Human Performance Laboratories (Australia)

UOW-CHAP-HPL-Report-049

PHYSIOLOGICAL EMPLOYMENT STANDARDS FOR FIREFIGHTERS: *REPORT 4:* RECOMMENDED PHYSIOLOGICAL EMPLOYMENT STANDARDS FOR THE FIREFIGHTERS OF FIRE & RESCUE NSW.

John A. Sampson, Hugh H.K. Fullagar, Nigel A.S. Taylor
and Herbert Groeller

Centre for Human and Applied Physiology
Faculty of Health and Behavioural Sciences
University of Wollongong

2013

Report commissioned by:
Fire & Rescue New South Wales
Sydney, NSW, Australia.

Physiological employment standards for firefighters: *Report 4:* Recommended physiological employment standards for the firefighters of Fire & Rescue NSW.

EXECUTIVE SUMMARY

Background:

This report summarises the last two of five research projects that led to the development of a pre-employment screening test for firefighters. The aim of this research was to facilitate the identification of capable and robust recruits for Fire & Rescue New South Wales (FRNSW), such that it could sustain the capability of its workforce, whilst simultaneously minimising the risk of injury to both firefighters and members of the community.

The current investigators were tasked with identifying screening tools and physiological standards to facilitate these objectives. The overall project involved five discrete Research Phases leading to the provision of a series of sensitive and specific screening tools, and employment standards (thresholds) that will maximise the identification of true positive (suitable) and true negative (unsuitable) outcomes during recruit screening. However, these tools were also aimed at minimising the number of false positive (recruiting the unsuitable) and negative results (rejecting the suitable) during screening. The critical legal and scientific steps leading to the development of *bona fide* physiological employment standards have been established (Table E-1), and these formed the framework for this research.

Table E-1: Procedural summary and framework for developing *bona fide* pre-employment screening tests and physiological employment standards. Steps performed within this research Phase have been coloured.

Project phase	Step	Description
0	1	Justify need for establishing employment standards
	2	Establish a Project Management Team
1	3	Familiarise Research Team with the trade
	4	Trade review and preliminary analysis of all tasks
	5	Identify the essential, physically demanding tasks
	6	Validate and approve the fire-fighting task list
	7	Employee survey: importance, difficulty, frequency of tasks
2	8	Characterise critical tasks: observe, measure, quantify
	9	Determine criterion fire-fighting tasks
	10	Validate and approve criterion fire-fighting tasks
3	11	Develop defensible physiological screening tests
	12	Standardise screening tests and administration

Project phase	Step	Description
	13	Validate and approve screening tests
4	14	Evaluate validity and reliability of screening tests
	15	Acknowledge and approve performance standard development
5	16	Develop physical performance standards
	17	Validate and approve performance standards
	18	Implement pre-employment screening
		Review the screening process and its outcomes: ongoing

In Phase One of this project, 106 permanent and retained firefighters were interviewed, and a list of essential (critical) fire-fighting tasks was established. In consultation with senior subject-matter experts, 31 tasks were identified for further investigation, and a workforce survey was administered. More than 1,000 operational firefighters participated (717 permanent, 272 retained, 22 incomplete), providing information relating to task importance, and performance frequency, duration and difficulty. This facilitated identification of the essential (critical) fire-fighting tasks for FRNSW (Taylor *et al.*, 2012a).

Research Phase Two was directed at quantifying the physical and physiological demands of performing these critical tasks under work simulations (Taylor *et al.*, 2012b). Operational firefighters from a wide range of skills, ages, body sizes and fitness levels participated (N=51), with the female representation reflecting that of the workforce. Simulations were designed by subject-matter experts and controlled by Training Officers.

The third research Phase focussed on analysing these critical fire-fighting tasks, and the physical and physiological demands they placed upon firefighters. This involved detailed task analysis, leading to the generation of a list of criterion tasks, and a valid prescription for the design of physiological screening tests for firefighters. From this distillation and the operational constraints imposed on subsequent screening by FRNSW, the Research Team proposed a preliminary format for the Physical Aptitude Test, recommending that the following physical tests be conducted as a test circuit (Groeller *et al.*, 2012a, 2012b):

- Criterion task one: two single-sided (unilateral) carriage task
- Criterion task two: holding tasks above and below shoulder height
- Criterion task three: hose drag
- Criterion task four: fire attack
- Criterion task five: firefighter rescue.

Aims:
The last two Research Phases were directed at two objectives. The first aim was to develop and trial possible screening tests for the Physical Aptitude Test, and thereby quantify the validity of each test item proposed by the Research Team (Groeller *et al.*, 2012a, 2012b). This was an integral aspect of the research, since screening tests must provide valid

measures of the key physiological attributes of capable firefighters. Without this step, these screening tests would be indefensible. The research undertaken to derive this information is described within Section Two of this report, and it provided the Research Team with a scientific justification for each test item included within the Physical Aptitude Test circuit.

Having achieved this outcome, the second objective was to administer the provisional test circuit to a large number of operational firefighters. Data from these trials were evaluated by the Research Team and Station Officers, and this enabled the development of recommended performance standards for these screening tests. This research is described within Sections Three and Four of this report.

Finalising screening tests and setting task performance rates

The Research Team proposed that individuals who could satisfactorily complete all of the criterion activities within each class description identified within Table 2 of this report, would possess the minimal physiological attributes necessary to perform the tasks of contemporary fire-fighting within FRNSW (Groeller *et al.*, 2012a, 2012b). With this in mind, the isolated screening tests were performed at intensities that would be reflective of the demands of contemporary fire fighting. However, the Research Team also sought opportunities to maximise testing efficiency, either by performing more easily administered predictive tests as opposed to task simulations, or by eliminating duplicate tests.

A sample of 14 non-firefighters (males and females) recruited from staff and student at the University of Wollongong (mean age: 31.9 years [range: 22-59]) participated in 19 task replications. Testing was performed whilst wearing shorts, t-shirt, boots, personal protective equipment and breathing apparatus used by FRNSW. Test scores were established from performance time, strength, endurance or distance covered. In some situations, it was necessary to determine the speed at which some tasks should be performed, and in these cases, the tasks were typically performed at maximal, moderate and slow speeds. The most appropriate task performance intensity was determined from interpolation of the resulting data, relative to the oxygen demand observed when simulated activity was performed by operational firefighters (Taylor *et al.*, 2012b).

Recommendation one

It is recommended that the two load-carriage tasks (hazmat simulation and ventilation fan carry [loaded box stepping]) be considered as necessary test items.

Recommendation two

Since hazmat work is typically interspersed with unloaded activity, and since the ventilation fan carry is a relatively brief task, then it is recommended that the 5-min duration used for these tasks within this evaluation be reduced when performed within the Physical Aptitude Test.

Recommendation three

None of the anticipated ladder-raise substitute tasks provided meaningful predictive information to warrant their substitution into the Physical Aptitude Test. It is recommended that FRNSW undertake an additional exploratory investigation of this

critical task. Such an investigation should establish whether or not a satisfactory correlation exists between the ladder-raise task and a strength-based simulation task. A study of this nature is most appropriately conducted on recruits, but with FRNSW content experts used to determine objective standards for satisfactory performance. These standards may then be correlated with a separate strength-based simulation of the ladder raise.

Recommendation four
Due to the high face validity of the static-hold activity, it is recommended that it be included in the Physical Aptitude Test. It is recommended this activity should occur after the single-sided, load-carriage tasks to best replicate operational scenarios.

Recommendation five
The recommended work rate for this task should match the lower 95% confidence interval previously observed in the field (1.41 $L.min^{-1}$: Taylor *et al.*, 2012b).

Verification and approval of the staged screening tests
These screening tests, in the form of a prototype Physical Aptitude Test, were submitted to the Project Management Team for consideration, endorsement and validation. Approval to progress to the final Research Phase (development of physiological employment standards for firefighters) was also sought, and this was granted (Project Management Team meeting: November 19th, 2012 [Appendix Six]).

Pilot testing of the Physical Aptitude Test (circuit) using lay personnel
Fourteen non-firefighters (University staff and students) performed the prototype Physical Aptitude Test. Subjects were instructed the test was to be performed as quickly as possible, but at walking speed, whilst wearing shorts, t-shirt, boots, personal protective equipment and breathing apparatus used by FRNSW. These participants also wore a portable gas analysis system and heart rate monitor, with the following physiological variables were monitored continuously: heart rate, oxygen consumption, minute ventilation tidal volume and breathing frequency. The Physical Aptitude Test was comprised of six test items:

- Test items one and two: Unilateral carrying tasks
- Test item three: Static-hold task
- Test item four: Cardiorespiratory hose-dragging task
- Test item five: Fire-attack simulation
- Test item six: Firefighter-rescue simulation.

From these pilot trials, three further recommendations were tendered for consideration prior to embarking in further testing.

Recommendation six:
Since all participants experienced significant cardiorespiratory strain and fatigue, it is recommended that participants should be advised to set a work tempo that would ensure test completion. It is further recommended that firefighters be instructed to complete the test at operational intensities.

Recommendation seven:
Technique clearly influenced performance during the two hose-drag tasks and the simulated firefighter rescue. It is therefore recommended that participants be given sufficient opportunity to select and learn the most appropriate techniques. A range of information and technical advice should be provided to candidates prior to testing.

Recommendation eight:
The height restriction of 1.25 m during the simulated firefighter rescue forced the use of postures that prevented a rapid response to this emergency and increased the risk of injury. It is recommended that this height restriction be further considered, and possibly made less difficult to reflect a more realistic operational posture.

