Journal of Occupational and Environmental Medicine, Publish Ahead of Print

DOI : 10.1097/JOM.0000000000001583

Manuscript title:

Validity and Reliability of Firefighting Simulation Test Performance

Authors:

Richard D M Stevenson (Ph.D.)[1,2], Andrew G Siddall (Ph.D.)[1,3], Philip J F Turner (M.Sc.)[1,4], James L J Bilzon (Ph.D.)[1]

[1]Department for Health, University of Bath, Bath UK

[2]Occupational Health Services, South Wales Fire & Rescue Service, Cardiff UK

[3]Occupational Performance Research Group, University of Chichester, Chichester UK

[4]Lancashire Fire & Rescue Service, Preston, Lancashire UK

Corresponding Author

Professor James Bilzon, Department for Health, University of Bath, Bath BA2 7AY

Email: J.Bilzon@bath.ac.uk. Telephone: 01225383174

Funding

This work was jointly funded by the Chief Fire Officer's Association, the FireFit Steering Group and the Fire Service Research and Training Trust.

Conflicts of interest

The authors express no conflict of interest

Acknowledgements

We are grateful to all the participating UK Fire and Rescues Services for use of their facilities and to the individuals that volunteered to act as participants for this study.

Abstract

Objective To assess the validity and reliability of a firefighting simulation test (FFST). **Methods** Sixty-nine operational firefighters completed a best-effort FFST on one occasion and twenty-two participants completed a further FFST. All participants completed a maximal treadmill test to determine cardiorespiratory fitness (VO_2max). **Results** Time to complete the FFST demonstrated a strong inverse relationship with VO_2max (r = - 0.73), although the prediction error was high. Reliability of the FFST was high (r = 0.84, p = 0.01), demonstrating a coefficient of variation of 4.5%. **Conclusions** The FFST demonstrated reasonable validity as a surrogate assessment of cardiorespiratory fitness for firefighting. The FFST also demonstrated good reliability. Given the apparent magnitude of the prediction error, the FFST would be best used as a training tool, rather than as a primary means of assessing cardiorespiratory fitness for firefighting.

Key Words: Firefighting, performance test, validity and reliability, physical fitness,

Introduction

Physically demanding public safety occupations (e.g. police, fire and ambulance) require employees to possess the physical and physiological attributes to perform their duties safely and effectively, preferably without experiencing excessive physical strain. Without a minimum level of physical fitness, particularly cardiorespiratory fitness, workers may be subjected to overexertion, which can increase the risk of injury to the employee and may also affect public safety (Baur, Christophi, Tsismenakis, Cook, & Kales, 2011; Baur, Leiba, Christophi, & Kales, 2012; Plat, Frings-Dresen, & Sluiter, 2012). With the improved understanding of the risks to employees in these physical occupations, in particular from excessive cardiovascular strain (Kales, Soteriades, Christophi, & Christiani, 2007; Maguire, Hunting, Smith, & Levick, 2002; Smith, Barr, & Kales, 2013), there is a recognition that employers have a duty of care to administer regular fitness assessments to ensure that their employees maintain appropriate minimum levels of physical fitness for safe and effective work (Adams, 2016; Health and Safety at Work Act, 1974).

Physical fitness tests in demanding occupations typically take the form of either criterion (i.e. job simulation) or surrogate (i.e. predictive) tests. Criterion tests are either discrete job tasks or simulations, whereas surrogate tests measure components of fitness that are associated with performance on job-related tasks (Milligan, Reilly, Zumbo, & Tipton, 2016). In an occupational setting, the type of test used will often depend on practical factors such as resource availability, financial constraints and/or safety considerations. However, since the aim of the fitness test is primarily to determine potential operational performance and, ultimately, suitability for employment, it is essential that the psychometric properties of the tests demonstrate an acceptable level of validity and reliability. While criterion tests naturally have content or 'face' validity, performances on these tests can be more markedly affected by factors such as weather conditions (where job simulations are performed outside) and test familiarisation than predictive tests (Boyd, Rogers, Docherty, & Petersen, 2015). Surrogate or predictive fitness tests (e.g. the Harvard step test) tend to be less complex, and can often be more easily administered in controlled conditions (i.e. a fitness or occupational health facility), which typically increases safety and test-retest reliability (Buckley, Sim, Eston, Hession, & Fox, 2004). However, predictive tests inevitably contain prediction error which, when applied to a workforce, can introduce bias and call into question their validity and occupational relevance when compared to criterion tests that more closely resemble the job. Identifying valid and reliable fitness tests for physically demanding jobs is important to help identify workers that can undertake their role effectively thus improving employee and public safety.

