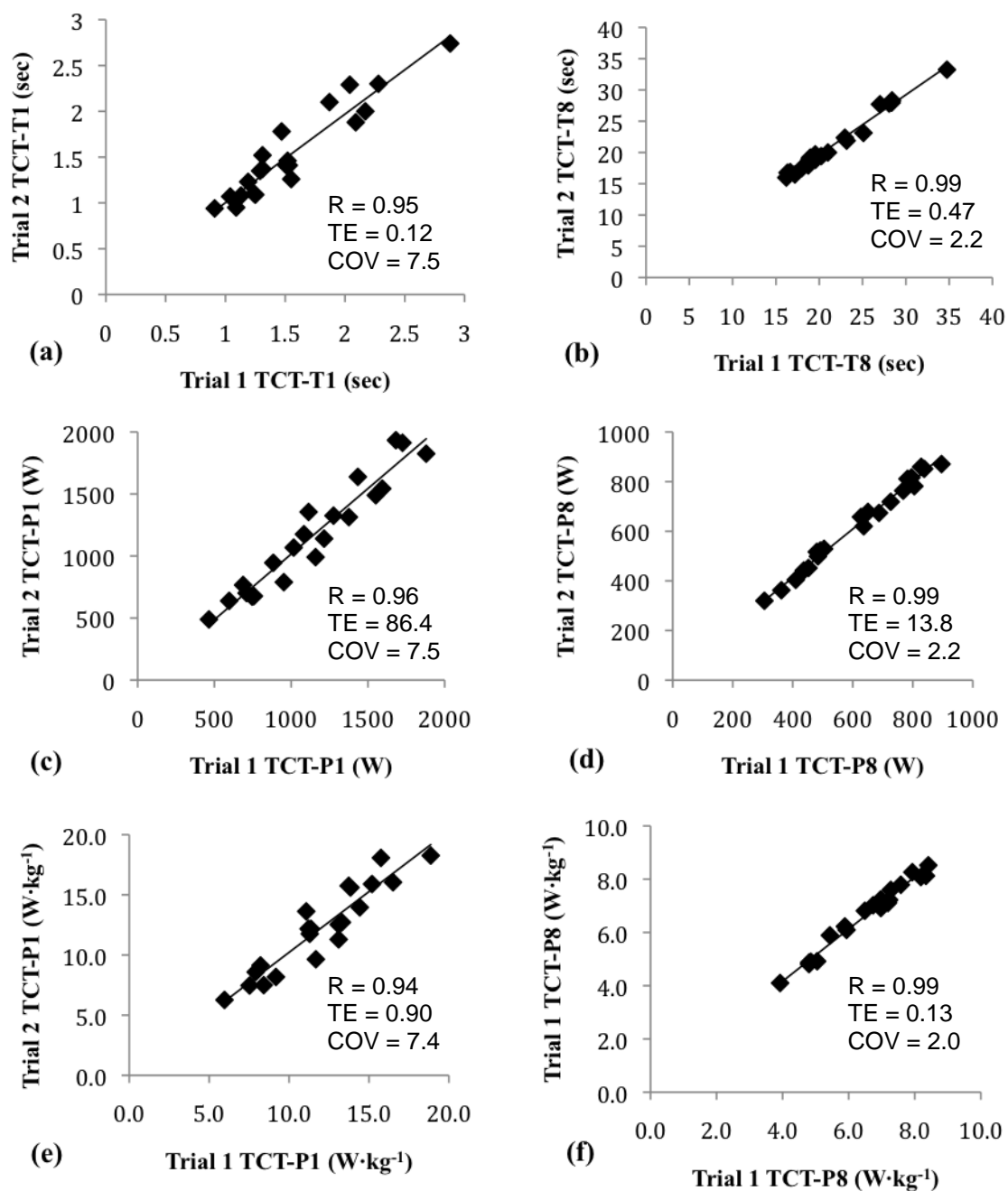


Figure 2. Reliability of (a) TCT-T1, (b) TCT-T8, (c) absolute TCT-P1, (d) absolute TCT-P8, (e) relative TCT-P1, and (f) relative TCT-P8 during 2 trials of the TCT ($n = 21$). Intraclass correlations (ICC), typical error (TE), and coefficients of variance (COV) are shown for each test-retest variable.

CHAPTER 4 – DISCUSSION

The purpose of this study was to evaluate the validity and reliability of a novel Tower Climb Test (TCT) as an appropriate field-based measurement tool for anaerobic performance in urban firefighters. This included examining the relationships that exist between the commonly used Wingate Anaerobic Test (WAT) and TCT performance. The WAT has been used as a valid and reliable measure of anaerobic power in various populations, and was employed in this study as the gold standard against which TCT measures were compared. Several strong and significant correlations were found between the TCT and WAT, supporting the hypothesis that the TCT can be used as an effective field-based assessment of occupation-specific anaerobic power in urban firefighters.

Participant Characteristics

Firefighters that volunteered to participate in the validity testing for the present study were initially selected by the fire department for an internal training event. The participants represented a heterogeneous sample of all levels of fire service that ranged from new recruits still on probation to officers with more than 20 years experience. The VO_2 peak test was conducted on the firefighters involved in the present study for descriptive purposes regarding their level of aerobic fitness. It has been recommended that firefighters meet a minimum requirement of 40 to 45

ml·kg⁻¹·min⁻¹ for safe and effective job performance (Lemon & Hermiston, 1977; Gledhill & Jamnick, 1992). VO₂ peak data from the present study was similar to or slightly greater than that previously reported in firefighting research (Lemon & Hermiston, 1977; Williams-Bell et al., 2009; Sheaff et al., 2010), indicating that the firefighters recruited for this study met aerobic performance standards for firefighting, and were not recruited based on their anaerobic training or capabilities.

TCT Power, Test Time, and Distance

Field tests analyzing anaerobic performance over 1 to 10 seconds rely primarily on the ATP-CP system (Margaria et al., 1966). Anaerobic alactic power represents the amount of work performed in a given unit of time, and is related to the maximal rate at which an individual can produce and utilize the ATP-CP system during high-intensity, very short duration tasks lasting no more than 4 to 5 seconds (Margaria et al., 1966; Foss & Keteyian, 1998; MacDougall et al., 1991). Anaerobic alactic capacity, on the other hand, refers to the total amount of ATP produced from intramuscular high-energy phosphates, and can be measured as the total work output during maximal short-term exercise lasting approximately 6 to 10 seconds (McArdle, Katch, & Katch, 2006; MacDougall et al., 1991).

Previous staircase testing has been conducted primarily to determine anaerobic alactic power through the use of the Margaria (Margaria et al., 1966) and Margaria-Kalamen (Kalamen, 1968) tests.

Although power generated during the TCT over the first (TCT-P1) and eighth (TCT-P8) flights in the present study are difficult to compare to other data reporting stair climb power using different methodologies, comparisons between the present study and previous research are presented in Table 7. Given the anaerobic alactic nature of the Margaria and Margaria-Kalamen tests, only TCT-P1 was selected for comparison with these other stair climbing tests. The mean time to climb the first flight of the tower (TCT-T1) was 1.44 ± 0.33 seconds and, like the Margaria and Margaria-Kalamen protocols, when completed with maximal effort, can be considered a measure of anaerobic alactic power (MacDougall et al., 1991).

Table 7

Comparison of Absolute and Relative Anaerobic Alactic Power from Present and Previous Research

Population	Anaerobic Alactic Power (W)	Anaerobic Alactic Power ($W \cdot kg^{-1}$)	Staircase Procedure	Reference
Power athletes (n = 6)	1175	15.4	Margaria	Komi et al., 1977
Hockey players (n = 13)	1154	14.9	Margaria	Komi et al., 1977
Speed skaters (n = 6)	1145	15.0	Margaria	Komi et al., 1977
Active males (n = 23)	970	12.9	Margaria	Komi et al., 1977
Nonathletic males (n = 23)	1652	22.1	Margaria-Kalamen	Kalamen, 1968
Untrained men (n = 82)	1413		Margaria-Kalamen	Mayhew & Salm, 1990
Male firefighters (n = 24)	1385	12.5	Modified Stair Climb Test	Present study

When compared to other data in Table 7, anaerobic alactic power from the present study is greater than that reported for other athletic and nonathletic populations using the Margaria protocol, but is less than that reported for active and inactive men using the Margaria-Kalamen protocol. This coincides with previous research showing that the Margaria-Kalamen Test tends to elicit higher power output scores compared to the original Margaria Staircase Test due to an increased running distance at the start of the test (6 metres instead of 2 metres), and a greater vertical distance (1.05 metres versus 0.70 metres) from climbing 3 stairs at a time instead of 2 (Kalamen, 1968). In the present study, a shorter approach to the base of the tower was incorporated which could have led to the decreased power output compared to other Margaria-Kalamen data. A shorter running start may have resulted in a slower speed as participants approached the staircase, and in turn, lowered power generation due to a greater time requirement to move from the first to the second timing instruments.

The present study required participants to travel an initial vertical distance of 1.75 metres from the base of the tower to the top of the first flight flight of stairs. This is greater than the 0.70 metre and 1.05 metre distances traveled in the Margaria and Margaria-Kalamen tests, respectively. Although greater power output data has been reported due to the increased vertical distance in the Margaria-Kalamen Test compared to the original Margaria protocol, it is inappropriate to conclude that the increased vertical distance in the present study should have led to

increased power due to greater vertical distance. Kalamen (1968) developed the protocols for the Margaria-Kalamen Test after determining the optimal stepping stride and vertical height. It was concluded that after testing 2, 3, and 4 steps per stride, covering distances of 0.697, 1.05, and 1.394 metres, respectively, that the 3 step stride and 1.05 vertical distance led to the greatest power output values (Kalamen, 1968). Therefore, although there is a greater vertical height in the present study compared to the Margaria and Margaria-Kalamen tests, it appears that it is greater than the ideal height to generate the highest possible power output. Since one of the objectives of this study was the development of a valid and reliable staircase test that could be easily administered on-site at fire training facilities, the height utilized in this study was selected based on ease of set-up and stabilization of the timing instruments, and the structure of the available training tower, which is representative of those available at various professional and volunteer fire departments in British Columbia. In addition, the TCT training tower included a 180° change of direction immediately after the participants passed the timing instruments to continue ascending the tower to the top of the eighth flight. Use of a training tower structured in this fashion may have also decreased power output compared to a training tower requiring movement of a mass over the same vertical distance in a continuous plane. Other studies assessing anaerobic performance instructed participants to run beyond the timing apparatus to

prevent deceleration influencing the final phase of the test (Clemons & Harrison, 2008).