Field trials of the Physical Aptitude Test using operational firefighters
The emphasis of this Section was to evaluate the capacity of the Physical Aptitude Test to evaluate the performance potential of prospective firefighters relative to the observed demands of the job, when successfully performed by operational firefighters. The prototype Physical Aptitude Test is comprised of six load-bearing tasks that are performed in a circuit format. This performance order reflected operational sequences and intensities that may be encountered during an incident. The test sequence, loads, distances, durations and other performance specifications are outlined below:

- ***Test item one:*** Unilateral load carriage (distance: 195 m [6.5 by 30-m shuttles]; load: 26 kg): hazmat simulation
- ***Test item two:*** Stepping with a unilateral load (36 steps [26 cm each]; load: 17.5 kg): ventilation fan carriage simulation
- ***Test item three:*** Static hold (load: 19.5 kg, fixed duration: 3 min [3 by 40-s hold and 20-s rest]): motor-vehicle rescue simulation
- ***Test item four:*** Charged hose drag (total distance: 300 m [10 by 30-m shuttles: 5 loaded and 5 unloaded]; resistive load: 265 N reel resistance including 11-kg hose mass): bushfire simulation
- ***Test item five:*** Fire attack (height restriction: 1.25 m; total distance: 60 m [2 by 30-m shuttles: 1 loaded and 1 unloaded recovery walk]; resistive load: 265 N reel resistance including 11-kg hose mass)
- ***Test item six:*** Firefighter rescue (height restriction: 1.55 m; distance: one 10-m excursion; resistive load: 550 N reel resistance including 30-kg mass attached to the harness of a breathing-apparatus backplate).

A sample of 148 firefighters across ranks up to, and including Superintendent volunteered to complete this test circuit. This sample was made up from metropolitan and regional firefighters from both permanent and retained stations (mean age: 39.1 years [range: 19-64]). All tasks were performed whilst wearing shorts, t-shirt, boots, personal protective equipment and breathing apparatus (12.5 kg) used by FRNSW. Trials took place at the University, and at Fire Stations and other centres in NSW (Alexandria, Blacktown, Campbelltown, Goulburn, Katoomba, Kirrawee, Queanbeyan). Firefighters were instructed that the test was to be performed at the fastest possible operational pace (*i.e.* at walking speed), with each person setting his/her own performance pace. Running was not permitted, and firefighters could stop to rest at any time within or between tasks, and for any duration,

except during the static-hold simulation.

In this final Research Phase, the prototype Physical Aptitude Test was shown to be a good reflection of the physiological and physical capacity of potential firefighters. It was not found to be dependent upon the skill of the test participants.

Recommendation nine: Clothing and equipment specifications for testing:
It is recommended that the Physical Aptitude Test be performed whilst carrying a load of 20 kg. This load should be comprised of the following items: underwear, t-shirt, shorts, socks and running shoes (all supplied by the person being tested), and thermal protective clothing, self-contained breathing apparatus (12.5 kg) and ankle weights (1 kg for each leg). The last three items are to be supplied by FRNSW.

Recommendation ten: Pass threshold and buffer zone (test items 1-4):
Based upon the performance speeds and work rates that would elicit the threshold metabolic costs seen in firefighters performing these tasks effectively, and at an acceptable level of operational performance, it was determined that the first four tasks should be completed within 15 min. It is therefore recommended that this 15-min time be considered as a screening threshold (performance standard) for the first four items of the Physical Aptitude Test.

It is further recommended that a buffer zone of 5% (45 s) be applied to this standard to minimise performance uncertainty. Individuals falling outside this buffer would fail the test at this point, and would not be permitted to continue. Those slower than the 15-min threshold, but inside the buffer zone, would be allowed to continue the test, with an opportunity to make up time on the last two test items.

Recommendation eleven: Pass threshold and buffer zone (test items 5-6):
It is recommended that pass threshold time for these two tasks be set at 1 min 53 s. It is again recommended that a 5% buffer (6 s) be used, with failing scores outside the buffer zone constituting absolute failure.

Recommendation twelve: Physical Aptitude Test pass and failure criteria:
It is recommended that persons satisfying the three static-position holds (item three), as well as both of the test threshold times, should be deemed to have passed the test.

However, it is recommended that four ways may exist for one to absolutely fail the Physical Aptitude Test, and that failure on any one of these criteria should represent test failure, independently of the performance on any of the other test components:

- failure to complete the three positional holds for the required duration: such individuals could be allowed to continue the test, but should still be required to pass this component as part of the complete test sequence
- failure to satisfy the time threshold for test items one-four, and with a score outside the buffer zone
- failure to satisfy the time threshold for test items five-six, and with a score outside the buffer zone

- the attainment of two test times beyond the threshold targets for test items one-four and five-six, both of which fall within the respective buffer zones.

It is further recommended that individuals completing the test in a time frame that falls within the 5% buffer zones for either of the test-item groups (*i.e.* items 1-4 or 5-6) should be identified as candidates for whom retesting might be appropriate.

Recommendation thirteen:
It is advised that all candidates be given a full briefing and be walked through the Physical Aptitude Test before being tested. It is further recommended that candidates be given adequate opportunity to practise each task, and be offered information and advice concerning optimal performance techniques.

AUTHORS

Dr. John A. Sampson, *Ph.D.*
University of Wollongong.
John has recently completed his Doctoral dissertation within the School of Health Sciences (University of Wollongong), examining musculoskeletal adaptation to resistance exercise. His primary research interest is encompassed within the physiological adaptations that are associated with exercise. The focus of his current and future work is in manipulating exercise training programmes to reduce total work, but without compromising the associated physiological adaptations observed during exercise training.

Hugh H.K. Fullagar, *B.Sc.*
University of Wollongong.
Hugh is a Master of Science student within the School of Health Sciences (University of Wollongong). His undergraduate training was in the discipline of exercise science, majoring in exercise physiology. His current research is centred around the development of valid and reliable physiological employment standards for Fire & Rescue NSW.

Assoc. Prof. Nigel A.S. Taylor, *Ph.D.*
University of Wollongong.
Nigel is a human stress physiologist with over 30 years of research experience, and with particular research emphases within exercise physiology, temperature regulation, and work-based physical and physiological assessments and interventions (occupational physiology). His research training was based in Europe and North America, and included extensive involvement with Defence organisations from both continents. He participates in a six-way research collaboration (Environmental Physiology and Ergonomics Research Exchange) involving leading research laboratories in France, Japan, Slovenia, Sweden and the United Kingdom (www.uow.edu.au/health/epere/index.html). Nigel's research often focuses upon the interface between the worker and the environment, and how human performance may be optimised under physical and environmental extremes. He has an extensive research background, with more than 300 publications. Nigel is the Chair of the ***IUPS Advisory Committee for Thermal Physiology and Pharmacology****. He is the Reviews Editor for the* **European Journal of Applied Physiology** *and an International Editorial Board member for seven other refereed journals.*

Dr. Herbert Groeller, *Ph.D.*
University of Wollongong.
Herb has over 20 years of experience working as an exercise physiologist with occupational, healthy active, sports-injured, aged and cardiac patient groups. He has also worked with various groups as a consultant, advising on the physiological demands and requirements of

specific work tasks, the implementation of functional testing regimens, and medical screening procedures and health initiatives. Herb's past research centred upon the impact of exercise habits on cardiovascular risk factors in middle-aged and older males. His current research focus is upon skeletal muscle strength adaptation within healthy and clinically relevant population samples.

CONTACT DETAILS

Centre for Human and Applied Physiology

Nigel A.S. Taylor, *Ph.D.*
Associate Professor
Centre for Human and Applied Physiology
www.uow.edu.au/health/chp/
School of Health Sciences
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522
Australia.
Telephone: 61-2-4221-3463
Facsimile: 61-2-4221-5945
Electronic mail: nigel_taylor@uow.edu.au

Fire & Rescue New South Wales:

Brendan Mott
Team Leader Health and Fitness
Health and Safety Branch
Fire & Rescue New South Wales
227 Elizabeth Street
Sydney, NSW 2000
Australia
Telephone: 61-2-9265-2800
Facsimile: 61-2-9265-2986
Electronic mail: Brendan.Mott@fire.nsw.gov.au

TABLE OF CONTENTS

SECTION 1: GENERAL INTRODUCTION

1.1 INTRODUCTION

This report arises from the final Research Phase in the development of, and the recommendations for physiological employment (occupational) standards for Fire & Rescue NSW. In the first Phase of this project, the physically demanding tasks of fire fighting were identified via focus-group sessions conducted at 11 Fire Stations (regional and metropolitan: 106 firefighters). Following consultation with Executive Staff and high-level, subject-matter experts from Fire & Rescue NSW, these tasks were consolidated into a list of 30 activities that were then evaluated using a workforce survey (>1,000 respondents). This last process led to the identification of 15 fire-fighting tasks deemed to be the most demanding, most critical and most frequently performed. These tasks became the critical, physically demanding activities of fire fighting, and this list was deemed to be a valid and representative subset of tasks performed by regional and metropolitan firefighters in NSW (Table 1; Taylor *et al.*, 2012a).

Table 1: The fifteen critical physical tasks of fire-fighting (Taylor *et al.*, 2012a).

Critical fire-fighting tasks
Critical task 1: Hazmat task
Critical task 2: Motor-vehicle rescue
Critical task 3: Rolling out hose (70 mm)
Critical task 4: Coupling hoses
Critical task 5: Locating and connecting to hydrant
Critical task 6: Drag charged 70-mm hose (lateral)
Critical task 7: Fire attack
Critical task 8: Firefighter down (rescue)
Critical task 9: Dragging charged 38-mm hose (uneven terrain)
Critical task 10: Stair climb dragging charged hose
Critical task 11: Prolonged use of hose (38 mm)
Critical task 12: Prolonged use of hose (70 mm)
Critical task 13: Ladder use (10.5 m)
Critical task 14: Stair climb with ventilation fan
Critical task 15: Using sledge axe to gain entry

The second Research Phase focussed upon observing, quantifying and evaluating the physical and physiological demands placed on firefighters performing these 15 tasks under controlled, but operationally relevant conditions (simulations). While these activities were not evaluated during actual fire fighting operations, each of the simulations was designed by a Training Officer, and then supervised by Station Officers whilst being performed by operational firefighters at work rates that replicated acceptable job performance at an incident. Operational firefighters (N=51; age range: 23-57 y; experience range: 1-29) participated in these simulations. Both male and female firefighters took part, with these

individuals being drawn from a range of Fire Stations to provide a representative mixture of task performance skills, as well as firefighter ages, body sizes and fitness levels. This wide variation was aimed at providing a fair and reasonable reflection of current operational firefighters within Fire & Rescue NSW. Since it would be inefficient to consider using all 15 activities within employment screening tests, the Research Team sought to exclude tasks if efficiencies could be gained without compromising the integrity of the process. To facilitate this outcome, a filtration algorithm (decision tree) was developed so that tasks could be culled to minimise the duplication of movement patterns and loads. On the basis of these analyses, 11 criterion tasks remained. Since there were movement pattern similarities across these tasks, then each criterion task was classified into one of four different movement categories (Table 2; Taylor *et al.*, 2012b).