Several studies have investigated the association between a variety of fitness indices with simulated firefighting performance. Expectedly, cardiorespiratory fitness exhibits strong associations with tasks of longer duration, where little to no correlation is observed with individual, short duration (< 30 seconds) firefighting tasks (Rhea, Alvar, & Gray, 2004; Williford, Duey, Olson, Howard, & Wang, 1999). Weak to moderate (but significant) correlations have been observed with combined task simulations lasting between 5 and 10 minutes in U.S (r = - 0.38) (Williford et al., 1999) and Norwegian (r = - 0.53) firefighters (Von Heimburg, Rasmussen, & Medbo, 2006). Cardiorespiratory fitness exhibits the strongest associations with firefighting simulations lasting approximately 10 minutes (r^2 = 0.57) (Williams-Bell, Villar, Sharratt, & Hughson, 2009) or longer (r = - 0.72) (Von Heimburg, Medbo, Sandsund, & Reinertsen, 2013), supporting the notion that cardiorespiratory fitness is an important determinant of task performance during sustained tasks.

Studies investigating the reliability of timed task simulations have, however, received comparatively little attention. One study investigated the variability in task performance during a Canadian Forces firefighter work simulation test (Boyd et al., 2015). Participants completed six best-effort attempts of a standardised firefighting simulation with 24 to 48 hours between trials. Even with thorough orientation procedures and controlled test conditions, a continual improvement in performance on the occupational task was observed. While the test-retest correlations between sequential trials were high (r = 0.957 to r = 0.988), significant variations in mean task performance were evident between all trials (Boyd et al., 2015).

In a series of recent studies, we identified and established job-related simulations involving the critical and most physically demanding firefighting tasks (Stevenson, Siddall, Turner, Stokes, & Bilzon, 2016). We also quantified the metabolic cost of these simulations and proposed a minimum cardiorespiratory fitness standard for operational performance (Siddall, Stevenson, Turner, Stokes, & Bilzon, 2016). Treadmill-based tests of minimum cardiorespiratory fitness are now used to assess the occupational fitness of UK Fire and Rescue Service personnel. However, it is unknown whether a simulation, composed of these firefighting tasks can be used as a valid and reliable criterion test of operational fitness. This study was performed to: (a) assess the validity of a firefighting simulation test (FFST) to estimate maximal cardiorespiratory fitness (VO_2max) and; (b) establish the test-retest reliability of the FFST.

Methods

Participants

Sixty-nine (64 male, 5 female) operational firefighters from seven UK Fire and Rescue Services (FRS) volunteered for this study and gave written consent to participate following written and verbal explanation of the test procedures. All study participants were trained operational personnel and considered medically fit for firefighting duties. The study was approved by the University of Bath Research Ethics Approval Committee for Health (REACH Reference number: EP 12/13 6).

Study design

Previous studies have determined the critical and most physically demanding tasks undertaken by UK firefighters and identified standardised simulations for each of these tasks (Stevenson et al., 2016). From this work a criterion firefighting simulation test (FFST) was developed by combining three of these tasks. These were selected based on the critical and physically demanding nature of the tasks and the ability to be easily replicated on a standard fire service training ground. After a familiarisation procedure, participants were required to complete this firefighting simulation in the quickest possible time along with a maximal treadmill test to determine their maximal oxygen uptake (VO_2max).

Firefighting Simulation Test (FFST)

The firefighting simulation consisted of 'equipment carry', 'casualty evacuation' and 'hose run' tasks (Stevenson et al., 2016) and are described in detail in Table 1.