Data from Margaria et al. (1966) and Kalamen (1968) have led to the development of age and gender-based norms for anaerobic power using the Margaria-Kalamen Test (Foss & Keteyian, 1996). Although there are some challenges to comparing data from this study with other data using a different protocol, normative data are shown in Appendix E. Given that data from this study fall between values previously cited using both Margaria and Margaria-Kalamen protocols, the norms have been used for comparative purposes. Anaerobic alactic power production, reflected as TCT-P1 in the present study, categorizes the participants as having “good” power output based on age and gender.

Clemons and Harrison (2008) developed a new stair sprint test in which participants ascended a single flight of stairs, 2.04 metres high. There was a 1.87 metre approach, and times were measured from the time the foot hit the first stair through to the eleventh stair. Protocols used in this test are the most similar to those used in the present study to measure anaerobic alactic power. Clemons & Harrison (2008) did not report completion times, but developed stair sprinting norms for power generation from college-age males ($n = 135$). Using this normative data, the 25 participants in the present study fell within the 70th to 75th percentiles, although they are approximately 18 years older than the participants in the study conducted by Clemons and Harrison (2008).

The inclusion of 8 flights of stairs in the TCT provided a unique way of testing anaerobic lactic performance. There is very limited information available for longer duration staircase tests as most popularized staircase protocols are less than 1 flight in length. Eight flights were chosen for this study given the occupational demands of firefighting, the nature of many firefighting drills, and the structure of the training tower and small multi-story buildings that urban firefighters may encounter. The TCT could be performed at many facilities since most fire departments will have training towers spanning multiple flights. The mean time of completion for this task (TCT-T8) was 21.16 ± 3.40 seconds, indicating that the fully completed TCT performed at maximal intensity would challenge the anaerobic lactic system more than shorter duration tests lasting only a few seconds focusing on measurement of anaerobic alactic power. Although performance times related primarily to the anaerobic lactic system usually last 30 to 90 seconds (MacDougall et al. 1991), glycolytic involvement and increased blood lactate levels have been measured in activities, including during the WAT, from as early as 5 to 10 seconds (Vandewalle, Peres, & Monod, 1987; Inbar et al., 1996; Beneke, Pollmann, Bleif, Leithauser, & Hutler, 2002). Anaerobic alactic, lactic, and aerobic contributions during the WAT have been measured at 31.1%, 50.3%, and 18.6%, respectively (Beneke et al., 2002). Like the WAT, the 21.16-second duration of the TCT would not be long enough to fully deplete anaerobic energy stores and measure anaerobic lactic capacity (Vandewalle et al, 1987; Inbar et al.,

1996). Although too long in duration to exclusively measure anaerobic alactic power, when performed at a maximal intensity, the 8-flight TCT does fall within a range to measure anaerobic lactic power (MacDougall et al., 1991; Inbar et al., 1996).

As expected, absolute and relative TCT-P8 were significantly lower than TCT-P1. Other anaerobic power tests show similar decrements in power output as the duration of the test increases (Vandewalle et al., 1987). The WAT, for example, usually yields the highest power during the first 5 seconds of the test, with continual power decreases over each subsequent 5 second increment, and the lowest power output during the last 5 seconds of the test (Inbar et al., 1996). During maximal performance tests lasting 10 to 180 seconds, the accumulation of hydrogen ions can interfere with ATP resynthesis and troponin-calcium binding capacity resulting in increased fatigue and decreased ability to maintain power production (Fitts, 1994; Sahlin, 1992).

WAT Power

WAT peak power (WAT-PP) and mean power (WAT-MP) were higher in this study compared to that reported in other firefighting studies (Sheaff et al., 2010; Williams-Bell et al., 2009), however previous studies have used a resistance of $75 \text{ g}\cdot\text{kg}^{-1}$ body weight (BW) compared to the heavier resistance ($85 \text{ g}\cdot\text{kg}^{-1}$ BW) employed in the present study. A difference in loading parameters can influence the power output generated

during the WAT, and the $75 \text{ g}\cdot\text{kg}^{-1}$ BW force setting originally suggested for the WAT has been reportedly too low for most adult and trained participants (Inbar et al., 1996).

Participants in the present study were instructed to attend the WAT session nourished and well-hydrated, without having had participated in strenuous activity within 24 hours. Williams-Bell et al. (2009) had participants perform the WAT at the end of a battery of other performance tests within the same day. The cumulative fatigue attained from a running treadmill VO_2 max test, 5 predicted 1-RM muscle strength tests, and 2 muscle endurance tests prior to the WAT could have also led to the lower power outputs reported. Furthermore, the participants in the study conducted by Williams-Bell et al. (2009) were classified as active males, and were not specifically firefighters.

Maud and Shultz (1989) developed normative tables for the WAT based on a resistance of $75 \text{ g}\cdot\text{kg}^{-1}$ BW in participants 18 to 28 years of age (See Appendix E). The absolute and relative WAT-PP, as well as absolute WAT-MP scores of the firefighters in this study would rank them in the 95th percentile and “well above average”. Relative WAT-MP of firefighters in the present study is classified as “above average” in the 80th to 85th percentile. Although it is difficult to draw conclusions from normative data using a different force setting, the classifications suggest that participants in this study have greater absolute WAT-MP compared to other male participants, but this is reduced when data are corrected for BW.

Relationships Between WAT Power and TCT Time

Several statistically significant inverse correlations were found between WAT power and TCT time, especially when data were corrected for BW. This provides evidence validating the effectiveness of the TCT in measuring anaerobic performance in urban firefighters. Individuals with a higher WAT power output were able to complete the occupation-specific TCT more quickly.

Misner et al. (1988) did not find significant relationships between 3-flight stair climb time with absolute WAT-PP ($r = -0.19$) or absolute WAT-MP ($r = -0.10$) in female firefighters. Patton and Duggan (1987) found that 200m sprint time (28.76 ± 1.64 seconds) was significantly correlated with relative WAT-PP ($r = -0.540$, $p < 0.05$) and relative WAT-MP ($r = -0.819$, $p < 0.001$), but not absolute WAT-PP ($r = -0.118$, $p > 0.05$) or absolute WAT-MP ($r = -0.370$, $p > 0.05$) in 14 military personnel (Patton & Duggan, 1987). 50m sprint time (7.14 ± 0.27 seconds) was significantly correlated with both relative and absolute WAT-PP (Patton & Duggan, 1987). Aziz and Chuan (2004) found the exact same correlation ($r = 0.63$, $p < 0.01$) between relative WAT-PP and 40m sprint time as was found in the present study between relative WAT-PP and TCT-T1.

TCT-T1 and TCT-T8 were the direct measurements taken during the TCT. Power was calculated for the TCT, and based on the formula, reflects performance time. While power calculations were required and were useful for the purposes of validating the TCT against the power

measurement obtained during the WAT, time on a fire scene may be the best representation of firefighter performance as it represents efficiency for completing a given task. At heavy workloads, firefighters who generate less power would be required to work at a higher percentage of their maximum ability to meet the demands, which may lead to greater exhaustion during the task and could result in increased time to completion and decreased job performance. The significant relationships between TCT times and WAT power outputs indicate that TCT time is a valid measure of anaerobic performance and could be used in future investigations involving the assessment of urban firefighters.

Relationships Between Absolute WAT and TCT Power

There were significant differences between all absolute TCT and WAT power values, except between TCT-P8 and both WAT-MP and WAT-MP_{TCT}. Statistically significant differences could be influenced by the varying demands of the tests, including the weight-bearing nature of stair climbing versus the weight-supported status of cycling, as well as the use of full PPE during the TCT versus exercise clothing during the WAT.

WAT-PP data were lower in the present study compared to TCT-P1, even though they were greater than that cited in previous research measuring WAT performance in firefighters. WAT-PP occurred over the first 5 seconds of the test, and was determined by using the equation:

$$\text{Peak Power (W)} = \frac{\text{Force (N)} \times \# \text{ rev} \times 6\text{m} \cdot \text{rev}^{-1}}{5 \text{ sec}} \quad (\text{Adams \& Beam, 2008})$$

Resistance was automatically applied during the WAT in the present study when participants reached 120rpm. This may not have been the maximum speed for every participant, and, in turn, may have decreased peak and mean power during the WAT because the participants would not have been able to increase speed and meet their personal optimal maximal cycling speed once the resistance was applied (Inbar et al., 1996). Inability to increase pedaling speed would have resulted in a lower number of revolutions in a given timeframe, and consequently decreased power output. Furthermore, the 85 g·kg⁻¹ BW resistance setting for the WAT in the present study may have been insufficient for some participants and may have lead to a sub-optimal maximal power output during the WAT as previously described compared to power output recorded during the TCT. Additionally, peak power occurred over the first 5 seconds of the WAT, but was calculated for the TCT as being 1.44 ± 0.33 seconds at TCT-T1.

Despite methodological variations in determining timing of peak power and limitations in the ability to compare specific values, the WAT had moderate to strong correlations with the TCT. Moderate to strong correlations have previously been reported between the WAT and other staircase tests (See Appendix A). Similar correlations have also been reported between WAT-PP and other field tests, including 40m run speed ($r = 0.84$; Inbar et al., 1996) and vertical jump height ($r = 0.70$; Tharp, Newhouse, Uffelman, Thorland, & Johnson, 1985), as well as between

WAT-MP and other field tests including 300m run speed ($r = 0.85$; Inbar et al., 1996), 300m run time ($r = -0.88$; Inbar et al., 1996), and vertical jump height ($r = 0.74$; Tharp et al., 1985).