Table 2: Criterion task movement classifications (Groeller *et al.*, 2012a).

Criterion task class	Class description	Criterion tasks
1	Single-sided carrying tasks	Hazmat task
		Locating and connecting to hydrant
		Ladder carriage (10.5 m)
		Stair climb with ventilation fan
2	Overhead push and hold tasks	Motor-vehicle rescue
		Ladder under-run (10.5 m)
		Using a sledge axe to gain entry
3	Cardiorespiratory dragging tasks	Fire attack
		Dragging charged hose (38 mm)
4	Critical strength task	Firefighter down (rescue)

The objective of Research Phase Three was to identify a battery of tests that could be used to either directly measure or to predict performance on each of these 15 tasks. The first step in this process was to evaluate each task with respect to three exclusion criteria:

- a low relative, whole-body physiological (metabolic) demand
- movement task duplication
- the availability of suitable substitution tasks.

Should any task satisfy one or more of these characteristics, it was considered for elimination. In this way, efficiencies within the proposed assessment battery could be found, without compromising the sensitivity or specificity of the test battery. Six operational constraints were then identified by Fire & Rescue NSW, and these were also considered when determining the suitability of each test for inclusion within the Physical Aptitude Test.

- environmental constraints
- equipment constraints
- height of the operating posture to avoid excessive heat and smoke exposure

- the structures and surfaces used during ambulatory and load-carriage tasks
- the mass to be used for the crucial strength task
- the correlation of generic lifting and locomotor activities with criterion tasks.

From this distillation and these operational constraints, the Research Team proposed a preliminary format for these firefighter assessments, recommending that the physical tests be conducted as a test circuit (Physical Aptitude Test: Groeller *et al.*, 2012a, 2012b):

- Criterion task one: single-sided (unilateral) carriage task or tasks
- Criterion task two: holding task or tasks (above and below shoulder height)
- Criterion task three: hose drag
- Criterion task four: fire attack
- Criterion task five: firefighter rescue.

1.1.1 Establishing legally defensible physiological employment (occupational) standards

The legal and scientific issues concerning the provision of genuine, certifiable (*bona fide*) and legally defensible physiological employment standards have been identified and described (Gledhill and Jamnik, 1992a, 1992b; Gledhill *et al.*, 2001; Constable and Palmer, 2000; Taylor and Groeller, 2003; Payne and Harvey, 2010; Jamnik *et al.*, 2012; Taylor and Billing, 2012; Tipton *et al.*, 2012). These steps (Table 3) formed the framework for this project, and facilitated answering two key questions:

- How certain can one be that those who are accepted into this job will be capable of successfully performing the necessary work-related tasks without exposing themselves to an undue risk of injury?
- How certain can one be that those who are deemed to be unacceptable, will actually be incapable of successfully performing the necessary work-related tasks, or that during the performance of these tasks, such individuals would expose themselves or others to an undue risk of injury?

1.1.2 Research aims

Having defined the physiological attributes necessary to successfully execute the critical tasks performed by operational firefighters within Fire & Rescue NSW, and an array of possible screening tests to evaluate these attributes in potential recruits, the last Research Phase was directed at two final objectives.

The first aim was to further develop and trial possible screening tests, and thereby quantify the validity of each test proposed by the Research Team (Groeller *et al.*, 2012a, 2012b). This was an integral aspect of this research, since screening tests must provide valid measures of the key physiological attributes of capable firefighters. Without this step, these screening tests would be indefensible. The research undertaken to derive this information is described within Section Two of this report, and it provided the Research Team with a scientific justification for each test item included within the final test circuit (Physical Aptitude Test).

Having achieved these outcomes, the second objective of this Research Phase was to administer the provisional test circuit (Physical Aptitude Test) to a large number of operational firefighters. Data from these tests were evaluated, and this enabled the

development of recommended performance thresholds (standards) for these screening tests. This research is described within Sections Three and Four of this report.

Table 3: Framework for developing *bona fide* pre-employment screening tests and employment standards. The current Research Phase is highlighted.

Project phase	Step	Description
0	1	Justify need for establishing employment standards
	2	Establish a Project Management Team
1	3	Familiarise Research Team with the trade
	4	Trade review and preliminary analysis of all tasks
	5	Identify the essential, physically demanding tasks
	6	Validate and approve the fire-fighting task list
	7	Employee survey: importance, difficulty, frequency of tasks
2	8	Characterise critical tasks: observe, measure, quantify
	9	Determine criterion fire-fighting tasks
	10	Validate and approve criterion fire-fighting tasks
3	11	Develop defensible physiological screening tests
	12	Standardise screening tests and administration
	13	Validate and approve screening tests
4	14	Evaluate validity and reliability of screening tests
	15	Acknowledge and approve performance standard development
5	16	Develop physical performance standards
	17	Validate and approve performance standards
	18	Implement pre-employment screening
		Review the screening process and its outcomes: ongoing

SECTION 2: EVALUATING TASK VALIDITY, FINALISING SCREENING TESTS AND SETTING TASK PERFORMANCE RATES

2.1 INTRODUCTION

This research was undertaken wholly at the University, and it involved the further development and trialing of possible screening tests for inclusion within the provisional test circuit (Physical Aptitude Test: PAT). This process permitted an evaluation of the validity of tests previously proposed by the Research Team (Groeller *et al.*, 2012a, 2012b), and thereby provided a scientific justification for including/excluding each test item. Moreover, this was an integral aspect of the current research, since screening tests must provide valid measures of the key physiological attributes of capable firefighters. Indeed, without this step, these screening tests would be difficult to defend.

2.1.1 Research aim

The aim of this research was to develop and trial possible screening tests for inclusion within the Physical Aptitude Test, and to evaluate the validity of these tests. In addition, to more closely replicate the metabolic demands of fire-fighting tasks, it was necessary to determine the speed at which several of these tests should be performed.

2.2 METHODS

2.2.1 The Project Management Team

Overall project management was undertaken through a Project Management Team. These individuals, their positions and their roles are summarised in Table 4.

Table 4: The Project Management Team.

Name	Position	Role
Alison Donohoe	Assistant Director Health and Safety (FRNSW)	Project Manager and Steering Committee Member
Darren Husdell	Director Human Resources (FRNSW)	Project Sponsor and Steering Committee Member
Jim Hamilton	Director Metropolitan Operations (FRNSW)	Steering Committee Member
Jim Smith	Acting Deputy Commissioner Emergency Management (FRNSW)	Steering Committee Member
Rick Griffiths	Acting Director Regional Operations (FRNSW)	Steering Committee Member
Geoffrey Parkes	Assistant Director Training (FRNSW)	Steering Committee Member
Brendan Mott	Team Leader Health and Fitness (FRNSW)	Research liaison and Steering Committee Member
Megan Smith	Manager Health Promotion (FRNSW)	Research liaison and Steering Committee Member

Name	Position	Role
Nigel Taylor	Associate Professor (UOW)	Scientific expertise
Herb Groeller	Senior Lecturer (UOW)	Scientific expertise
John Sampson	Lecturer (UOW)	Scientific expertise
Hugh Fullagar	Postgraduate student (UOW)	Data collection and analysis

2.2.2 Experimental subjects

A sample of eleven male and three female University staff and students (22-59 years old: Table 5) participated, with individuals of tall, medium and short stature targeted. Most individuals took part in each of the task simulations, and all subjects provided written, informed consent and completed a screening questionnaire prior to commencing procedures approved by the Human Research Ethics Committee (University of Wollongong). This screening procedure was aimed at identifying and eliminating individuals for whom the performance of these simulations might be considered an unacceptable health risk.

Table 5: Characteristics of the experimental subjects.

Subject	Gender	Age (y)	Height (cm)	Body mass (kg)
S1	Male	59.0	184.0	75.0
S2	Male	46.0	189.0	95.0
S3	Male	35.0	172.5	78.0
S4	Male	24.0	172.7	82.1
S5	Male	28.0	181.0	69.0
S6	Male	23.0	173.0	65.5
S7	Male	32.0	180.0	78.4
S8	Male	25.0	179.0	69.5
S9	Female	32.0	175.3	69.0
S10	Male	40.0	177.0	73.5
S11	Female	34.0	157.0	56.3
S12	Male	24.0	174.0	66.6
S13	Male	22.0	190.0	93.7
S14	Female	23.0	160.0	61.5
Mean		31.9	176.1	73.8
SD		10.59	9.72	11.09

2.2.3 Physiological measurements

2.2.3.1 Oxygen consumption and the data acquisition system

Oxygen consumption was measured using a portable open-circuit, expired gas analysis and ventilation system (Figure 1: Metamax 3B, Cortex Biophysik, Leipzig, Germany: mass = 1.82 kg). This involved separate determination of minute ventilation (turbine), and expired oxygen (electro-chemical cell) and carbon dioxide concentrations (infra-red). Data were recorded on a breath-by-breath basis and reported at 5-s intervals. Equipment was calibrated at the start of each test day, with calibration verification performed throughout each day.

Figure 1: The open-circuit, expired gas analysis and ventilation system.

2.2.3.2 Heart rate

Heart rates were monitored continuously from ventricular depolarisation (Polar Electro Sports Tester, Kempele, Finland). In most circumstances, these monitors were integrated into the data acquisition system for measuring oxygen consumption, with both sets of data being simultaneously recorded. These data were sampled on a breath-by-breath basis.