The tasks were selected to be completed in the order described (equipment carry, casualty evacuation, hose run) such that physical demand (reported previously) (Siddall et al., 2016) was incremental (i.e. least demanding to most demanding). Prior to undertaking the firefighting simulation, each firefighter was given full instruction of the task protocol and completed 2-3 attempts in the two weeks prior to the start of the testing procedure.

On the day of assessment, participants were instructed to avoid strenuous exercise and to eat and drink as normal. Participants completed the FFST in full firefighting ensemble (i.e. tunic, leggings, boots, flash-hood, helmet, gloves (total mass ~8.2 kg)), whilst carrying, but not breathing on, a self-contained breathing apparatus (SCBA) set (total mass 12.0 kg) during the casualty evacuation component of the simulation. Participants had to don and doff the SCBA set prior to and following the casualty evacuation elements of the task. Participants were asked to complete the FFST in the fastest time possible whilst adhering to standard operating procedures, manual handling and safety regulations. The time taken to complete each of the three stages/tasks of the FSST was recorded, as well as perceived exertion (Borg, 1982) at the end of the FFST.

Test retest reliability

At least 7-days later, a sub-sample of 22 participants (20 male, 2 female) completed a second best-effort attempt of the FFST to examine test-retest reliability. Both best-effort attempts were performed at the same fire station and at approximately the same time of day using the same equipment and pre-test conditions.

Cardiorespiratory fitness test

Participants also performed a maximal running protocol (in standard gym kit: running shoes, shorts and t-shirt) on a motorised treadmill (Life Fitness, Cambridge, UK) to determine their maximal oxygen uptake (VO_2max). This was conducted at the participants' designated fire stations in on-site gym facilities. Prior to the exercise test, anthropometric data (i.e. body mass, height, estimated body fat [BodyStat 1500, Bodystat Ltd., UK]) were collected. For the exercise test a 2-3 minute self-selected warm up was completed prior to commencing the test protocol which consisted of completing 4 (or 5 if required), 3-minute stages of walking or running at a constant speed with a 3% increase in gradient at the end of each stage until volitional exhaustion. Participants wore a portable breath-by-breath gas analyser (K4-B2, Cosmed, Italy). Peak VO_2max was established by calculating the mean of the highest minute of continuous oxygen consumption.

Statistical analysis

Statistical analysis was undertaken using SPSS version 22 (IBM, New York, USA). Mean anthropometric and performance (VO_2 max and task time) data were compared between male and female firefighters using independent t-tests. A Pearson's correlation coefficient was used to determine the relationship between the FFST time and VO_2 max. Shapiro-Wilk's tests were conducted to assess for normality of data distribution. Standard error of estimate (SEE) statistics were calculated to determine the size of the mean error from the estimation plot. Agreement between the FFST and VO_2 max scores was also assessed by determining the 95% limits of agreement (95% LoA) as previously described by Bland and Altman (1986). Test-retest reliability data were examined using Pearson's correlation coefficients between FFST attempts and a paired t-test was used to identify differences between mean performance times. Statistical significance was set at $p \leq 0.05$. The variability between attempts was assessed using coefficient of variation (CV).

Results

The mean (and standard deviation) firefighter physical and performance characteristics are presented in Table 2. All participants completed all tasks successfully/correctly and with "very hard" to "maximal" perceived exertion/effort (a rating of perceived exertion of ≥ 17 (Borg, 1982).

The male firefighters in this study were 8-years older ($p = 0.04$), 20-kg heavier ($p < 0.01$) and had a body mass index 5.4 $kg \cdot m^{-2}$ greater ($p = 0.01$) than their female colleagues (Table 2). Interestingly, there were no significant differences between males and females in height ($p = 0.19$) or estimated body fat percentage ($p = 0.12$), despite the fact that males were 3 cm taller and 4 percentage points leaner than the females. The mean time to complete the firefighting simulation for all firefighters was 608 (± 90) seconds, with male firefighters (600

± 77 seconds) completing the task significantly quicker ($p < 0.01$) than the female firefighters (706 ± 57 seconds).