In the present study, TCT-P8 was the TCT performance variable that had the greatest number of strong and significant correlations with WAT measurements. The highest r -value was found between TCT-P8 and the mean power produced in the WAT at the time corresponding with the TCT completion time, $\text{WAT-MP}_{\text{TCT}}$ ($r = 0.869$). This finding was expected because WAT data were corrected to better reflect the completion time of the TCT. It has been shown that power output decreases rapidly during longer duration tests (Vandewalle et al, 1987; Inbar et al., 1996), and since TCT-T8 averaged approximately 21 seconds, the power measured during the WAT for each participant between TCT-T8 to 30 seconds may not have provided the most accurate analysis when comparing the WAT and TCT. Typically the lowest WAT power output occurs over the last 5 seconds of the test (Adams & Beam, 2008), and therefore, if the TCT did not elicit the same level of fatigue due to the nature or duration of the test, it was deemed appropriate to exclude this data from the WAT. This is corroborated by the fact that although statistical significance was shown, WAT-MP, which represents the mean power over the full 30-second duration of the WAT, had lower r -values with TCT-P8 compared to every other WAT power variable. If time plays a role in fatigue development during the TCT as it has been shown to do during the WAT and other

anaerobic performance tests, there needs to be some consistency for test duration.

Using data from the WAT performance beyond the timeframe equivalent to the TCT could reflect an individual's fatigability, and fatigue associated with WAT performance can be highly variable between individuals (Inbar et al., 1996). Individuals that are more aerobically trained may have less performance decrement over a longer duration test compared to individuals that are more anaerobically trained (Inbar et al., 1996). A high correlation has been shown between 20-second and 30-second WAT performance (Vandewalle et al., 1987), and a 15 to 20 second WAT protocol has been suggested as a better anaerobic test duration for ease of testing and to minimize the influence of increasing contributions of the aerobic system (Inbar et al., 1996). Nonetheless, the 30-second WAT remains the most commonly used anaerobic assessment tool, and, given the extensive body of research produced using this protocol, continues to remain the gold standard anaerobic test against which other tests are compared (Inbar et al., 1996). Importantly, a strong correlation was observed between $WAT-MP_{TCT}$ and $WAT-MP$ ($r = 0.835$), and supports the use of $WAT-MP_{TCT}$ data for analysis of the TCT results.

TCT-P8 was also significantly correlated with WAT-PP2 and WAT-PP. Given the alactic nature of climbing one flight of stairs, power measures in the first 2 and 5 seconds of the WAT were expected to have strong and significant relationships with TCT-P1 compared TCT-P8,

however the r-values between TCT-P1 with both WAT-PP2 and WAT-PP were lower than those with TCT-P8. Nonetheless, these significant relationships are still exceptionally strong, and are within range of previously reported research. Ayalon, Inbar, and Bar-Or (1974) reported a correlation of $r = 0.79$ between WAT-PP and power during the Margaria Staircase Test, and Patton and Duggan (1987) reported a moderate correlation of $r = 0.62$ using a modified Margaria protocol. Limited previous research exists comparing the WAT with stair climb tests lasting more than a few seconds or a single flight of stairs, so it is difficult to determine whether or not the higher r-value between WAT-PP with a longer stair climb test were unique to this study or if it should have been expected. It was also found in the present study that TCT-T8 and TCT-P8 data had higher intraclass correlations, with lower error and coefficients of variance than TCT-T1 and TCT-P1 data (See Test-Retest Reliability Section, p. 43), which may help to explain why r-values between the WAT and TCT-P1 were lower than those of TCT-P8.

Relationships Between Relative WAT and TCT Power

Significant and strong relationships also existed between WAT and TCT power when data were corrected for BW. Like absolute power, the highest r-values existed between relative TCT-P8 with relative WAT- MP_{TCT} , WAT-PP2, and WAT-PP. Patton and Duggan (1987) reported $r = 0.64$ and $r = 0.74$ when comparing relative power using a modified Margaria protocol

with WAT-PP and WAT-MP, respectively. The present study found a similar correlation and statistically significant relationship ($p = 0.001$) between relative TCT-P1 with relative WAT-PP ($r = 0.620$), but not between relative TCT-P1 with relative WAT-MP ($r = 0.097$). The lack of significant correlation between TCT-P1 and WAT-MP was expected and replicated what was observed in the absolute data. Based on the physiological demands of the activity, Flight 1 data would be more representative of WAT-PP rather than that of mean power over a 30-second cycling test.

Significant differences were found between all relative TCT and WAT power values. While initially unexpected, relative TCT-P8 was lower compared to all relative WAT power values. In comparison, absolute TCT-P8 was not significantly different than WAT-MP or $WAT-MP_{TCT}$. This might be explained by the weight of the gear and the weight-bearing nature of stair climbing versus cycling. The act of carrying body mass, as well as approximately 26kg of protective gear up 8 flights of stairs would have required greater muscle involvement and may have led to increased fatigue and decreased power production compared to the non-weight-bearing WAT. Furthermore, the greater decrement in relative versus absolute power may be reflective of the fact that even though the PPE weighed the same for each participant, it was an overload equating to a different percentage of their body mass. The WAT, on the other hand, used an overload determined as a percentage of the participant's body mass. Despite the differences in raw scores, the moderate to strong correlations

between the TCT and the WAT indicate that relative power during the TCT could also be a useful measure to report anaerobic performance in urban firefighters.

WAT and TCT Blood Lactate

Peak blood lactate values were measured during both the WAT and TCT and were within expected ranges of anaerobic testing. This provides further support for the use of the TCT as a valid field-based anaerobic performance test. Previous studies have reported blood lactate values up to $13 \text{ mmol}\cdot\text{L}^{-1}$ during strenuous firefighting activities (Gledhill & Jamnick, 1992; von Heimburg, Rasmussen, & Medbo, 2006), providing additional evidence of the TCT being an appropriate test procedure for firefighters. Although the blood lactate measurements were within an expected range, all TCT post-test values were significantly lower than those of the WAT. The higher lactate values observed following the WAT compared to following the TCT likely reflects the different time requirements for each test. The WAT was, on average, nearly 9 seconds longer than the TCT and at a maximal anaerobic intensity would be expected to produce more lactate simply given the increased duration (Powers & Howley, 2008; Inbar et al., 1996).

The different methodologies employed for the WAT and TCT to measure post-test lactate values could also account for differences in the means between the values reported in the 2 tests. Once participants had

completed their ascent of the training tower during the TCT, they immediately turned around and descended to the start line as quickly as possible. The 1-minute and 3-minute post-exercise blood lactate timeframe began once participants reached the base of the tower, and from there participants walked approximately 50m to the testing area for blood lactate analysis. After the WAT, participants continued cycling for 20 seconds against zero resistance before walking approximately 5m to the lactate testing area; the 1- and 3-minute post-exercise time frames commenced immediately upon completion of the 30-second WAT.

Variation in measurement protocols for each test could have influenced shuttling of lactate from muscle into the bloodstream, leading to faster clearance and potentially increased aerobic utilization of lactate metabolism. The full-body nature of stair climbing would have required greater stabilizing activity and muscle involvement compared to the WAT, resulting in enhanced muscle pump and increased blood flow that could have facilitated clearance and mobilization of lactate produced during the ascent of the TCT. During the descent, then, there may have been an increased opportunity to use lactate as a fuel source by other tissues through gluconeogenesis, which may have led to the lower blood lactate values observed after the TCT compared to the WAT. Although the comparison of lactate values between the WAT and TCT may have been interfered with by methodologies, the lactate values obtained during the

TCT provide evidence that the TCT can be used as an appropriate anaerobic test for performance in urban firefighters.

Reliability of the TCT

The reliability of the TCT was measured using a test-retest procedure in 21 active males and females. Intraclass correlations (ICC) for all TCT time and power data ranged from 0.94 to 0.99, and represent very high reliability (Hopkins, 2000). Data from the full 8-flight TCT had higher ICC, with lower error and COV compared to the 1-flight TCT data. In combination with the validity test findings, these results provide additional support of the 8-flight TCT being an appropriate, valid, and reliable test.

Test-retest reliability of the TCT was at least as strong as that previously reported for other anaerobic performance tests, including the Margaria and Margaria-Kalamen tests, and the WAT (Inbar et al., 1996; Sawka, Tahamont, Fitzgerald, Miles, & Knowlton, 1980; MacDougall et al., 1991; Patton et al., 1985; Sargeant, Hoinville, & Young, 1981; McCartney, Heigenhauser, Sargeant, & Jones, 1983; Coggan & Costill, 1984; Ayalon et al., 1973; Patton et al., 1985). The uniqueness of cycling, particularly in non-cycling populations, as well as the shortness of previous staircase tests could build error and decrease reliability compared to the longer duration TCT employed in this study.

Furthermore, participants in the present study were not specifically instructed to perform the TCT stepping 2 or 3 steps at a time as required in

the Margaria and Margaria-Kalamen protocols. This could have minimized any learning effect during the TCT because participants could have taken a stepping stride that they were comfortable with and was more representative of their natural gait pattern. The Margaria-Kalamen test has been criticized previously, for example, because the 3-step stride pattern was challenging for some participants (Clemons & Harrison, 2008). Task familiarity of stair climbing has been previously reported as one of the major benefits of staircase testing (Margaria et al., 1966; Kalamen, 1968).

Although test-retest trials for TCT reliability were not conducted in the same firefighters who participated in the validity component of this study, the stair climbing task in the TCT would arguably be more reproducible amongst a population of firefighters who would have already been more familiar with the training tower utilized and the weight and restrictiveness of the PPE. Even in the non-firefighting population used in the reliability component of this study, the mean TCT-T8 difference between Trial 1 and Trial 2 was only 0.42 seconds. Mean completion times for the 2 trials were 21.81 ± 5.03 seconds and 21.38 ± 4.86 seconds for Trials 1 and 2, during reliability testing, and are very comparable to the 21.16 ± 3.40 second times measured in the 25 firefighters tested during the validity testing session. The test-retest measures obtained for TCT demonstrate it to be a highly reliable and reproducible field test of anaerobic alactic and lactic performance.

Conclusion

This study provides strong evidence that the TCT is an effective field-based assessment of occupation-specific anaerobic power in urban firefighters. Benefits of the TCT include:

- 1) high ecological validity to urban firefighting;
- 2) performance can be expressed as time to complete the test or power produced;
- 3) familiarity of a commonly performed task resulting in minimal learning effect;
- 4) acceptance of test procedures by the firefighting population and increased participant recruitment potential for future researchers;
- 5) minimal equipment requirements and ease of set-up;
- 6) minimal experience required for testers; and
- 7) ability to measure anaerobic performance amongst firefighters on-site at various fire training facilities with minimal disruption.

The TCT is an occupation-specific test that can be easily conducted and can minimize challenges inherent to testing within this population. The ability to calculate power as a measure of anaerobic performance during the TCT enhances the applicability of this test in that it could be administered at various facilities, and not just the training tower that was utilized for this study. Use of the TCT in future investigations could contribute to the growing body of research supporting firefighter health and

safety, and the continued development of physical fitness recommendations and guidelines.