2.2.4 Overview of the fire-fighting task simulations

In this stage of Research Phase Four, the Research Team quantified and evaluated the physical and physiological demands of performing proposed screening test items (simulations) that were aimed at replicating the movement patterns, the muscular demands and the metabolic costs of the criterion tasks identified within Table 2. The intention of these trials was to determine the validity and the predictive capability of several proposed test items, and to establish suitable work rates for other tasks. For instance, it was necessary to determine walking speeds for the load-carriage tasks that would most closely approximate the metabolic demands observed during Research Phase Two (Taylor *et al.*, 2012b).

However, the fire attack and firefighter rescue were not simulated (Table 2), as these activities were deemed to be essential fire-fighting tasks, and it was anticipated that each would be performed in the provisional test circuit (Physical Aptitude Test) as high-speed tasks that demanded near-maximal efforts.

Each of these task activities or simulations was performed with participants wearing exercise clothing (shorts, t-shirt, running shoes), the thermal protective ensemble and personal protective equipment used by firefighters (7.7 kg), self-contained breathing apparatus (12 kg) and, on occasion, a portable data collection device (respiratory gas analysis system: 1.82 kg). This clothing, and the additional load carried by the participants, replicated the equipment worn and carried by firefighters within Fire & Rescue NSW.

2.2.4.1 Single-sided carrying tasks

For the first movement pattern (single-sided carrying tasks), two criterion tasks were identified (Table 2): a horizontal load-carriage task (the hazmat task) and a vertical load-carriage task (stair climb with ventilation fan). The construct validity[1] of these carrying tasks was established by first identifying the physiological attributes necessary to successfully perform each simulation (Taylor *et al.*, 2012b: see Tables 16 and 46), and then by designing screening tests that would provide an appropriate evaluation of these attributes (Groeller *et al.*, 2012a, 2012b).

Four unilateral, load-carriage tasks (Simulations 1A-1D) were evaluated within this category, and these were designed to replicate the physiological demands of the two criterion tasks (Taylor *et al.*, 2012b; Fullagar *et al.*, 2012):

- hazmat incident simulation (1.61 $L.min^{-1}$)
- stair climb with a ventilation fan (1.49 $L.min^{-1}$).

The average duration of the hazmat incident simulation in Research Phase Two was 15.2 min (Taylor *et al.*, 2012b). Whilst this time provided a more realistic reflection of steady-state physiological strain, it would not be feasible to use this duration within a test circuit (Physical Aptitude Test) consisting of multiple exercise assessments. However, an equivalent steady-state physiological strain could be imposed over a shorter duration within a longer test circuit. To do this, an answer to the following question was needed:

> Which horizontal walking speed would be required to replicate the metabolic demand of the hazmat incident simulation (Taylor *et al.*, 2012b)?
>
> - To provide data that would facilitate this prediction, it was necessary to collect data from this simulation performed at a minimum of three speeds: maximal safe speed (Simulation 1A) and two sub-maximal walking speeds (Simulation 1B).
>   - The provision of this answer leads to the establishment of

[1] Construct validity relates to attributes (constructs) that are associated with real-world scenarios, and can be viewed as an evaluation of how well a test measures these attributes. For instance, whilst we know that fire-fighting performance is a multi-faceted quality, we can measure some physiological attributes in isolation that are themselves related to job performance, or that underpin job performance (*e.g.* strength and local-muscle endurance). A screening test with high construct validity will measure these attributes, and will be well correlated with independent indices of the same physiological attributes.

criterion-related validity[2] for this task.

Since Fire & Rescue NSW would not have access to either a uniform set of stairs or even a sufficient number of stairs at every testing location in the State, then it was necessary to determine whether or not the stair climb task could be substituted with another activity. Three further questions were therefore evaluated:

- Does performance on the horizontal load-carriage task provide a valid prediction of one's ability to perform the vertical load-carriage task?
  - This helps to establish criterion-related validity.
- If the predictive power of the horizontal load-carriage task was not sufficiently strong, then could a box-stepping task be used as a valid surrogate for performance prediction on the vertical load-carriage task?
  - This also helps to establish criterion-related validity.
- If box stepping proved suitable, what cadence would be required to replicate the metabolic demand observed in Phase Two of this project for the carriage of a ventilation fan up stairs (Taylor *et al.*, 2012b)?
  - To provide data to facilitate this prediction, it was necessary to collect data at a minimum of three speeds: maximal safe speed (Simulation 1C) and two sub-maximal stepping speeds (Simulation 1D).
    - This helps establish the criterion-related validity of this task.

2.2.4.1.1 Simulation 1A: Hazmat incident replication

In this activity, participants performed a task designed to replicate the loads carried and the carriage distance observed within the hazmat incident simulation (Taylor *et al.*, 2012b). Subjects completed the replication at a self-selected maximal, but safe, walking speed. A 26-kg, liquid-filled plastic container with a mass equivalent to 50% of the heaviest object moved in a two-man hazmat simulation, was carried unilaterally over a 32-m synthetic grass track (Figure 2). The activity was performed as a series of "out and back" walks, with each walk being comprised of a 32-m walk out, followed by a 180°-turn at a fixed marker, and a 32-m return trip. This constituted one lap (64 m), with participants alternating between unloaded and loaded laps for a total of 16 laps. Thus, this sequence was designed to simulate walking into the hazmat incident exclusion zone (lap one: unloaded) and the subsequent carriage of a hazardous material container out of this zone (lap two: loaded). Whilst firefighters are required to carry objects of varying size, and often in pairs, it was important to standardise the object carried by all individuals, and for each lap to be performed with the same object. This standardisation was needed so that the required

[2] Criterion-related validity (also known as empirical validity) is typically established through the use of correlation analyses. For example, it is used to compare outcome measures from a predictive test with an established criterion reference. In this instance, the predictive capacity of the proposed screening test to reflect the physiological strain observed during tasks simulated by operational firefighters (the criterion reference: Taylor *et al.*, 2012b) will be evaluated, and the following question will be addressed: does the physiological strain (*e.g.* oxygen consumption) observed during the test match that observed during the corresponding work simulation? This will establish the predictive validity of individual tests proposed for use within the Physical Aptitude Test.

predictive precision of the walking speed would be achieved. Nevertheless, the face validity[3] and the content validity[4] of this simulation were both very high.

Figure 2: Hazmat incident replication.

Subjects were instructed to complete the task as quickly as possible at walking speed (one foot always in ground contact) without running, while maintaining a safe upright posture throughout the activity. However, participants could stop at any time, and change hands as required. Performance was assessed as the time taken to complete these 16 laps.

2.2.4.1.2 Simulation 1B: Hazmat incident replication (three walking speeds)

This activity was a repetition of the above simulation (Simulation 1A), but now with participants walking at three different absolute speeds, and for a 5-min duration at each speed (Figure 2). Two of these speeds were the same for all subjects, with the first (slow speed) set at 30% of the average maximal walking speed observed in Simulation 1A, and the second (medium speed) set at 60% of the average maximal walking speed. The third, 5-min trial was completed at maximal walking pace. Trials were performed in a randomised order with 5 min rest between trials. A researcher monitored walking speed to ensure that it remained constant. However, two participants from the original 14 were excluded from

[3] Face validity refers to the apparent similarity between a test or simulation and a work-related activity. Whilst this is an entirely subjective form of validity, and the most superficial form, it is often the most important form for workers, for their acceptance of a test as a realistic approximation of work tasks will determine their acceptance of, and their subsequent support for the activity within a test battery. Indeed, face validity can even be evaluated by lay people (*e.g.* the activity looks like it resembles a fire-fighting task).

[4] Content validity can also be subjective in nature, but it is one level higher than face validity. That is, it relates to the extent that test items (or a test battery) approximate fire-fighting tasks. When a test has good content validity, subject-matter experts would advise that the test is a fair and reasonable approximation of a fire-fighting task(s).

these assessments as the medium speed required was quicker than the best time they achieved during Simulation 1A. Participants were again instructed how to maintain a safe, upright posture, and they were again permitted to stop at any time and change hands as required. Oxygen consumption was measured continuously.

2.2.4.1.3 Simulation 1C: Loaded stair climb

The aim of this simulation was to replicate the ventilation fan carriage examined within Phase Two of this Project (Taylor *et al.*, 2012b), but with attention directed to the leading firefighter. This individual frequently walks backwards up the stairs, as the width of most staircases is too narrow to permit walking abreast with the ventilation fan. Therefore, to increase both the face validity and the content validity of this simulation, participants carried a 17.5-kg mass (liquid-filled plastic container: 50% of the mass of the ventilation fan)[5] in one hand, and completed the task walking backwards (Figure 3). The simulation commenced with a 7.3-m walk on level ground to the base of the stairs, and continued with the ascent of 64 steps (4 storeys). Participants were instructed to complete the task as quickly as possible, but at walking speed (one foot always in ground contact). Running was not permitted. Instructions were provided in how to maintain a safe, upright posture. Rest stops were permitted, as was the changing of the load between hands. Performance was assessed as the time to complete the task once.

Figure 3: Loaded stair climb.

2.2.4.1.4 Simulation 1D: Loaded box stepping (three stepping rates)

Participants performed three loaded, 5-min box-stepping trials. In each trial, the same 17.5-

[5] Screening tests are typically performed on individuals, so this required a load representing half the mass of the object. Whilst it is recognised that the load distribution for paired tasks varies according to stature differences, differences between the centre of mass of each individual relative to that of the object, the skill of each carrier and so on, this first-level approximation is acceptable in this circumstance.

kg mass (liquid-filled plastic container) was carried in one hand. Stepping involved a continuous step cycle onto, and down from a stable box (26 cm), with the stepping cadence set via an electrical metronome. However, to make this a more simple activity for the widest possible range of recruits, stepping up was performed in a forward direction (Figure 4). Three stepping cadences were used: one represented each individual's maximal stepping rate observed during Simulation 1C (*i.e.* each individual stepped at the rate matching his/her maximal cadence observed on the backward stair climb activity), and the other two stepping rates were set at the same absolute cadences for all subjects. These were set at 30% (slow cadence) of the average maximal stepping rate observed in Simulation 1C, and at 60% of the average maximal cadence (medium cadence). One participant from the original 14 was excluded from these assessments as the medium cadence was faster than the maximal cadence achieved during Simulation 1C. Trials were performed in a randomised order with at least a 5-min rest between trials. Participants were instructed in how to maintain a safe, upright posture and informed that they were able to stop at any time, and to change hands as required. Oxygen consumption was measured continuously.