Figure 1 shows the relationship between FFST task completion time (in seconds) and measured VO_2 max (maximal treadmill test). The time to complete the FFST was highly inversely correlated with cardiorespiratory fitness (r = - 0.73, p = 0.01). The SEE was equivalent to 55 seconds on the FFST.

Data describing the physical characteristics of participants used in the reliability analysis are described in Table 3. The relationship between the FFST performance time trials are presented in Figure 2.

Participants that performed the two reliability trials were, on average, quicker during the second trial (595 (± 74) seconds) compared to the first (612 (± 83) seconds), although this difference was not statistically significant (p = 0.09). Test-retest reliability of the FFST was high, revealing a strong relationship between the mean completion time of trial 1 and trial 2 (r = 0.84, p = 0.01). The coefficient of variation between the two tests was calculated to be 4.5%, equivalent to 27 seconds.

Discussion

This study was undertaken to investigate the validity of a best-effort performance on a firefighting simulation test (FFST) to estimate the minimal maximal cardio-respiratory fitness (VO_2max) required for operational firefighting and to determine the test-retest reliability. The time taken to complete the FFST demonstrated a strong inverse correlation with VO_2max demonstrating that performance on the FFST is strongly determined by cardiorespiratory fitness. Furthermore, the FFST simulation demonstrate a high degree of reliability (r = 0.842). However, the prediction error associated with the simulation (SEE of 6.13 ml.kg^{-}

$^{1}.min^{-1}$), the 95% limits of agreement with VO_2 max (16.2 $ml.kg^{-1}.min^{-1}$) and the test-retest coefficient of variation (4.5%) suggest it may not be suitably accurate to supersede already widely-used predictive treadmill tests and may not be considered an appropriate criterion fitness test if used in isolation.

Whilst previous studies have identified equivocal findings in the relationship between firefighting task performance and cardiorespiratory fitness, test protocols used have varied widely (Von Heimburg et al., 2013). Studies investigating individual tasks or short-duration simulations have reported little to no association with cardio-respiratory fitness (Rhea et al., 2004; Williford et al., 1999). However, work involving firefighting simulations lasting more than a few minutes have reported stronger and statistically significant associations with overall task performance (Von Heimburg et al., 2006; Williford et al., 1999), with the strongest correlations in firefighting simulations lasting greater than 10 minutes (Williams-Bell et al., 2009); (Von Heimburg et al., 2013). Interestingly, a number of these studies have also reported that when isolated from the complete simulation, specific tasks involving stair climbing (Williford et al., 1999) and casualty rescue (Von Heimburg et al., 2006) have elicited stronger correlations compared with the overall task performance. This supports the notion that the specific type or nature of the physical task(s) involved in the simulation, may be as important as the duration of the task itself, in terms of determining the reliance on and strength of association with cardiorespiratory fitness. In support of this, stair climbing and casualty rescue tasks have been identified as being some of the most aerobically demanding tasks in firefighting (Elsner, 2008; Gledhill & Jamnik, 1992; Siddall et al., 2016). These factors suggest that the combination of several different tasks, and subsequently, the total duration of the simulation in the current study may have facilitated the validity as an operationally relevant test of cardiorespiratory fitness.

Other methodological differences may also affect the relationship between the fitness variable and firefighting task performance. Two of the longest duration simulations described in the literature, used specifically designed firefighter applicant (Williams-Bell et al., 2009) or incumbent (Von Heimburg et al., 2013) simulation tests, where the former separated firefighting activities with periods of (walking) recovery (Williams-Bell et al., 2009). The nature of urgency by which the firefighting tasks are completed may also be of importance. Participants completing the simulations in the studies described were instructed to complete the tasks, "with no unnecessary waste of time" (Von Heimburg et al., 2006), "at a steady and rapid pace" (Williford et al., 1999), or "as quickly as possible" (Rhea et al., 2004; Von Heimburg et al., 2013). It is therefore impossible to know whether participants in these studies achieved best-effort performances, which would be more likely to elicit a stronger correlation with maximal cardiorespiratory fitness. In the present study, participants completed the FFST in the fastest possible time without walking or recovery between tasks. Furthermore, upon completion of the FFST, all participants reported an RPE of 17 or more. This represents the most likely reason why the relationship between task completion time and cardiorespiratory fitness in this study was stronger than those associations reported in the wider literature.