REFERENCES

- Adams, G.M., & Beam, W.C. (2008). *Exercise physiology: Laboratory manual (5th ed.)*. New York: McGraw-Hill.
- Adams, J., Roberts, J., Simms, K., Cheng, D., Hartman, J., & Bartlett, C. (2009). Measurement of functional capacity requirements to aid in development of an occupation-specific rehabilitation training program to help firefighters with cardiac disease safely return to work. *American Journal of Cardiology*, *103*, 762-765.
- Angerer, P., Kadlez-Gebhardt, S., Delius, M., Raluca, P., & Nowak, D. (2008). Comparison of cardiocirculatory and thermal strain of male firefighters during fire suppression to exercise stress test and aerobic exercise testing. *American Journal of Cardiology*, *102*, 1551-1556.
- Ayalon, A., Inbar, O., & Bar-Or, O. (1973). Relationships among measurement of explosive strength and anaerobic power. In R.C. Nelson & C.A. Morehouse (Eds.), *Biomechanics IV* (pp.572-577). Baltimore: University Park Press.
- Aziz, A. R., & Chuan, T. K. (2004). Correlation between tests of running repeated sprint ability and anaerobic capacity by Wingate cycling in multi-sprint sports athletes. *International Journal of Applied Sports Sciences*, *16*, 14-22.

- Baechle, T. R., & Earle, R. W. (Eds.). (2000). *Essentials of strength training and conditioning* (2nd ed.). Champaign, IL: Human Kinetics.
- Baker, J., Grice, J., Roby, L., & Matthews, C. (2000). Cardiorespiratory and thermoregulatory response of working in firefighter protective clothing in a temperate environment. *Ergonomics*, *43*, 1350-1358.
- Barr, D., Gregson, W., & Reilly, T. (2010). The thermal ergonomics of firefighting reviewed. *Applied Ergonomics*, *41*, 161-172.
- Baumgartner, T.A., & Jackson, A.S. (1991). *Measurement and Evaluation in Physical Education and Exercise Science* (4th ed.). Wm.C.Brown: Dubuque, IA.
- Beneke, R., Pollmann, C., Bleif, I., Leithaeuser, R. M., & Huetler, M. (2002). How anaerobic is the Wingate anaerobic test for humans? *European Journal of Applied Physiology*, *87*(4), 388-392.
- Bilzon, J.L., Scarpello, E.G., Smith, C.V., Ravenhill, N.A., & Rayson, M.P. (2001). Characterization of the metabolic demands of simulated shipboard Royal Navy firefighting tasks. *Ergonomics*, *44*, 766-780.
- Carey, M. G., & Thevenin, B. J. -. (2009). High-resolution 12-lead electrocardiograms of on-duty professional firefighters: A pilot feasibility study. *The Journal of Cardiovascular Nursing*, *24*(4), 261-267.
- Clemons, J., & Harrison, M. (2008). Validity and reliability of a new stair sprinting test of explosive power. *Journal of Strength & Conditioning Research*, *22*(5), 1578-1583.

- Coggan, A. R., & Costill, D. L. (1984). Biological and technological variability of three anaerobic ergometer tests. *International Journal of Sports Medicine*, 5(3), 142-145.
- Davis, P.O., & Dotson, C.O. (1978). Heart rate responses of firefighters. *Journal of Ambulatory Electrocardiology*, 1, 15-18.
- Davis, P.O., & Dotson, C.O. (1987). Physiological aspects of fire fighting. *Fire Technology*, 23(3), 280-291.
- Del Sal, M., Barbieri, E., Garbati, P., Sisti, D., Rocchi, M. B. L., & Stocchi, V. (2009). Physiologic responses of firefighter recruits during a supervised live-fire work performance test. *Journal of Strength and Conditioning Research*, 23, 2396-2404.
- Dotan, R., & Bar-Or, O. (1983). Load optimization for the Wingate anaerobic test. *European Journal of Applied Physiology*, 51, 409-417.
- Dreger, R., Jones, R., & Petersen, S. (2006). Effects of the self-contained breathing apparatus and fire protective clothing on maximal oxygen uptake. *Ergonomics*, 49, 911-920.
- Fahy, R.F. (2005). *US firefighter fatalities due to sudden cardiac death, 1995-2004*. National Fire Protection Association; Quincy, MA.
- Fahy, R., & Leblanc, P. (2002). On-duty deaths 2002. *National Fire Protection Association Journal*, 97, 56-63.

- Findley, B., Brown, L.E., & Whitehurst, M. (2002). Anaerobic power performance of incumbent female firefighters. *Journal of Strength & Conditioning Research*, 16(3), 474-476.
- Firefit Web site. (2010). Retrieved December 12, 2011, from <http://www.firefit.com>
- Fitts, R.H. (1994). Cellular mechanisms of muscle fatigue. *Journal of Applied Physiology*, 104, 551-558.
- Foss, M.L., & Keteyian, S.J. (1998). *Fox's physiological basis for exercise and sport* (6th ed.). New York: McGraw-Hill.
- Fox, E., Bowers, R., & Foss, M. (1993). *The physiological basis for exercise and sport* (5th ed.). Dubuque, IA: Wm. C. Brown.
- Gledhill, N., & Jamnik, V.K. (1992). Characterization of the physical demands of firefighting. *Canadian Journal of Sports Science*, 17, 207-213.
- Hall, J.R. (2005). *Fire in the US and Canada*. National Fire Protection Association; Quincy, MA.
- Hart, G. (1999). *Job demands analysis – City of Vancouver Fire Department*. Human Effort Evolutionary Health; Vancouver, BC.
- Holmer, I., & Gavhed, D. (2007). Classification of metabolic and respiratory demands in firefighting activity with extreme workloads. *Applied Ergonomics*, 45-52.
- Hopkins, W.G. (2000). *Summarizing data: Precision of measurement*. Retrieved December 27, 2011, from <http://sportsci.org>

- Howell, D.C. (2008). *Fundamental statistics for the behavioral sciences* (6th ed.). Belmont, CA: Thomson Wadsworth.
- Inbar, O., & Bar-Or, O. (1986). Anaerobic characteristics in male children and adolescents. *Medicine and Science in Sports and Exercise*, 18, 264-269.
- Inbar, O., Bar-Or, O., & Skinner, J.S. (1996). *The Wingate anaerobic test*. Champaign, IL: Human Kinetics.
- Jackson, A. S. (1994). Pre-employment physical evaluation. *Exercise and Sport Science Reviews*, 22, 53-90.
- Kalamen, J.L. (1968). *Measurement of maximum anaerobic power in man*. Unpublished PhD Thesis, University of Ohio.
- Kales, S.N., Soteriades, E.S., Christophi, C.A., & Christiani, D.C. (2007). Emergency duties and deaths from heart disease among firefighters in the United States. *New England Journal of Medicine*, 356, 1207-1215.
- Kamon, E., & Belding, H.S. (1971). Heart rate and rectal temperature relationships during work in hot humid environments. *Journal of Applied Physiology*, 31, 472-477.
- Karter, M.J., & Molis, J.L. (2011). *US firefighter injuries – 2010*. National Fire Protection Association; Quincy, MA.
- Komi, P. V., Rusko, H., Vos, J., & Vihko, V. (1977). Anaerobic performance capacity in athletes. *Acta Physiologica Scandinavica*, 100(1), 107-114.

- Lemon, P.W., & Hermiston, R.T. (1977). The human energy cost of firefighting. *Journal of Occupational Medicine*, 23, 558-562.
- MacDougall, J.D., Wenger, H.A., & Green, H.J. (Eds.). (1991). *Physiological testing of the high performance athlete* (2nd ed.). Champaign, IL: Human Kinetics.
- Margaria, R., Aghemo, P., & Rovelli, E. (1966). Measurement of muscular power (anaerobic) in man. *Journal of Applied Physiology*, 21, 1662-1664.
- Maud, P.J., & Shultz, B.B. (1986). Gender comparisons in anaerobic power and anaerobic capacity tests. *British Journal of Sports Medicine*, 20(2), 51-54.
- Mayhew, J.L., & Salm, P.C. (1990). Gender differences in anaerobic power tests. *European Journal of Applied Physiology*, 60, 133-138.
- McArdle, W.D., Katch, F.I., & Katch, V.L. (2006). *Essentials of exercise physiology* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- McCartney, N., Heigenhauser, G. J., Sargeant, A. J., & Jones, N. L. (1983). A constant-velocity cycle ergometer for the study of dynamic muscle function. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, 55, 212-217.
- Misner, J.E., Boileau, R.A., Plowman, S.A., Elmore, B.G., Gates, M.A., Gilbert, J.A., & Horswill, C. (1988). Leg power characteristics of female firefighter applicants. *Journal of Occupational Medicine*, 30(5), 433-437.

- Moran, D.S., Montain, S.J., & Pandolf, K.B. (1998). Evaluation of different levels of hydration using a new physiological strain index. *American Journal of Physiology-Regulatory, Integrative, and Comparative Physiology*, *44*(3), R854-R860.
- Nelson, M. D., Haykowsky, M. J., Mayne, J. R., Jones, R. L., & Petersen, S. R. (2009). Effects of self-contained breathing apparatus on ventricular function during strenuous exercise. *Journal of Applied Physiology*, *106*, 395-402.
- Patton, J. F., & Duggan, A. (1987). An evaluation of tests of anaerobic power. *Aviation, Space & Environmental Medicine*, *58*(3), 237-242.
- Patton, J. F., Murphy, M. M., & Frederick, F. A. (1985). Maximal power outputs during the Wingate anaerobic test. *International Journal of Sports Medicine*, *6*(2), 82-85.
- Powers, S.K., & Howley, E.T. (2009). *Exercise physiology: Theory and application to fitness and performance* (7th ed.). New York: McGraw-Hill.
- Romet, T. T., & Frim, J. (1987). Physiological responses to firefighting activities. *European Journal of Applied Physiology*, *56*, 633-638.
- Rossi, R. (2003). Firefighting and its influence on the body. *Ergonomics*, *46*, 1017-1033.
- Rubin, S. (1987). Core temperature regulation of heart rate during exercise in humans. *Journal of Applied Physiology*, *62*, 1997-2002.
- Sahlin, K. (1992). Metabolic factors in fatigue. *Sports Medicine*, *13*, 99-107.