Figure 4: Loaded box stepping.

2.2.4.2 Overhead pushing and holding tasks

The second movement pattern involved pushing and holding tasks, with three criterion tasks being identified (Table 2): motor-vehicle rescue, ladder under-run (10.5 m) and using a sledge axe to gain entry. Following consideration of advice from senior subject-matter experts within the Project Management Team regarding the last task (*i.e.* "the availability of alternative tools to assist with this task" [Rapid Intervention Kit]: Appendix One: Groeller *et al.*, 2012a), it was deemed that this task should be eliminated from the proposed screening

test, as it could be performed with less effort if the firefighter chose to use these tools. Thus, it could be argued that such a task should not be used to evaluate the physiological capability of potential recruits. This left two criterion tasks that each involved the holding of objects above and below shoulder height, and five different tasks (Simulations 2A-2E) were evaluated within this test category. The construct validity of these tasks was established by Taylor *et al.* (2012b: see Tables 18 and 44) and Groeller *et al.* (2012a).

Using the methods described below, several questions were evaluated:

- Firstly, it was requested that Fire & Rescue NSW determine whether the motor-vehicle rescue holding task (operating hydraulic shears for victim extrication from a motor vehicle) was to be performed at a fixed absolute or at some relative height.
    - Three relative heights were chosen by Fire & Rescue NSW for this task (Appendix Four): below the knees, at waist height and at eye level (with the arms positioned below the tool).
    - The shape and size of the load used, its total mass (19.5 kg) and its mass distribution must replicate the hydraulic shears.
        - These decisions increase both the face and content validity of the simulation to be used within the proposed screening test.
- Secondly, since this criterion task is typically performed as a series of work and rest cycles, the Research Team needed to determine whether or not the duration of the work phase affected the metabolic demand of this task.
- Finally, the Research Team explored the possibility that performance on the ladder under-run (10.5 m) task could be predicted from performance on a generic, single-lift task (box lift and place).
    - This contributes to the criterion-related validity of this assessments.

2.2.4.2.1 Simulation 2A: Hydraulic shears static-hold simulation

In this simulation, participants first performed three trials, each lasting 6 min. Every trial was a series of intermittent holds and rests, using a replica hydraulic tool (19.5 kg). The tool was held in three positions (1 min) relative to the stature of each participant: below knee height, between knee and hip heights, and between the shoulder and eye level (Figure 5). The position sequence was randomised across subjects and performed twice. Whilst the duration at each relative height remained constant (1 min), the duration of the static hold was varied: (a) 20-s hold and 40-s rest, (b) 30-s hold and 30-s rest, and (c) 40-s hold and 20-s rest. The resting phase involved placing the load onto a bench at the same height, and resting briefly. A researcher controlled the timing of these hold and rest durations. In addition, each participant performed a fourth trail which involved a static hold in each position until failure; defined as the inability to hold the mass at the required height, or the loss of the safe holding posture. Participants were instructed in how to maintain a safe, upright posture and were informed that they were able to stop at any time. Oxygen consumption was measured continuously.

Figure 5: Hydraulic shears, static-hold simulation.

2.2.4.2.2 Simulation 2B: Ladder-raise - task performance

All subjects were required to perform a single-person ladder raise from the ground to the vertical position (resting against an external wall), immediately followed by lowering the ladder to the ground. No assistance was provided to any subject. The ladder was a 10.5-m extension ladder (49.5 kg) provided by Fire & Rescue NSW. This is a highly specific, direct task replication with very high face and content validity. Before attempting the activity, participants were taught how to safely perform the lift, as specified by Fire & Rescue NSW: squat and grasp one end of the ladder with both hands (top rung); stand to raise this end of the ladder to shoulder height; extend both arms to elevate the end of ladder above the head (arms fully extended: both elbows straight and with both hands still holding the top rung); walk forwards by reaching to each successive rung and then extending it above the head (one hand must always be gripping a ladder rung); continue until the ladder is positioned against the side of the building. Participants were instructed to make contact with every ladder rung in succession. Whilst it is normal to place one hand on each rung, participants were instructed to use both hands if desired. Participants performed the raise and lower stages as quickly as possible whilst walking (one foot on the ground at all times). Running was not permitted. A 30-s rest period was provided between the completion of the ladder raise and its subsequent lowering. Performance times for the separate raise and lower stages were recorded. The feet of the ladder were held in position by one researcher. Another researcher acted as a safety monitor, holding a safety line that was secured to the top rung of the ladder, and connected via a pulley to the wall at the point to which the ladder would come to rest. This line was not held taught, but at sufficient length so that it could not fall if accidentally released. Failure to lift the ladder or incorrect lifting technique (loss of a neutral spine, excessive lordosis, stooped lifting) defined task failure.

2.2.4.2.3 Simulation 2C: Box lift and place - maximal load

A maximal lift and place task was used to replicate the movement pattern observed within the ladder raise. Subjects were required to lift a box (30.5 * 30.5 * 30.5 cm with lifting handles on both sides) from the ground, and to place it on a platform 1.5 m above the ground. The box was returned to the ground by the researchers. The aim of this activity was to determine the maximal load that each participant could safely lift and place. The task

began with a gender-specific mass: 10 kg (females) and 20 kg (males). An additional mass, selected by the participant within a 2.5-10 kg range, was added following each successful lift. A 2-min rest was provided between successive lifts, and this cycle continued until the participant could no longer safely lift and place the box. Subjects were instructed in how to correctly and safely lift and place this box. Incorrect lifting technique (*i.e.* loss of a neutral spine, excessive lordosis, stooped lifting), sliding the box onto the shelf or the loss of box control each defined test failure. In additional, a trial ended at volitional termination.

2.2.4.2.4 Simulation 2D: Box lift and place - endurance
This activity was a replication of Simulation 2C, but with two variations. Firstly, the box mass was set at 75% of the mass each person lifted in Simulation 2B. Secondly, the subject returned the box to the ground after each successful lift. The aim of this task was to determine how many times each participant could successfully and safely lift and place this box. Participants were instructed in how to safely lift and place the box. The task was terminated at volitional fatigue, or when the participant could no longer safely lift and place the box. Incorrect lifting technique (*i.e.* loss of a neutral spine, excessive lordosis, stooped lifting), sliding the box onto the shelf or the loss of box control defined test failure.

2.2.4.2.5 Simulation 2E: One-repetition maximum shoulder press
Participants performed a maximal (front) shoulder press within a lifting machine (seated), ensuring a controlled vertical movement. Participants were instructed to place both hands on the bar at a width corresponding to the maximal handgrip width on a ladder rung. Following a warm-up with the bar plus 10 kg, the researcher asked each participant: "How many more repetitions do you think you could perform using this mass?" From the answer provided, an estimated additional load to reach a one-repetition maximum[6] was derived. This process continued until a one-repetition maximum was achieved. Failure was defined as an inability to move the loaded bar through the required range of motion and with full extension at the elbows.

2.2.4.3 Cardiorespiratory dragging tasks
Two cardiorespiratory dragging tasks were simulated within the third criterion task class (Table 2): the fire attack and dragging a charged hose (38 mm: bushfire simulation). The construct validity of these tasks was established by Taylor *et al.* (2012b: see Tables 28 and 32) and Groeller *et al.* (2012a). It had already been determined that the former task would be included within the provisional screening test circuit, and its further evaluation has not been described within this report, with one exception:

- It was requested that Fire & Rescue NSW determine the height restriction to be used during this simulation, since the Research Team had been advised by Training Officers that firefighters were instructed to adopt safe operating postures (below the neutral plane) to avoid excessive heat and smoke exposure.
    - This height was deemed to be 1.25 m (Appendices Three and Four).
    - This restriction increases the face and content validity of the test.

[6] The maximal mass a participant was able to correctly lift once, and through the desired range of motion.

Whilst the latter task was essential for inclusion, due to its relevance to both metropolitan and regional firefighters, the typical duration of this task (>50 min) did not lend itself well to an efficient task within a test circuit. Fortunately, it is well established that the oxygen consumption of sub-maximal work will stabilise after 3-5 min (Henry and DeMoor, 1956; Di Prampero *et al.*, 1970; Whipp and Wasserman, 1972), so this will permit a shorter test duration. Accordingly, the hose-drag replication was performed over 6 min, and this was designed to replicate the average physiological demand of this criterion task (Taylor *et al.*, 2012b): 1.63 $L.min^{-1}$. Thus, one key question required an answer:

- Which walking speed would be required to replicate the metabolic demand observed for the bushfire hose-drag simulation (Taylor *et al.*, 2012b)?
    - The answer to this question provides criterion-related validity.

2.2.4.3.1 Simulation 3A: Hose-drag replication

This task was designed to replicate the demands of dragging a charged fire hose (38 mm) over bush terrain (bushfire simulation), and within an urban setting. The frictional load that would faithfully replicate the resistance experienced by the firefighter was unknown, but it would vary with the nature of the surface over which the hose was to be dragged. This load was determined using pin-loaded, hanging scales connected in series with a charged 38-mm hose. The hose was then dragged 30 m (the length of one hose) at both slow and fast walking speeds, over concrete and grass surfaces. The average force developed over this distance (265 N [27 kg]) was adopted, and this was applied to a custom-made line and reel resistance loader (Figure 6).

Figure 6: The hose-drag replication and resistance simulator.

In this simulation, work at two different intensities was performed in an alternating manner. The first stage of each work cycle was heavier, with the subject holding the branch of a 5.6-m section of 38-mm hose filled with sand (11 kg) and doubled over, giving a 2.8-m length of hose. The weighted hose acted as the interface with the resistance loader, and permitted the application of both the required a vertical and horizontal resistances that were determined for this activity (above). This combination contributed to both the face and content validity of this task. The participant walked away from (pulling against) the

resistance loader that had been set to provide the resistive force defined above. This walking distance was 30 m. On reaching a marker, participants passed the hose branch to the researcher and then walked back to the resistance device: the unloaded or light work stage. On reaching the loader, the researcher handed the hose branch back to the subject, who then performed the next hose-drag simulation. This two-stage sequence was repeated until 6 min elapsed (Figure 6). Participants were instructed to complete the task as quickly as possible at walking speed (one foot always in ground contact), but without running, and they were permitted to stop and rest at any time. Task performance was measured from the total distance covered. Oxygen consumption was measured continuously.