Despite the strong relationship between FFST performance and cardiorespiratory fitness observed in this study, only 54% of the total variance in task time could be explained by cardiorespiratory fitness. Additionally, the SEE of 6.13 $ml \cdot kg^{-1} \cdot min^{-1}$ equates to 55 seconds on the firefighting simulation. Whilst the SEE of 4.5% is similar to other reported studies (Williams-Bell et al., 2009), the value of this error in absolute terms indicates that this test may not be suitable to use as a stand-alone test to accurately determine fitness for duty in UK firefighters. The current minimum cardiorespiratory fitness standard for UK firefighters is 42.3 $ml \cdot kg^{-1} \cdot min^{-1}$, which would equate to a pass score of 10 minutes and 44 seconds on the

FFST. However, the SEE equates to 55 seconds on the simulation test. As such, in isolation, this test may not be suitably accurate to determine whether a firefighter was fit or safe enough to successfully undertake firefighting duties without undue physical strain.

The test-retest reliability of the FFST was strong (r = 0.84) and the variability between the two trials was 4.5% or 27.1 seconds, which is similar to other studies involving firefighting simulations (Boyd et al., 2015; Von Heimburg et al., 2013). Whilst the correlations reported by Boyd et al. (2014) were higher (r = 0.95 – 0.98) between each consecutive pair of six trials and the variation between trials smaller (2.6%) during the latter part of the study, these tasks were completed 24 to 48 hours apart in an indoor, temperature-controlled facility. This was in contrast to the current study where trials were completed outdoors, one to two weeks apart and where natural variations in the ambient conditions were present, which likely explains the weaker correlations observed.

Considering the work of Boyd et al. (2015), improvements in best-effort times on the FFST may have been achieved where further task familiarisation was possible. However, as this study was conducted using operational firefighters, whilst on-duty, emergency duties and other work commitments made this difficult to achieve. Indeed, due to time restrictions, it was also not possible to collect strength or muscular endurance performance data from the firefighters used in this study. This may have been useful, allowing the determination of the effects of multiple fitness characteristics on FFST performance, particularly as both strength and muscular endurance have been reported to be important determinants of firefighting performance in specific tasks (Stevenson, Siddall, Turner, & Bilzon, 2017).

The FFST investigated in this study has been shown to be a reasonably valid and highly reliable test for predicting cardiorespiratory fitness. However, the error terms observed for the validity and reliability raise questions of the efficacy of the test when aiming to apply the minimum fitness standard for UK firefighters (Siddall et al., 2016). The use of this test in isolation and without prior health and fitness screening may not be suitable to effectively identify firefighters that are above/below the fitness standard and therefore capable of safe and efficient operational firefighting performance. Therefore, this type of test may be better utilised when included within a larger fitness management process to ensure firefighters are at least above a minimum level of cardiorespiratory fitness before undertaking this test. This would minimise the possibility for error and, more importantly, help ensure the safety of the incumbent performing the FFST. Importantly, the FFST can now be used as part of a regular operational firefighter fitness training programme.

Acknowledgements

Special thanks to Avon, Buckinghamshire, Devon & Somerset, Leicestershire Manchester, Merseyside and West Yorkshire Fire & Rescue Services for their support throughout the project. This work was jointly funded by the Chief Fire Officer's Association, the FireFit Steering Group and the Fire Services Research and Training Trust (Project Code RE-FH1085).

References

Adams, E. M. (2016). Human rights at work: Physical standards for employment and human rights law. *Appl Physiol Nutr Metab, 41*(6 Suppl 2), S63-73. doi: 10.1139/apnm-2015-0552

Baur, D. M., Christophi, C. A., Tsismenakis, A. J., Cook, E. F., & Kales, S. N. (2011). Cardiorespiratory fitness predicts cardiovascular risk profiles in career firefighters. *J Occup Environ Med, 53*(10), 1155-1160. doi: 10.1097/JOM.0b013e31822c9e47

Baur, D. M., Leiba, A., Christophi, C. A., & Kales, S. N. (2012). Low fitness is associated with exercise abnormalities among asymptomatic firefighters. *Occup Med (Lond), 62*(7), 566-569. doi: 10.1093/occmed/kqs112

Bland, J. M. & Altman, D. G. (1986). Statistical-methods for assessing agreement between 2 methods of clinical measurement. *Lancet*, *1*, 307-310

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med Sci Sports Exerc, 14*(5), 377-381.