- Sargeant, A. J., Hoinville, E., & Young, A. (1981). Maximum leg force and power output during short-term dynamic exercise. *Journal of Applied Physiology*, 51(5), 1175-1182.
- Sawka, M. N., Tahamont, M. V., Fitzgerald, P. I., Miles, D. S., & Knowlton, P. I. (1980). Alactic capacity and power: Reliability and interpretation. *European Journal of Applied Physiology & Occupational Physiology*, 45(2), 109-116.
- Sheaff, A.K., Bennett, A., Hanson, E.D., Kim, Y., Hsu, J., Shim, J.K., Edwards, S.T., & Hurley, B.F. (2010). Physiological determinants of the Candidate Physical Ability Test in firefighters. *Journal of Strength and Conditioning Research*, 24(11), 3112-3122.
- Smith, D.J., & Stokes, S.M. (1985). Load optimization in anaerobic power testing of elite athletes. *Canadian Journal of Applied Sport Sciences*, 10, 30.
- Smith, D. L., Manning, T. S., & Petruzzello, S. J. (2001). Effect of strenuous live-fire drills on cardiovascular and psychological responses of recruit firefighters. *Ergonomics*, 44, 244-254.
- Smith, D. L., & Petruzzello, S. J., Kramer, J.M., & Misner, J.E. (1997). The effects of different thermal environments on the physiological and psychological responses of firefighters to a training drill. *Ergonomics*, 40, 500-510.

- Stair climb for clean air. (n.d.). Retrieved December 12, 2011, from [http://www.bc.lung.ca/campaigns_and_events/stair climb_for_cleanair.html](http://www.bc.lung.ca/campaigns_and_events/stair_climb_for_cleanair.html)
- Tharp, G.D., Newhouse, R.K., Uffelman, L., Thorland, W.G., & Johnson, G.O. (1985). Comparison of sprint and run times with performance on the Wingate anaerobic test. *Research Quarterly for Exercise and Sport*, 56, 73-76.
- TriData Corporation. (2002). *Firefighter fatality retrospective study, April 2002*. Prepared for Federal Emergency Management Agency, United States Fire Service, National Fire Data Center. Arlington, VA.
- Vandewalle, H., Peres, G., & Monod, H. (1987). Standard anaerobic exercise tests. *Sports Medicine*, 4(4), 268-289.
- von Heimburg, E.D., Rasmussen, A.K., & Medbo, J.I. (2006). Physiological responses of firefighters and performance predictors during a simulated rescue of hospital patients. *Ergonomics*, 49, 111–126.
- Washburn, A.E., LeBlanc, P.R., & Fahy, R.F. (1998). 1997 firefighter fatalities. *National Fire Protection Association Journal*. 92, 50-62.
- Wasserman, K. (1987). Determinants and detection of anaerobic threshold and consequences of exercise above it. *Circulation*, 76(6), 29-39.
- Williams, B.E., Peterson, S.R., & Douglas, K. (1996). Physiological responses of training officers during live fire training. *Fire Engineering*, 149(9), 45-52.

Williams-Bell, F.M., Villar, R., Sharratt, M.T., & Hughson, R.L. (2009).
Physiological demands of the firefighter Candidate Physical Ability
Test. *Medicine and Science in Sports and Exercise*, 41, 653–662.

APPENDIX A – REVIEW OF LITERATURE

The unique stressors of firefighting combined with a high prevalence of injury and on-duty death have garnered the attention of researchers and firefighting organizations, including the International Association of Firefighters (IAFF), the International Association of Fire Chiefs (IAFC), and the National Firefighter Protection Association (NFPA), for decades. While earlier research was largely conducted in laboratory settings, recent research includes more field-based studies trying to more closely replicate the specific stressors of firefighting. This chapter will summarize relevant research related to firefighter injuries and on-duty death, physiological demands of firefighting, and challenges imposed the use of personal protective equipment (PPE) and a self-contained breathing apparatus (SCBA). This chapter will also describe commonly used anaerobic performance laboratory and field tests, as well as considerations in occupational fitness testing to ensure validity and reliability in testing.

Injuries and On-Duty Death in Firefighting

Although data for Canadian firefighters is scarce, a report was developed for the NFPA comparing American and Canadian fire statistics from 2001 and found the following: the United States has approximately 18 times higher per capita reports of fire incidence; civilian fire death rates are higher in the US and Canada than any other industrialized countries in the world, although the deaths per million people have decreased by

approximately one-third in both countries between 1977 and 2001; the fire loss rate is similar, but slightly higher in Canada, when fire damage is expressed as a percentage of gross domestic product (Hall, 2005). In 2001, there were 337 and 1754 civilian deaths and injuries, respectively, and \$1.421 billion in property damage in Canada (Hall, 2005).

The NFPA has tracked and analyzed injuries and fatalities of on-duty volunteer and career firefighters in the United States since 1977 (Fahy, 2005). On average, excluding data resulting directly from terrorist attacks on September 11, 2001, approximately 80,000 firefighters are injured and 100 firefighters die annually while on-duty (Karter & Molis, 2011; Fahy, 2005; Fahy & LeBlanc, 2002; Washburn, LeBlanc, & Fahy, 1998; TriData Corporation, 2002). According to Karter & Molis (2011), approximately 45% of injuries to firefighters occur on the fireground. The leading cause of injury is overexertion and strain (25%), followed by falls, jumps and slips, contact with an object, exposure to fire products, and extreme weather (Karter & Molis, 2011). Similarly, up to 45% of fatalities in firefighters occur on the fireground due to sudden cardiac death (Fahy, 2005). According to Kales, Soteriades, Christophi, and Christiani (2007), approximately 32% of these deaths occur during fire suppression. Other causes of deaths occurring in the line of duty are due to responding to or returning from alarms involving emergencies, training exercises, non-fire emergency medical and rescue calls, and other non-emergency tasks, including fire prevention, inspection, and maintenance (Fahy, 2005; Kales et al., 2007).

It has been shown that firefighters spend only 1 to 5% of their time each year in fire suppression activities, yet those activities account for the majority of injuries and fatalities and pose a risk 10 to 100 times greater than non-emergency duties (Kales et al., 2007; Fahy, 2005). There are increased physical demands during fire suppression, but firefighters may have inadequate physical fitness, or other underlying cardiovascular risk factors (Kales et al., 2007). Occupational factors that have been attributed to the high prevalence of injury and death in firefighters include: emergency response, dangerous duties, chemical exposure, heat stress, shift work, high noise levels, psychological stress, and irregular and heavy physical exertion (Kales et al., 2007; Carey & Thevenin, 2009; Smith, Manning, & Petruzzello, 2001; Rossi, 2003). The analyses by Kales et al. (2007) and Fahy (2005) showed decreased risk of injury and sudden cardiac death when responding to less physically demanding emergencies and medical calls, and support the evidence that firefighters have a much greater incidence compared to other emergency medical workers who often deal with less physically demanding emergencies. Furthermore, these analyses show that physical training drills caused a relatively high incidence rate and, therefore, support the link of irregular and heavy physical exertion to injuries and fatalities amongst firefighters (Kales et al., 2007; Fahy, 2005).

Physiological Demands of Firefighting

Activities of firefighting include lifting and carrying heavy equipment up and down stairs, hoisting ladders, handling and maneuvering charged and uncharged hoses, breaking down doors and walls, fighting fires, auto extrication, property salvage, ceiling overhaul and working with heavy tools overhead, and search and rescue of victims (Hart, 1999; Barr, Gregson, & Reilly, 2010).

Laboratory and field tests documenting the physiological demands of these activities have shown that firefighters are required work at high intensities for sustained periods. Romet and Frim (1987) found that heart rates exceeded 150 beats per minute during firefighting, while Smith et al. (2001) found that firefighters reached age-predicted maximum heart rate (MHR), with mean values up to 189 beats per minute (bpm) during repeated bouts of firefighting drills. Bilzon, Scarpello, Smith, Ravenhill, and Rayson (2001) found that firefighting drills required work at 82% VO_2 max ($43 \text{ ml}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$), and 89 to 92% MHR. Gledhill and Jamnik (1992) found that VO_2 max reached $44 \text{ ml}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$ and HR reached 181 bpm during various firefighting tasks, and Davis and Dotson (1978) found that firefighters were required to work at 97% VO_2 max. Holmer and Gahved (2007) found that firefighters worked at 80% VO_2 max and 94% MHR during simulated non-fire work tasks, and reported that during a series of simulated firefighting activities, the highest VO_2 values occurred during tower climb activities, eliciting an oxygen uptake of $3.55\text{L}\cdot\text{min}^{-1}$. Lemon and

Hermiston (1977) found that firefighters sustained work at 60 to 80% VO_2 max even when not wearing an SCBA or exposed to the external stressors of an actual fire.

The PPE and SCBA that firefighters must wear weigh approximately 26.2kg combined, and have been shown to affect VO_2 max in laboratory settings compared to exercise in regular clothing (Dreger, Jones, & Peterson, 2006; Baker, Grice, Roby, & Matthews, 2000). Dreger et al. (2006) found VO_2 max decreased 17% with a corresponding decrease in ventilation and tidal volume while wearing PPE and a SCBA. Baker et al. (2000) found that VO_2 and HR were significantly higher when subjects were wearing PPE compared to sporting gear. Use of a SCBA during sustained activity increased core temperature, intrathoracic, and esophageal pressure, and reduced end systolic and diastolic cavity areas, stroke area, and plasma volume (Nelson, Haykowsky, Mayne, Jones, & Peterson, 2009).

The thermal stress firefighters face also increases the high physiological demands. Williams, Peterson, and Douglas (1996), and Angerer et al. (2007) measured HR values that met or exceeded age-predicted maximums, and estimated temperatures during live fire training of 500 to 800 degrees F. The effects of thermoregulation have been well documented and have shown increased HR, temperature, and blood lactate when exercising in a hot conditions compared to a thermoneutral conditions (Nelson et al., 2009; Smith et al., 2001; Smith, Petruzzello,

Kramer, & Misner, 1997; Del Sal et al., 2009). The thermal loads imposed on firefighters caused dehydration, and reduced stroke volume and plasma volume due to sweating (Smith et al., 2001; Moran, Montain, & Pandolf, 1998). Smith et al. (2001) found that stroke volume decreased up to 25% when comparing pre and posttest values. Reductions in stroke volume caused a compensatory cardiovascular drift response characterized by an increase in HR to maintain cardiac output and meet the demands of the activity (Rubin, 1987; Kamon & Belding, 1971).