2.2.4.3.2 Simulation 3B: Hose-drag replication (two walking speeds)
This activity was a repetition of the above simulation (Simulation 3A), but now with participants walking at two different absolute speeds, and for a 6-min duration at each speed (Figure 6). These speeds were the same for all subjects, with the first (slow speed) set at 30% of the average maximal walking speed observed in Simulation 3A, and the second (medium speed) set at 60% of the average maximal walking speed. Trials were performed in a randomised order with 5 min rest between trials. A researcher monitored walking speed to ensure that it remained constant. Participants were again instructed how to maintain a safe, upright posture, and they were again permitted to stop at any time and change hands as required. Oxygen consumption was measured continuously.

2.2.4.4 Strength task
The final criterion task classification was a strength task (Table 2). The firefighter rescue simulation was to be used for this evaluation, with the construct validity of the task being established in earlier Research Phases (Taylor *et al.*, 2012b [see Table 30]; Groeller *et al.*, 2012a). However, this activity also required a height restriction:

- Fire & Rescue NSW was ask to determine the height restriction for this simulation, so that firefighters were forced to adopt a safe operating posture (below the neutral plane) to avoid excessive heat and smoke exposure.
    - This height was originally set at 1.25 m, but the difficulty of this restriction prompted consultation with Fire & Rescue (Appendices Three and Four), and this led to the height being set at 1.55 m.
    - This restriction increases the face and content validity of the test.

2.2.5 Data analysis
Within this report, descriptive statistics including the means, data ranges (minimal and maximal limits) and 95% confidence intervals are reported. In addition, two forms of correlation analysis (Pearson's product-moment and Spearman's rank-order correlation) were performed to investigate the predictive power of tasks within each class description.

2.3 RESULTS AND DISCUSSION
2.3.1 Single-sided carrying tasks
2.3.1.1 Simulation 1A: Hazmat incident replication
The hazmat incident replication was designed as a time-to-completion activity, such that the task was set at 16 laps (1,024 m), and subjects were asked to perform the simulation at their maximal walking pace. Two subjects were unable to complete this task, and terminated the

activity prior to completing the required 16 laps. Data from these individuals were excluded from subsequent calculations.

The average performance time for the simulation was 589.9 s (SD 141.6: range: 436.0-973.0 s), and this was attained through the participants generally maintaining a constant unloaded walking speed, but with a gradual, fatigue-induced reduction in their loaded walking speed (Figure 7). The average walking speed over the entire simulation was 1.74 $m.s^{-1}$ (range: 1.05-2.35 $m.s^{-1}$).

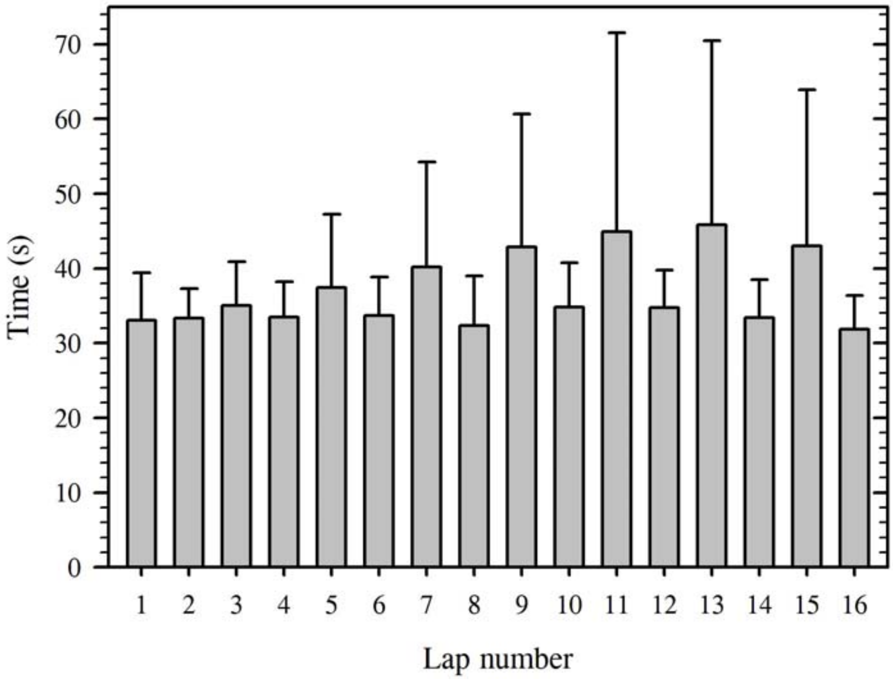

Figure 7: Hazmat incident simulation: 16 laps at maximal walking pace, alternating between loaded (odd laps) and unloaded walking. Data are means with standard errors of the means (N=14).

2.3.1.2 Simulation 1B: Hazmat incident replication (three walking speeds)

2.3.1.2.1 Physiological strain

This repetition of Simulation 1A was completed with participants walking for 5 min at each of two absolute speeds, and at their own maximal walking pace. One limitation with measuring the metabolic demand of work simulations is that some tasks are performed for quite brief periods. As a consequence, participants may fail to achieve metabolic steady states, since this can take several minutes to develop. Thus, the metabolic demand may have been under-estimated. To evaluate this possibility, oxygen consumption and heart rate data obtained during the fastest of these trials are presented in Figure 8, with clear steady states being evident.

Maximal performance data for all subjects during 5-min, best-time hazmat simulations are presented in Table 6. The distance covered had a 95% probability of falling between 330-444 m, whilst the 95% probability ranges for the heart rate and oxygen consumption responses were 148-168 $beats.min^{-1}$ and 1.45-1.91 $L.min^{-1}$.

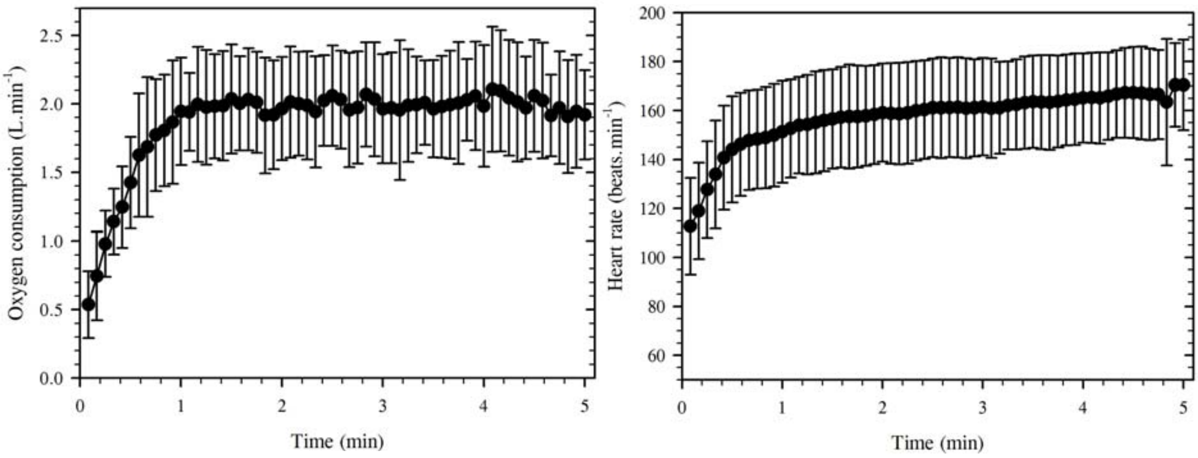

Figure 8: Oxygen consumption and heart rate data during hazmat incident simulation (5 min) performed at maximal walking pace. Data are means with standard errors of the means (*N*=14).

The average lap time for the maximal 5-min simulation was 41.4 s. To calculate the slow and medium walking speeds, this value was rounded to 40 s, and the slow speed was set at 68 $s.lap^{-1}$ (*i.e.* 40 s + 0.7 * 40) whilst the medium speed was set to 56 $s.lap^{-1}$ (*i.e.* 40 s + 0.4 * 40). However, these speeds remained greater than the average speed of the two subjects who failed to complete the 16-lap simulation (S11 and S14), so these individuals were omitted from the last two trials. Tables 7 summarises data from these two walking speeds. During these speed-controlled replications, the mean heart rate had a 95% probability of falling between 115-145 $beats.min^{-1}$ (slow) and 121-147 $beats.min^{-1}$ (medium). Similarly, the mean absolute oxygen cost could be expected to fall between 1.20-1.54 $L.min^{-1}$ (slow) and 1.32-1.62 $L.min^{-1}$ (medium).

Table 6: Maximal performance: 5-min hazmat simulation.

Subject	Distance (m)	Heart rate ($b.min^{-1}$)	Oxygen consumption ($L.min^{-1}$)
S1	481	145	2.05
S2	437	129	2.09
S3	368	179	1.7
S4	356	138	1.45
S5	544	172	2.28
S6	512	169	1.99
S7	448	164	2.24
S8	388	183	1.71
S9	288	159	1.32
S10	424	132	1.92

S11	160	184	1.09
S12	448	154	1.12
S13	352	141	1.64
S14	212	164	---
Mean	387	158	1.74
SD	109.2	19	0.4

Table 7: Hazmat incident simulation (5 min) performed at two walking speeds.

Subject	Heart rate (b.min^{-1})		Oxygen consumption (L.min^{-1})	
	Slow	Medium	Slow	Medium
S1	119	125	0.96	1.08
S2	100	101	1.32	1.72
S3	165	150	2.05	1.69
S4	150	140	1.2	1.19
S5	116	121	1.09	1.2
S6	132	132	1.46	1.72
S7	126	151	1.16	1.61
S8	178	176	1.15	1.2
S9	152	153	1.55	1.71
S10	115	107	1.57	1.59
S12	100	---	1.41	---
S13	104	113	1.54	1.5
Mean	130	134	1.37	1.47
SD	26	23	0.29	0.25

When this task was evaluated using operational firefighters who performed field-based simulations, the average heart rate response was 134 beats.min^{-1} and the mean oxygen cost was 1.61 L.min^{-1}. Therefore, these three walking speeds elicited mean physiological responses that were above, below and close to the strain observed in these firefighters.