Boyd, L., Rogers, T., Docherty, D., & Petersen, S. (2015). Variability in performance on a work simulation test of physical fitness for firefighters. *Appl Physiol Nutr Metab, 40*(4), 364-370. doi: 10.1139/apnm-2014-0281

Buckley, J. P., Sim, J., Eston, R. G., Hession, R., & Fox, R. (2004). Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *Br J Sports Med, 38*(2), 197-205.

Elsner, K. L., Kolkhorst, F. W. (2008). Metabolic demands of simulated firefighting tasks. *Ergonomics, 51*(9), 1418-1425. doi: 10.1080/00140130802120259

Gledhill, N., & Jamnik, V. K. (1992). Characterization of the physical demands of firefighting. *Canadian Journal of Sport Sciences-Revue Canadienne Des Sciences Du Sport, 17*(3), 207-213.

Health and Safety at Work Act. (1974). *Elizabeth II. Chapter 22. (1974) London: The Stationary Office. Available at:* http://www.legislation.gov.uk/ukpga/1974/37/contents *(Accessed 22nd November 2016).*

Kales, Stefanos N., Soteriades, Elpidoforos S., Christophi, Costas A., & Christiani, David C. (2007). Emergency duties and deaths from heart disease among firefighters in the United States. *New England Journal of Medicine, 356*(12), 1207-1215. doi: 10.1056/NEJMoa060357

Maguire, B. J., Hunting, K. L., Smith, G. S., & Levick, N. R. (2002). Occupational fatalities in emergency medical services: a hidden crisis. *Ann Emerg Med, 40*(6), 625-632.

Milligan, G. S., Reilly, T. J., Zumbo, B. D., & Tipton, M. J. (2016). Validity and reliability of physical employment standards. *Appl Physiol Nutr Metab, 41*(6 Suppl 2), S83-91. doi: 10.1139/apnm-2015-0669

Plat, M. C., Frings-Dresen, M. H., & Sluiter, J. K. (2012). Which subgroups of fire fighters are more prone to work-related diminished health requirements? *Int Arch Occup Environ Health, 85*(7), 775-782. doi: 10.1007/s00420-011-0720-x

Rhea, M. R., Alvar, B. A., & Gray, R. (2004). Physical fitness and job performance of firefighters. *J Strength Cond Res, 18*(2), 348-352. doi: 10.1519/R-12812.1

Siddall, A. G., Stevenson, R. D., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2016). Development of role-related minimum cardiorespiratory fitness standards for firefighters and commanders. *Ergonomics*, 1-9. doi: 10.1080/00140139.2015.1135997

Smith, D. L., Barr, D. A., & Kales, S. N. (2013). Extreme sacrifice: sudden cardiac death in the US Fire Service. *Extrem Physiol Med, 2*(1), 6. doi: 10.1186/2046-7648-2-6

Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2017). Physical Employment Standards for UK Firefighters: Minimum Muscular Strength and Endurance Requirements. *J Occup Environ Med, 59*(1), 74-79. doi: 10.1097/JOM.0000000000000926

Stevenson, R. D., Siddall, A. G., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2016). A task analysis methodology for the development of minimum physical employment standards. *Journal of Occupational and Environmental Medicine. In Press.*

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*(2), 111-126. doi: 10.1080/00140130500435793

Von Heimburg, E., Medbo, J. I., Sandsund, M., & Reinertsen, R. E. (2013). Performance on a Work-Simulating Firefighter Test Versus Approved Laboratory Tests for Firefighters and Applicants. *International Journal of Occupational Safety and Ergonomics, 19*(2), 227-243.