The high physical exertion, use of PPE and SCBA, and the thermal loads inherent to firefighting will require increased reliance on anaerobic metabolism. While there has been considerable attention paid to the assessment of aerobic requirements related to firefighting, there has been less focus on the measurement of anaerobic contributions to occupation-specific metabolic demands. It has been reported that structural firefighting requires a strong anaerobic component ranging from 35% to 60% of the metabolic demands of the activity. Bilzon et al. (2001) found that firefighting tasks required 35% to 41% of the energy demands to be met anaerobically, while Davis and Dotson (1987) reported anaerobic contributions equivalent to approximately 40% (Lemon & Herminston, 1977; Davis & Dotson, 1987). During various firefighting activities, firefighters were required to work at up to 97% VO_2 max (Davis & Dotson, 1978), and blood lactate values were reported to be as high as $13 \text{ mmol}\cdot\text{L}^{-1}$ (Gledhill & Jamnick, 1992).

Firefighters with lower levels of aerobic fitness will be required to rely more on anaerobic metabolism to meet the occupational demands they are faced with relative to more aerobically fit counterparts. Lemon and Herminston (1977) reported 59.84% anaerobic contribution during various firefighting tasks in participants with lower aerobic power ($\text{VO}_2 \text{ max}$ less than $40 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$), compared to 52.73% anaerobic contribution in participants with a $\text{VO}_2 \text{ max}$ greater than $40 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$. Bilzon et al (2001) found that individuals with a $\text{VO}_2 \text{ max}$ greater than $43 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ were able to work more aerobically, and although females with a lower $\text{VO}_2 \text{ max}$ were able to complete the same amount of work as their male counterparts, they did so by working at a higher percentage of $\text{VO}_2 \text{ max}$ and producing more energy anaerobically. The intermittent nature of firefighting and the reliance on anaerobic metabolism supports the inclusion of anaerobic performance testing in firefighters.

Employment Testing, Analysis, Validity, and Reliability

The use of valid and reliable field tests in physically demanding occupations can assist in matching the demands of an occupation with individuals that are most physiologically capable of handling the workload and can enhance education, awareness, and training amongst workers (Jackson, 1994). Validity depends on reliability, and is arguably the most important characteristic in measurement and evaluation (Jackson, 1994). Validity refers to the degree to which a test measures what reports to

measure, and can be determined by comparing two measures on one group of participants (Hopkins, 2000).

Content validity relies on subjective decision-making to consider whether or not a certain test is appropriate to represent the given intention (Baumgartner & Jackson, 1991). Content validity has been used successfully with fitness testing and assessment, but has been less successful with more complex tests to assess sporting skills and has received some criticism given its subjective nature (Baumgartner & Jackson, 1991). Content validity has been used successfully in occupational settings to develop work-sample tests specific to firefighting, including climbing a ladder or stairs, dragging a dummy, and pushing, pulling, and carrying equipment (Jackson, 1994).

Criterion-related, or ecological, validity depicts how closely a test is predictive of job performance by analyzing data using a Pearson product-moment correlation coefficient (r) to compare field test scores to a criterion measure (Baumgartner & Jackson, 1991). R-values range from zero to 1, where an r -value approaching 1 indicates a significant relationship and that concurrent validity exists between then 2 tests, and an r -value approaching zero indicates no significant relationship between the tests and therefore little or no validity exists (Howell, 2008; Baumgartner & Jackson, 1991). Graphical representations of validity data should be represented so that the test scores that are being validated are assigned as the observed, independent variable on the X axis, while the criterion data is assigned as

the dependant variable on the Y axis (Hopkins, 2000). In other words, the observed value is used to predict the true value (Hopkins, 2000).

In order for a test to be valid, it must also be reliable. Reliability measures the reproducibility of results by testing the same participants with the same methods two or more times (Hopkins, 2000). Primary measures of test-retest reliability are intraclass correlations and typical error (Hopkins, 2000). Typical error is also referred to as within-subject standard deviation or standard error of measurement, and can be expressed as a raw value, or as the coefficient of variation (COV) (Hopkins, 2000). COV represents the typical error expressed as a percentage of the mean score, and usually ranges from 1% to 5% when used for athletic performance tests (Hopkins, 2000). The validity of test-retest reporting should be interpreted cautiously due to the influence of sample size and inter-individual variability using intraclass correlations, especially because most researchers usually only report correlations and do not report typical error and COV (Vandewalle, Peres & Monod, 1987).

Anaerobic Performance Testing

Valid and reliable laboratory and field tests are used by researchers assess anaerobic power and capacity. Maximal tests analyzing anaerobic performance over 1 to 10 seconds rely primarily on the ATP-CP system (Margaria, Aghemo, & Rovelli, 1966). Anaerobic power represents the amount of work done in a given unit of time ($P = \text{force} \times \text{distance}/\text{time}$), and

is related to the maximal rate at which an individual can produce and utilize the ATP-CP system during high-intensity, short duration tasks (Foss & Keteyian, 1998; MacDougall, Wenger, & Green, 1991). Anaerobic capacity, on the other hand, refers to the total amount of ATP produced through both alactic and lactic pathways (MacDougall et al., 1991). Anaerobic performance capacity can be measured as the total work output during maximal exercise in the short term (up to approximately 10 seconds), intermediate term (approximately 30 seconds), and long term (approximately 90 seconds) (MacDougall et al., 1991).

The Wingate Anaerobic Test (WAT) is a 30-second maximal effort cycle test and is the most widely accepted test of anaerobic leg performance (MacDougall et al., 1991). The WAT was developed in the 1970's, and provides measurements of peak power (PP), mean (average) power (MP), and fatigue index (FI) to assess anaerobic power and capacity (Inbar, Bar-Or, & Skinner, 1996). Energy production during the WAT is primarily anaerobic lactic (approximately 70%), with anaerobic alactic (approximately 15%), and aerobic (approximately 15%) contributions (MacDougall et al., 1991).

During the WAT, participants cycle on an ergometer at a maximal speed for 30 seconds against a predetermined resistance. Originally, the test was designed for adolescents and resistance was set at $7.5 \text{ g}\cdot\text{kg}^{-1}$ body weight (BW) (Inbar et al., 1996). The WAT has been adapted to utilize resistance loads of up to $10 \text{ g}\cdot\text{kg}^{-1}$ BW to suit various populations, including

active participants, and anaerobically trained athletes (Adams & Beam, 2008; Inbar et al., 1996). According to Dotan and Bar-Or (1983), the optimal load for eliciting the highest MP in male physical education students was $8.7 \text{ g}\cdot\text{kg}^{-1} \text{ BW}$ and higher in trained athletes. An inverted-U relationship has been suggested to exist between resistance and MP, but that PP appears to occur at higher loads (Dotan & Bar-Or, 1983).

PP is considered the highest power output occurring over 5 seconds of the test, and represents the maximum rate at which the ATP-CP system is capable of generating ATP (Inbar et al., 1996). PP usually occurs over the first 5 seconds of the test. MP is the average power sustained over the entire 30-second bout (Inbar et al., 1996). Power output data are often also expressed as relative power by dividing the absolute power by body weight (kg). Smith and Stokes (1985) measured PP outputs of $16.2 \text{ W}\cdot\text{kg}^{-1}$ in sprint speed skaters, and $13.5 \text{ W}\cdot\text{kg}^{-1}$ in volleyball players, $11.8 \text{ W}\cdot\text{kg}^{-1}$ in football players. Nonathletic males (aged 35-44) elicited $8.6 \text{ W}\cdot\text{kg}^{-1}$ (Inbar & Bar-Or, 1986).

The FI is the difference between PP and the lowest 5-second power divided by the peak power, and then expressed as a percentage (Inbar et al., 1996; Adams & Beam, 2008; MacDougall et al, 1991). The FI represents the maximal capacity of the ATP-CP and glycolytic systems to produce ATP and indicates that individuals with a higher fatigue index are less able to sustain the workload and maintain their power output over the

30-second test duration due to neuromuscular fatigue (Inbar et al., 1996; Adams & Beam, 2008).

The WAT has been used to determine the physiological demands of the firefighter Candidate Physical Abilities Test (CPAT), a content-based, 8-event circuit of simulated firefighting tasks. Male firefighters who successfully passed the criterion for the CPAT produced a PP of 633W during the WAT compared to 413W in men who did not successfully complete the CPAT (Williams-Bell, Villar, Sharratt, & Hughson, 2009). Sheaff et al. (2010) found that 26 male firefighters produced a PP of $841 \pm 29\text{W}$, or $9.4 \pm 0.3\text{W}\cdot\text{kg}^{-1}$.

Anaerobic power has also been previously measured using the WAT with female incumbent firefighters. Findley et al. (2002) measured WAT-PP and WAT-MP at $451.6 \pm 69.6\text{W}$ and $314.1 \pm 32.7\text{W}$, respectively, amongst 10 female incumbent firefighters using a force of $9.0\text{ g}\cdot\text{kg}^{-1}\text{ BW}$. Misner et al. (1988) assessed 150 female firefighters using $8.6\text{ g}\cdot\text{kg}^{-1}\text{ BW}$ and measured $494.0 \pm 84.7\text{W}$ and $398.2 \pm 56.9\text{W}$ for WAT-PP and WAT-MP, respectively. Findings from these 2 studies are contradictory to previous suggestions that a higher force should elicit a higher WAT-PP, however, power output was not significantly different between the 2 studies and might be explained by a large difference in sample size. Dotan and Bar-Or (1983) also suggested that the optimal load for yielding the highest mean power in adult female physical education students was $85\text{ g}\cdot\text{kg}^{-1}$, and therefore the $90\text{ g}\cdot\text{kg}^{-1}$ load might have had negative consequences on

power output given the inverted-U relationship between force and MP. As expected, the female firefighters had lower power output compared to that reported from male incumbent firefighters. It has been suggested that women typically do not tend to perform as well in anaerobic-based activities due to less efficient skeletal configuration, higher percent body fat, lower lean tissue, and lower peak lactic acid level in blood and muscle after maximal effort activity (Inbar et al., 1996). Findley et al. (2002) concluded that incumbent female firefighters should focus on the development of lower body power to meet occupational demands; however, these findings could also indicate that the WAT did not provide an accurate reflection of occupation-specific anaerobic performance for the female firefighters participating in the study.