2.3.1.2.2 Predictive power of the 5-min hazmat simulation

The 16-lap hazmat incident simulation was a fire-fighting criterion task. However, when developing recruit screening tests, it is often more efficient to use shorter-duration

predictive tests, particularly when screening many individuals across a variety of physiological attributes. Thus, the predictive utility of the 5-min maximal walking simulation was evaluated, and this was achieved by comparing performance outcomes from these two tests: time to complete the 16 laps and the maximal distance covered in 5 min. This comparison is illustrated in Figure 9, and 57%[7] of the variance in the 16-lap performance times can be explained on the basis of the distance covered during the maximal, 5-min simulation. Therefore, this 5-min replication can be considered to have a good predictive relationship with the longer, 16-lap simulation.

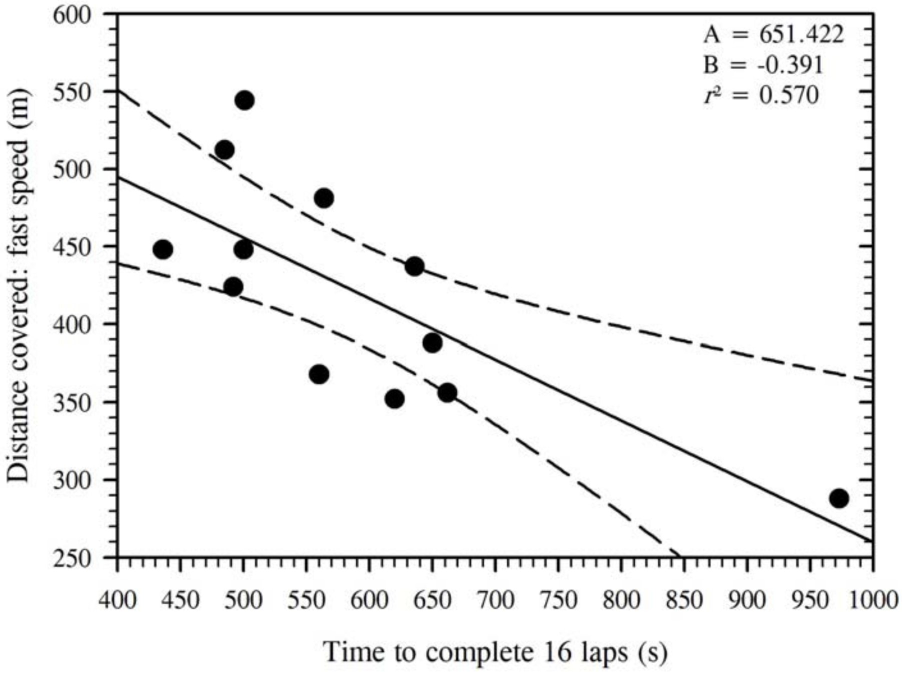

Figure 9: The correlation between the time to complete the 16-lap hazmat incident simulation at maximal walking speed, and the distance covered in a maximal, 5-min simulation of the same task. Individual scores are presented with linear modelling and 95% confidence intervals (*N*=12).

When the hazmat simulation was characterised using operational firefighters, the mean absolute oxygen cost was 1.61 $L.min^{-1}$, and there was a 95% probability of this mean falling between 1.47-1.75 $L.min^{-1}$ (Taylor *et al.*, 2012b). In a screening test designed to replicate the metabolic demand of this activity, it would be necessary to perform a closely related task at about the same oxygen cost. To achieve this, oxygen consumption data from each of the three 5-min simulations were used, with oxygen consumption first being regressed against the distance walked. From this relationship, one could predict with moderate confidence the walking distance that would elicit about the same oxygen cost as that observed in firefighters. Rather than just simulate the average oxygen cost of this task, it was deemed more appropriate to simulate the value at the lower end of the 95% probability range (1.47 $L.min^{-1}$), for this corresponded with values observed in operational firefighters

[7] This predictive power is indicated by the square of the correlation (r^2) shown on the Figure. For this type of field-based work, values around 0.6 are considered to be strong, and possess good predictive power.

who could still successfully complete the task, but at the lower end of the range of these physiological responses. Such a choice would minimise the risk of false negative recruiting; rejecting individuals who possess the necessary physical capacity to perform the simulation. This approach is illustrated in Figure 10, and was applied consistently to the relevant data sets within subsequent parts of this experimental phase.

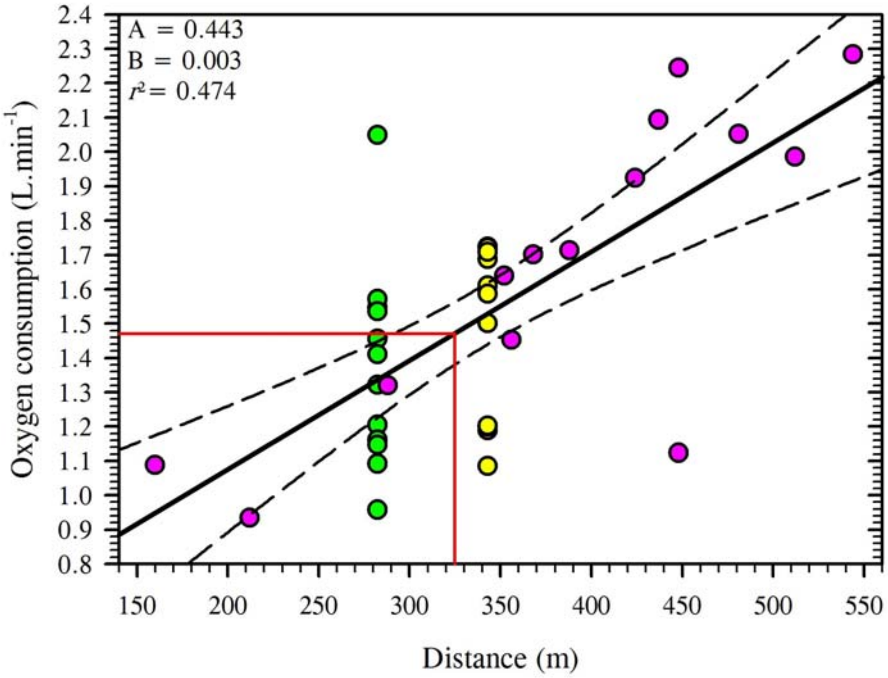

Figure 10: Predicting the walking speed required to elicit an oxygen cost of 1.47 L.min^{-1} during a 5-min hazmat incident simulation. Individual scores are presented (N=12) for the three stepping rates (fast = purple, medium [60%] = yellow, slow [30%] = green), with linear modelling (solid line) and 95% confidence intervals (dashed lines) applied to these data.

2.3.1.3 Simulation 1C: Loaded stair climb

Performances for the loaded stair climb (64 steps) are presented in Table 8. This was a criterion task (ventilation fan carriage) from which the stepping rates for the box stepping were determined (100%, 60%, 30%), with the aim of replicating the physiological demands of this task. It was anticipated that task performance would be powerfully correlated with performance on the hazmat simulation, but this was not the case (r^2=0.176).

2.3.1.4 Simulation 1D: Loaded box stepping (three stepping rates)

2.3.1.4.1 Physiological strain

Since one of the stepping rates used corresponded with each subject's maximal stair-climb speed, but only performed over a 5-min duration, then the capacity of this stepping rate to successfully replicate the metabolic steady state for this task was evaluated. This is illustrated in Figure 11, with steady states evident for both oxygen consumption and heart rate. The three stepping rates used in this activity were 61 (fast: range: 45-106), 36 (range: 23-54) and 18 steps.min^{-1} (range: 11-27), and the corresponding physiological responses for each subject are also shown in Table 8.

Table 8: Stair climb performance times and physiological responses to box stepping performed at three rates referenced to the stair climb performance.

Subject	Time (s)	Heart rate (b.min^{-1})			Oxygen consumption (L.min^{-1})		
		Slow	Medium	Fast	Slow	Medium	Fast
S1	82	116	124	118	0.82	1.02	1.8
S2	72	89	101	103	1	1.38	1.06
S3	67	148	166	147	0.94	1.21	1.36
S4	45	109	121	141	0.97	1.11	2.29
S5	50	106	113	142	0.88	1.12	1.78
S6	53	116	128	134	1.01	1.23	1.57
S7	57	117	124	146	0.95	1.41	2.12
S8	53	164	172	173	0.84	1.07	1.8
S9	96	136	144	143	1.19	1.21	1.24
S10	56	90	97	108	1.03	1.38	1.6
S11	106	---	---	171	---	---	1.21
S12	86	89	91	127	0.96	1.07	1.11
S13	85	92	96	110	0.94	1.24	1.56
S14	85	124	142	165	0.91	1.2	1.19
Mean	70.9	115	125	138	0.96	1.21	1.55
SD	19.2	23	26	23	0.09	0.13	0.38

2.3.1.4.2 Predictive power of the loaded box stepping task

The predictive utility of the 5-min maximal box stepping simulation was evaluated by comparing the oxygen cost of this test with the maximal performance on the stair-climb simulation. This comparison is illustrated in Figure 12, and it reveals that 48% of the variance in the stair-climb performance can be explained on the basis of the oxygen consumption observations. It was therefore considered that this simulation possessed a good predictive relationship with the criterion task.

The association between the cardiorespiratory demands of the criterion fire-fighting simulation and box-stepping performance was now investigated. In a previous characterisation of this criterion activity, the mean absolute oxygen cost was found to be 1.49 L.min^{-1}, with a 95% probability of this value falling between 1.23-1.75 L.min^{-1} (Taylor *et al.*, 2012b). Thus, to minimise the number of false negative outcomes during screening, the lower end of this range was used to predict the stepping rate for this task.

This is illustrated in Figure 13, revealing a very strong predictive power for field research (approximately 70%), and from which it was determined that the box stepping rate should be 11 stepping cycles per minute.