Williams-Bell, F. M., Villar, R., Sharratt, M. T., & Hughson, R. L. (2009). Physiological demands of the firefighter Candidate Physical Ability Test. *Med Sci Sports Exerc, 41*(3), 653-662. doi: 10.1249/MSS.0b013e31818ad117

Williford, H. N., Duey, W. J., Olson, M. S., Howard, R., & Wang, N. (1999). Relationship between fire fighting suppression tasks and physical fitness. *Ergonomics, 42*(9), 1179-1186. doi: 10.1080/001401399185063

Figure Legends

Figure 1. Relationship between completion time (seconds) on the firefighting simulation versus VO_2max (n = 69, r = - 0.734, r^2 = 0.539), with line of best fit.

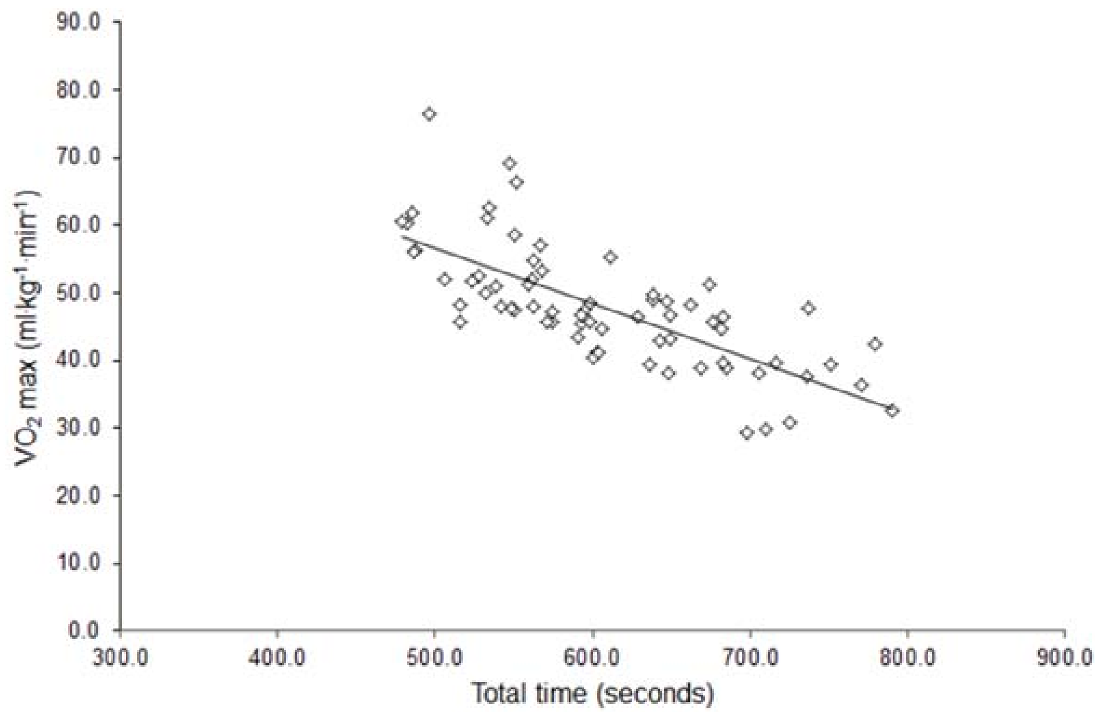

Figure 2. Test-retest reliability: relationship between the first and second attempts at the firefighter simulation test (n = 22, r = - 0.842, r^2 = 0.708).