Although highly popularized as a gold-standard measure for anaerobic power, use of the WAT may not provide the most appropriate or feasible assessment of anaerobic firefighting performance. According to MacDougall et al. (1991), test selection should be relevant and specific to the modes of activity and muscles typically used by a population. For firefighters, measurement of anaerobic performance during common firefighting tasks could be more useful than measurement on a cycle ergometer. Use of a cycling test limits the ability to consider the specific physical demands of firefighting, which mainly include weight-bearing locomotion, and eliminates the increased physiological strain imposed by

the weight, restrictiveness, and insulation of the PPE and increased work of breathing through an SCBA.

The Firefit Challenge is a popular fitness competition amongst the firefighting community that combines various fire-specific tasks. The events are done in sequence with no rest between tasks in the following order: stair climb, hose hoist, forcible entry, run, hose advance, and victim rescue (*Firefit*, 2010). The Firefit Challenge is highly publicized and has gained international attention in Canada, the US, New Zealand, and England (*Firefit*, 2010). The event, dubbed by participants as the “toughest two minutes in sports”, is done in full PPE and relies significantly on the anaerobic system to complete firefighting tasks. The current world record for completing the Firefit Challenge is 1 minute, 16 seconds (*Firefit*, 2010). The use of stair climbing specifically is also used for other firefighter public events including charitable fundraisers such as the Grind in the City Charitable Stair climb and the Stair climb for Clean Air (*Firefit*, 2010; “Stair climb for clean air”, 2011).

Researchers have used stair climbing to assess anaerobic performance in various populations. Advantages of the staircase tests include: inexpensive equipment requirements, minimal training and specialized skill requirements for testers and participants, applicability to a number of populations, and increased compliance from participants due to the simplistic nature of the test and minimal associated fatigue (Margarita et al., 1966).

The Margaria Staircase Test is a short-term anaerobic test in which participants ascend a flight of stairs, 2 at a time, as fast as possible (Margaria et al., 1966). A short, 2 metre “running start” is incorporated on a flat surface leading to the staircase to reduce the acceleration phase on the staircase and to reach a maximal climbing speed sooner in the test (Margaria et al., 1966). Timing lights are placed on the eighth and twelfth stairs, covering a vertical distance of 0.70m (Margaria et al., 1966). The Margaria Staircase Test has highly repeatable results as evidenced by scores within $\pm 4\%$ when testing occurred within the same session or on alternate days (Margaria et al., 1966).

An alternate stair climb test that has been utilized extensively to assess anaerobic performance is the Margaria-Kalamen Power Test. Kalamen (1968) modified the original Margaria test so that participants approach the staircase from 6m instead of 2m, and ascend 3 stairs at a time instead of 2. Switch mats are placed on the third and ninth stairs, and cover a vertical distance of 1.05m (Kalamen, 1968).

According to MacDougall et al. (1991), power output can be calculated in watts (W) for the Margaria and Margaria-Kalamen tests using the equation:

$$\text{Power (W)} = \frac{\text{Weight (kg)} \times 9.8 \times \text{Vertical Distance (m)}}{\text{Time (sec)}}$$

Original calculations by the researchers expressed power output values in $\text{kg}\cdot\text{m}\cdot\text{sec}^{-1}$, from the equation:

$$\text{Power} = \frac{\text{Weight (kg)} \times \text{Vertical Distance (m)}}{\text{Time (sec)}} \quad (\text{Foss \& Keteyian, 1998})$$

Data expressed in $\text{kg}\cdot\text{m}\cdot\text{sec}^{-1}$ can be converted to watts by multiplying by 9.807 (McArdle, Katch, & Katch, 2006; Baechle & Earle, 2000).

Komi, Rusko, Vos, and Vihko (1977) used the Margaria Staircase Test to determine anaerobic power in 89 athletes and 31 controls. Anaerobic power in male subjects ranged from 854W ($12.9 \text{ W}\cdot\text{kg}^{-1}$) for long distance runners, to 1175W ($15.4 \text{ W}\cdot\text{kg}^{-1}$) for athletes competing in power events (100-400m running, throwing and jumping sports, and decathlon). Other athletes eliciting the highest power outputs were hockey players (1154W , $14.9 \text{ W}\cdot\text{kg}^{-1}$), and speed skaters (1145W , $15.0 \text{ W}\cdot\text{kg}^{-1}$). Long distance runners (854W , $12.9 \text{ W}\cdot\text{kg}^{-1}$) and cross-country skiers (869W , $12.5 \text{ W}\cdot\text{kg}^{-1}$) had the lowest power outputs amongst all athletes.

The Margaria-Kalamen Test tends to elicit higher power output scores compared to the original Margaria Staircase Test. Kalamen (1968) measured 1652W in 23 nonathletic males, while Mayhew and Salm (1990) measured 1413W in 82 untrained men, and Maud and Shultz (1986) measured 1648W in 54 physically active male college students.

Data from Margaria et al. (1966) and Kalamen (1968) have led to the development of age and gender-based norms for anaerobic power using the Margaria-Kalamen Test (Foss & Keteyian, 1996). Full normative data is outlined in Appendix E, although it should be noted that these norms are based on combined data from Margaria et al. (1966) and Kalamen (1968), despite methodological variations and differences in observed power data.

Calculation of absolute power during staircase testing is mass-specific, so consideration should be made in the interpretation of data. The power that a heavier person generates during stair climbing is greater than a lighter counterpart, even if they score the same time and have covered the same distance (McArdle et al., 2006).

Staircase tests have previously been validated against the WAT. Misner et al. (1988) did not find significant relationships between a 3-flight stair climb time with absolute WAT-PP ($r = -0.19$) or absolute WAT-MP ($r = -0.10$) in female firefighters. Patton and Duggan (1987) found that 200m sprint time (28.76 ± 1.64 seconds) was significantly correlated with relative WAT-PP ($r = -0.540$, $p < 0.05$) and relative WAT-MP ($r = -0.819$, $p < 0.001$), but not absolute WAT-PP ($r = -0.118$, $p > 0.05$) or absolute WAT-MP ($r = -0.370$, $p > 0.05$) in 14 military personnel (Patton & Duggan, 1987). 50m sprint time (7.14 ± 0.27 seconds) was significantly correlated with both relative and absolute WAT-PP (Patton & Duggan, 1987). Aziz and Chuan (2004) reported $r = 0.63$ between relative WAT-PP and 40m sprint time.

ICC's have been reported between 0.85 to 0.90 for the Margaria Test (Inbar et al., 1996; Sawka, Tahamont, Fitzgerald, Miles, & Knowlton, 1980). WAT-PP and WAT-MP have been reported with an ICC ranging from 0.90 to 0.98 (MacDougall et al., 1991; Patton, Murphy, & Frederick, 1985; Ayalon et al., 1974), with a COV of 5.4% to 8.6% (Coggan & Costill, 1984; Sargeant, Hoinville, & Young, 1981; McCartney, Heigenhauser,

Sargeant, & Jones, 1983). Clemons and Harrison (2008) reported an ICC of 0.99 for a stair sprinting test up 11 stairs on a straight staircase.

Conclusion

Laboratory and field testing has been an essential part of firefighting preemployment standards and research. While there has been considerable attention paid to the assessment of aerobic requirements related to firefighting, there has been less focus on the measurement of anaerobic contributions to occupation-specific metabolic demands. Anaerobic performance testing has commonly been measured in various populations using Wingate and staircase testing. Testing in specific populations will be most applicable and relevant when it represents the actual demands of the activity. Development of an occupation-specific test that can easily be conducted at various fire departments will minimize challenges inherent to testing within this population, and will contribute to the growing body of research supporting firefighter health and safety, and the continued development of physical fitness recommendations and guidelines.

APPENDIX B – DATA COLLECTION SHEETS

TCT and WAT

Name: _____ **Group:** _____ **Age:** _____

Wingate Test

Date: _____ Testers: _____

Weight: _____ kg Basket Load: _____ kg

Pre-test Lactate: _____ (measured by: _____)

Peak Power: _____ W Mean Power: _____ W

Peak Power: _____ W/kg Mean Power: _____ W/kg

Post-test Lactate (measured by: _____):

1 min Post-test Lactate: _____ 3 min Post-test Lactate: _____

Comments:

Tower Climb Test

Date: _____ Testers: _____

Weight (no gear): _____ Weight (full gear): _____

Pre-test Lactate: _____ (measured by: _____)

Timing Light ID #: _____

Split 1: _____ **Split 2:** _____ **Split 5:** _____

Split 3: _____ Split 4: _____

Post-test Lactate (measured by: _____):

1 min Post-test Lactate: _____ 3 min Post-test Lactate: _____

Comments:

APPENDIX C – DATA COLLECTION PROTOCOLS

Tower Climb Test

Equipment:

- Training tower
- Timing lights (Brower Speedtrap Systems)
- Stopwatches
- Whistle
- Duct tape
- Measuring tape
- Cones
- Spare 9V batteries
- Scale
- Clipboard with data collection sheets
- Pens
- Blood Lactate Supplies: Test Strips, Analyzers, Lancets, Gauze, Sharps Container, Alcohol Swabs, Band Aids, Bleach solution (10%)

Set-up:

1. Place cones at the start of the concrete pad at the base of the training tower (1.7m from first stair).
2. Duct tape timing lights to the handrails at:
 - i. The base of the tower
 - ii. The top of the first flight of stairs (height = 1.75m)
 - iii. The top of the eighth flight of stairs (height = 13.89m)
3. Ensure receiver from Brower Speedtrap Systems is collecting data from all timing lights.

Procedures:

Familiarization and Practice:

1. Explain the protocol (described below) to the participant.
2. Explain the purpose of the blood lactate analyzers.

Data Collection

1. Obtain participant's baseline weight in station gear
2. Have participant don PPE pants and boots and researcher or assistant takes baseline blood lactate.
3. Have participant don the rest of PPE and reweigh them.
4. Have participant begin 2 minute warm-up on a stationary cycle.
5. Recap of testing procedure.
6. Participants start at the start line, and on their own time, sprints up the stairs as fast as possible using only leg power. Participants are not permitted to use their arms on the handrails during the ascent of the training tower.
7. When the participant crosses the first set of timing lights at the base of the tower, the researcher or research assistant at the base of the tower will blow a whistle so that the researcher or research assistant at the top of the tower can press start on a back-up stopwatch in case of equipment malfunction.
8. At the top of the tower, participants must touch the handrails at the end of the platform before turning around for descent.