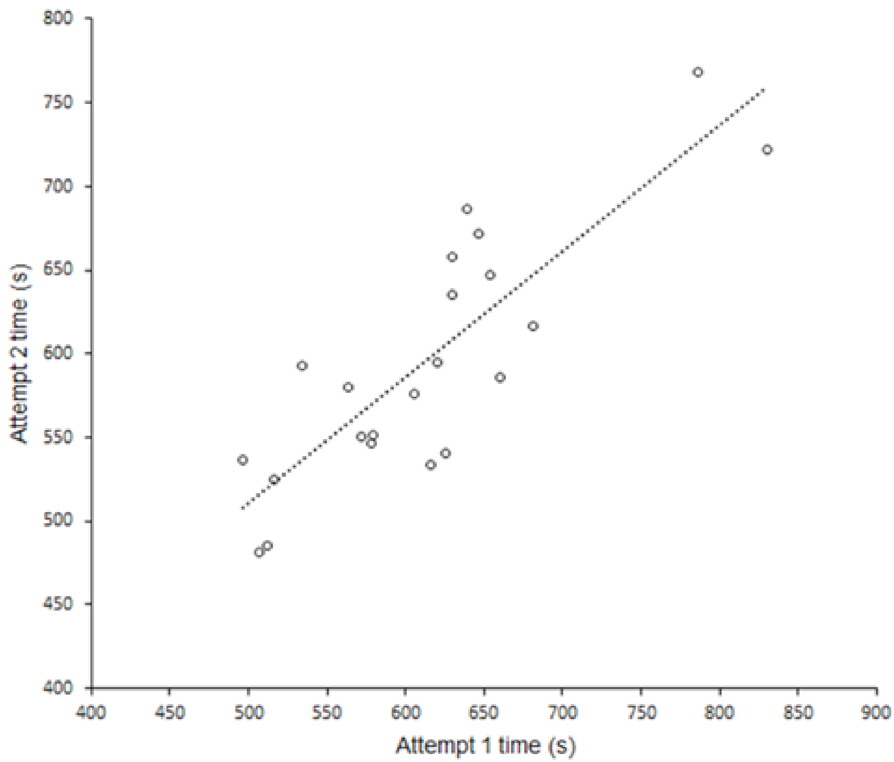

Table 1. Description of the consecutive components of the firefighting simulation

Component	Description
1. Equipment carry	A simulated equipment-handling task carrying firefighting equipment. Performed by walking a flat 25 m course while carrying a 25 kg barbell. Consisting of completing 8, 25 m shuttles (total distance 200 m), followed by:
2. Casualty evacuation	A simulated entry to (no actual door), and rescue of an unconscious casualty from, an industrial building. Consisting of: Dragging a charged hose reel 25 m; walking back 25 m; dragging another section of the hose 50 m; dragging a 55 kg dummy 50 m (total distance 150 m), followed by:
3. Hose run	A simulated water relay task to establish a water supply from a fire hydrant to a fire appliance 100 m apart using a total of four standard 70 mm hoses (weighing 13 kg) completed over a flat 25 m course. Consisting of: running 8 x 25 m, carrying 2 hose 75 m and 1 hose 25 m, rolling out 1 hose 25 m, then another hose 25 m; running 50 m, carrying 2 hose 25 m and 1 hose a further 25 m; rolling out 1 hose 25 m, then another hose 25 m; running 8 x 25 m (total distance 700 m).

Table 2. Study participant physical and performance characteristics (mean ± SD).

	All (n=69)	Male (n=64)	Female (n=5)
Age (y)	40 ± 8	41 ± 8	33 ± 4*
Mass (kg)	85.8 ± 12.8	87.2 ± 12.0	67.2 ± 5.4*
Height (cm)	178 ± 6	178 ± 6	175 ± 6
Body Mass Index ($kg \cdot m^{-2}$)	27.0 ± 3.6	27.4 ± 3.4	22.0 ± 2.5*
Estimated body fat (%)	19.7 ± 5.5	19.4 ± 5.5	23.4 ± 4.7
VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$)	47.8 ± 9.0	48.0 ± 9.2	45.5 ± 4.2
Simulation task time (s)	608 ± 90	600 ± 77	706 ± 57*

* *Significant difference between male and female firefighters ($p < 0.05$).*

Table 3. Reliability study participants' physical and performance characteristics (mean ± SD).

Variable	All (n=22)
Age (y)	42 ± 7
Mass (kg)	81.9 ± 11.6
Height (m)	1.78 ± 0.06
Body Mass Index ($kg \cdot m^{-2}$)	26.0 ± 3.0
Estimated body fat (%)	18.2 ± 4.9
Simulation task attempt 1 time (s)	612 ± 83
Simulation task attempt 2 time (s)	595 ± 74