9. During descent, participants must use the handrails for support and their feet are required to touch each stair.
10. When the participant returns to the base of the training tower, an assistant will begin a post-test stopwatch and escort them inside of the firehall for measurement of blood lactate 1 and 3 minutes post-test.
11. Record the split times from the Brower Timing Systems receiver.
12. Offer the participant water.
13. Researcher will ensure proper clean up and of all instruments used for collecting blood lactate and HR.

Wingate Anaerobic Test

Equipment:

- Monark cycle ergometer and weight plates
- Laptop and cables
- Scale
- Stopwatches
- Clipboard with data collection sheets
- Pens
- Blood Lactate Supplies: Test Strips, Analyzers, Lancets, Gauze, Sharps Container, Alcohol Swabs, Band Aids, Bleach solution (10%)

Set-up:

1. Adjust seat height of ergometer so that participant has a slight bend in knee with feet on pedals.
2. Adjust foot straps so feet are secure on pedals.
3. Ensure weight basket is in upright position.

Procedures:

Familiarization and Practice:

1. Explain the protocol (described below) to the participant.
2. Explain the purpose of the blood lactate analyzers.

Data Collection

1. Obtain participant's baseline weight in exercise gear and calculate resistance setting at $85 \text{ g}\cdot\text{kg}^{-1}$ body weight.
2. Have researcher or assistant takes baseline blood lactate.
3. Recap of testing procedure.
4. Have participant begin 2 minute warm-up at 60 to 80rpm at 0 resistance.
5. Cue participant to increase pedal frequency to as fast as possible. Ensure that resistance is applied automatically by the Monark software at 120 rpm.
6. Upon completion of the 30-second test, manually remove resistance and begin stopwatch.
7. After participant cycles against zero resistance for 20 seconds, have them return to blood lactate measurement area and record 1 and 3 minute post-test blood lactate.
8. Record time and power data from Monark software.
9. Offer the participant water.
10. Researcher will ensure proper clean up and of all instruments used for

collecting blood lactate and HR.

Maximal Graded Exercise Test on the Treadmill

Equipment:

- Treadmill
- Metabolic Measurement Cart
- Breathing Apparatus – fire helmet with modified head piece, Rudolph valve, hose, and nose clip
- Heart Rate monitors
- Towels
- Clipboard with data collection sheet
- Pens

Procedures:

Familiarization and Practice:

1. Explain the protocol (described below) to the participant.
2. Explain the purpose of the heart rate monitors and monitoring devices.
3. Have participants practice walking on treadmill if they are unaccustomed.

Data Collection:

1. Measure participant body weight and height.
2. Put heart rate monitor on participant.
3. Have participant don firefighter PPE and running shoes.
4. Reweigh participant in PPE.
5. Place nose clip and head piece on participant along with Rudolph valve in the mouth. Adjust accordingly.

6. Recap of testing procedure.
7. Have participant walk on treadmill for 3 minutes at 3.5mph, 0% grade.
8. Next 2 minutes increase grade to 2%.
9. Increase grade by 2% every 2 minutes until ventilatory threshold (VT) is reached ($RER > 1.0$).
10. At VT, increase grade by 2.0% each minute. If the participant reaches a maximum grade of 16%, increase speed by 0.5mph each minute until the criteria for VO_2 max is met which include:
 - i. Attainment of predicted maximum heart rate ($220 - \text{age}$);
 - ii. A rise in VO_2 of less than $2 \text{ ml } \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ with an increase in workload;
 - iii. A respiratory exchange ratio (RER) greater than 1.15;
 - iv. Volitional exhaustion.
11. Remove Scott harness and tank, jacket, helmet with head piece, and Rudolph valve from participant.
12. Ask for RPE rating.
13. Participant has cool down on treadmill at light walking speed and is given water.
14. Participant steps down from treadmill and continues to cool down and stretch.

Notes:

- At least one other research assistant will be present at test as a spotter.

- Researcher will download and print data from the metabolic cart.
- Rudolph valve and hose will be sanitized for each participant.

Blood Lactate Testing Instructions

Equipment:

- Gloves
- Lactate strips and calibration strip (Arkray Lactate Pro Test Strips)
- Alcohol swabs (Loris Medium)
- Disposable lancet tips (Accu-Chek Softclix Pro lancets)
- Biohazard Can
- Regular garbage can
- Hand sanitizer and antibacterial wipes
- Lactate analyzer (Arkray Lactate Pro)
- Gauze pads (Source 2" x 2" non-sterile)
- Lancet (Accu-Chek Softclix Pro)
- Sharps container (BD Sharps collector)
- Band-aids Beaker with dilute bleach (soaking lancets)

Equipment Calibration:

- Calibrate lactate analyzer using check strip & calibration strip and insert new test strip (see Lactate Pro instructions for more details)

Procedures:

Testing preparation:

1. Wash hands, put clean, new gloves on
2. Prep bleach solution - MUST PREP NEW BLEACH SOLUTION EACH DAY (bleach solution only effective for 1 day) - dilution 1/10 (bleach/water)
3. Disinfect the table and equipment where you will be taking blood - set out equipment and supplies you will be using for data collection

Data collection:

1. Load lancet with tip
2. Wipe participant's finger with alcohol swab – discard swab
3. Wipe excess alcohol from finger with gauze
4. Puncture fingertip with lancet
5. Gently squeeze finger to start flow of blood
6. Wipe first drop of blood with gauze
7. Gently squeeze finger to encourage flow of blood
8. Collect blood sample
9. Apply gauze and pressure to fingertip to stop blood flow, participant holds onto gauze until next blood sample - gauze is then disposed of in biohazard can
10. Dispose of used lancet tip in sharps container
11. After recording lactate value dispose of used lactate test strip in biohazard bag
12. Dispose of gauze wrappers and alcohol wrappers in the regular garbage can
13. For multiple samples:
 - i. Use fresh alcohol swab & gauze for each sample (do not usually need to lancet each time for the same participant)
 - ii. Check to see if blood is still flowing (may have to perform gentle squeeze)

- iii. Excessive squeezing will cause erroneous results
- iv. NOTE: watch for blood spray or splatter when squeezing
(point finger away from face)

14. Apply Band-aid at end of test
15. Discard gloves in appropriate garbage receptacle at the end of each test – use a fresh pair of gloves for each new participant
16. All equipment (lactate analyzer, lancet, bench top, etc) needs to be cleaned with alcohol or bleach between participants

Post-test:

1. Ensure all garbage has been disposed of appropriately (see above)
2. Bleach all surfaces that may have been in contact with blood or body fluids
3. Surfaces - spray with bleach , leave for 5 min, then wipe with cloth or paper towel
4. SPILLS – cover with paper towel, spray with bleach, leave for 5 min, then wipe with paper towel, re-spray with bleach & re-wipe
5. Wipe lancet with alcohol & soak tip in dilute bleach solution between tests & at end
6. Wipe analyzers, pens, keyboard, mouse, (anything you touch with gloves) with alcohol &/or bleach
7. Put all equipment and unused supplies away in appropriate storage places

If your participant becomes light headed, queasy or faints, get them to lie down immediately, raise feet slightly. Stay with the person even if they say they feel okay. Accompany them to the washroom, or out of the room. Watch for skin color, dilated pupils. If they did faint and have fallen, check for injuries. Get an accident report form and fill it out.

APPENDIX D – RESULTS FOR TCT EXCLUDING PPE MASS

Table 8

Mean (SD) Calculated Flight 1 and Flight 8 Power Output Excluding PPE

Mass During TCT Validation Testing

	N	Absolute Power without PPE Mass (W)
TCT-P1	24*	1109.73 (255.91)
TCT-P8	25	584.59 (104.43)

* Due to equipment malfunction, one TCT Flight 1 data set was excluded from analysis

Table 9

Pearson Correlation Coefficients for TCT and WAT Power Excluding PPE

Mass During TCT Validation Testing

	TCT-P1 Without PPE Mass (W)	TCT-P8 Without PPE Mass (W)
WAT-PP2 (W)	0.766 **	0.885 **
WAT-PP (W)	0.715 **	0.878 **
WAT-MP (W)	0.365	0.686 **
WAT-MPTCT (W)	0.654 *	0.892 **

* $p < 0.05$, ** $p < 0.01$

APPENDIX E – NORMATIVE DATA FOR OTHER ANAEROBIC PERFORMANCE TESTS

Table 10

Margaria-Kalamen Stair Sprint Test Absolute Power Guidelines for Males

Classification	20-30 years	30-40 years	40-50 years
Poor	< 1,039	< 834	< 637
Fair	1,039 – 1,363	834 – 1,089	637 – 824
Average	1,364 – 1,716	1,090 – 1,373	825 – 1,029
Good	1,717 – 2,059	1,374 – 1,648	1,030 – 1,226
Excellent	> 2,059	> 1,648	> 1,226

(Fox, Bowers, & Foss, 1993; Baechle & Earle, 2000)

Table 11

WAT Peak Power Percentile Norms for Males (n = 62)

Percentile Rank	Watts	W·kg ⁻¹
95	866.9	11.08
90	821.8	10.89
85	807.1	10.59
80	776.7	10.39
75	767.9	10.39
70	757.1	10.20
65	744.3	10.00
60	720.8	9.80
55	706.1	9.51
50	689.4	9.22
45	677.6	9.02
40	670.8	8.92
35	661.9	8.63
30	656.1	8.53
25	646.3	8.34
20	617.8	8.24
15	594.3	7.45
10	569.8	7.06
5	530.5	6.57

(Maud & Schultz, 1989)

Table 12

WAT Mean Power Percentile Norms for Males (n = 60)

Percentile Rank	Watts	W·kg ⁻¹
95	676.6	8.63
90	661.8	8.24
85	630.5	8.09
80	617.9	8.01
75	604.3	7.96
70	600.0	7.91
65	591.7	7.70
60	576.8	7.59
55	574.5	7.46
50	564.6	7.44
45	552.8	7.26
40	547.6	7.14
35	534.6	7.08
30	529.7	7.00
25	520.6	6.79
20	496.1	6.59
15	484.6	6.39
10	470.9	5.98
5	453.2	5.56