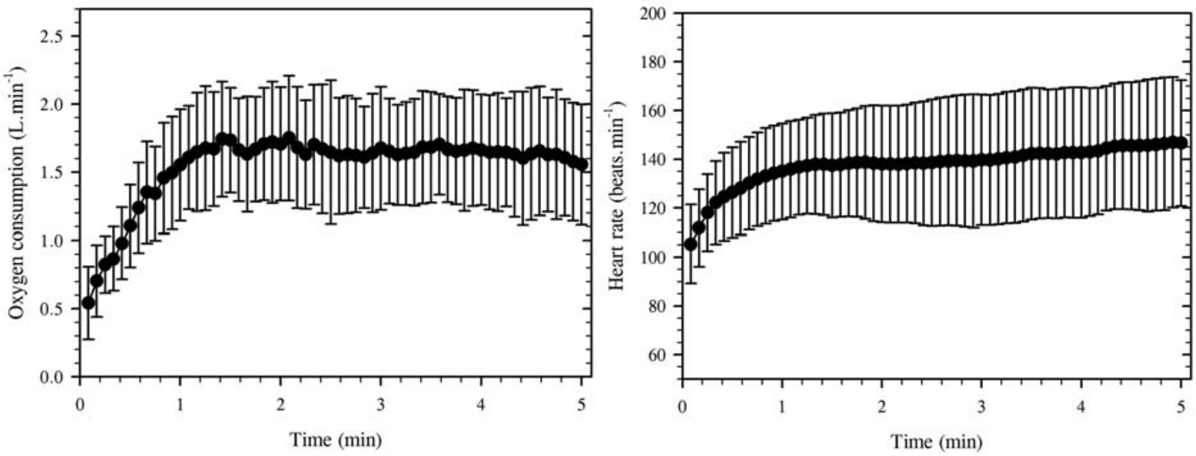

Figure 11: Oxygen consumption and heart rate data during box stepping activity performed at a rate matching the maximal stair-climb pace. Data are means with standard errors of the means (N=14).

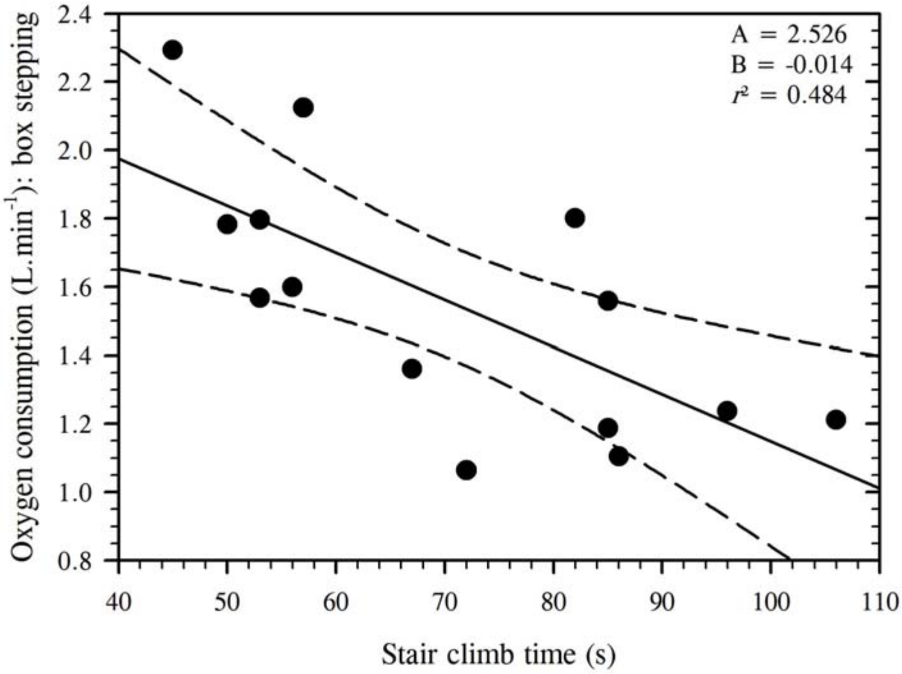

Figure 12: The correlation between performance time on the loaded stair climb task and the oxygen cost of a box-stepping activity (5 min) performed at an equivalent speed. Individual scores are presented with linear modelling and 95% confidence intervals (N=14).

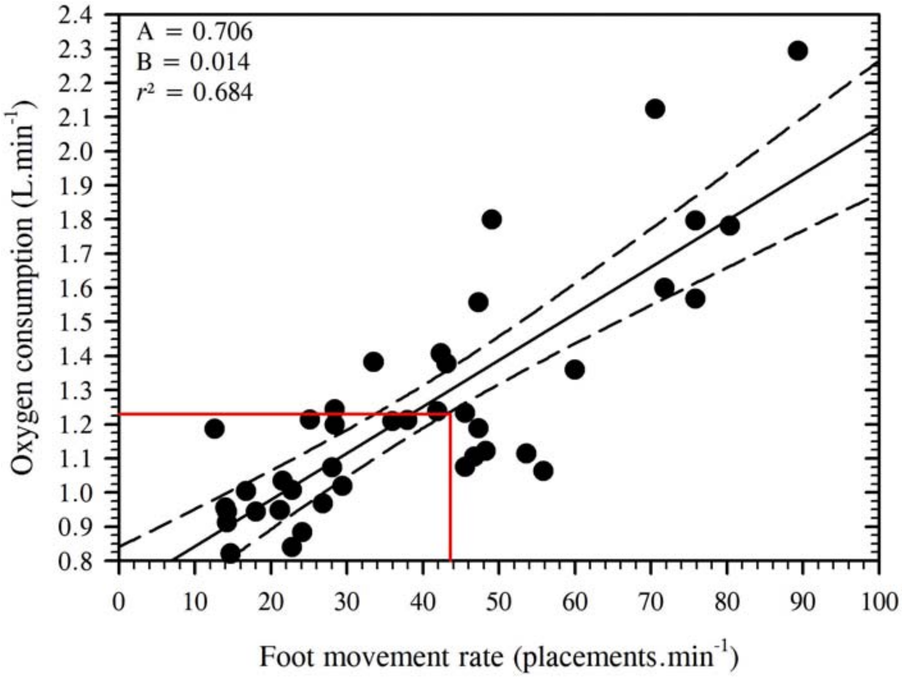

Figure 13: Predicting the box-stepping rate that would elicit an oxygen cost of 1.23 L.min^{-1} during a 5-min ventilation-fan carry simulation. Individual data are presented (N=14) with linear modelling (solid line) and 95% confidence intervals (dashed lines). Four foot movements (placements) are required to complete one stepping cycle (upwards and downwards).

2.3.2 Overhead pushing and holding tasks

2.3.2.1 Simulation 2A: Hydraulic shears, static-hold simulation

2.3.2.1.1 Physiological strain

Steady-state heart rates and oxygen consumption data for the static-hold simulation are summarised in Figure 14. It is clear that the oxygen cost was greater ($P<0.05$) when this activity was performed below knee height, when compared with the other two postures. This is attributable to the change in muscle recruitment demanded in this posture, which now relies much more heavily upon the large thigh muscles to hold both the body mass and the load in this semi-squat position. Conversely, the heart rate response increased as the height of the static hold was elevated. With this type of task, there is a greater activation of smaller muscle groups as the mass is raised, and this is invariably associated with a significant elevation in heart rate. In this simulation, the heart rate for the highest position was significantly greater than observed for the low-hold position ($P<0.05$). Table 9 presents data obtained from trials during which the duration of each static hold was modified, whilst the hold-rest cycle was maintained at 60 s.

The mean absolute oxygen cost observed when operational firefighters performed this task was 1.25 L.min^{-1}, with the 95% confidence interval being 1.15-1.35 L.min^{-1} (Taylor *et al.*, 2012b). Thus, the cardiorespiratory demand of this partial replication did not match the demand observed in operational firefighters. The reason for this was that the current task was focussed only on one aspect of the vehicle extraction (the static-hold phase with the

hydraulic tools), and excluded all other parts of the activity. Thus, whilst the strength demand was high for this holding tasks, it had a much lower metabolic demand.

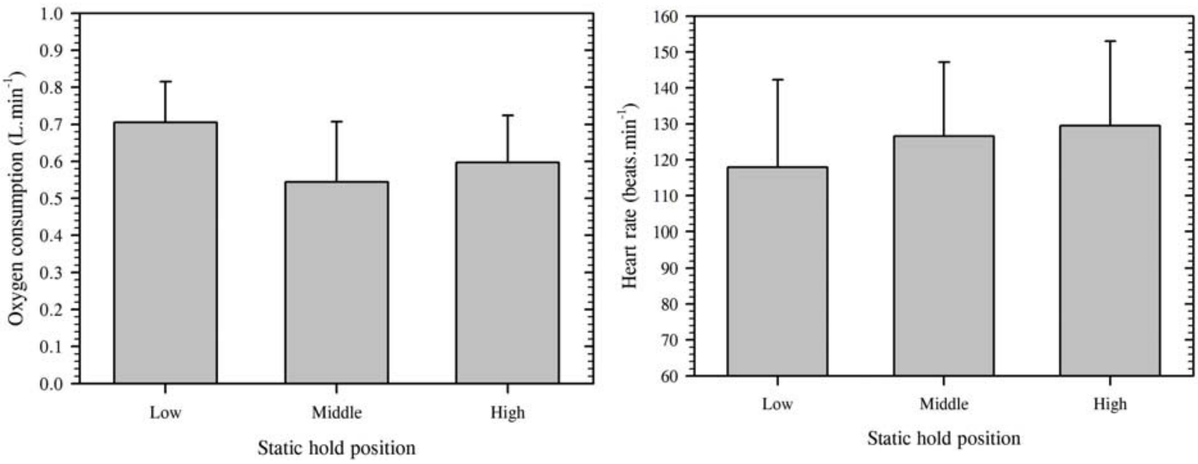

Figure 14: Oxygen consumption and heart rate data during a simulated static hold of the hydraulic shears at each of three relative heights: below the knees (low), between the knees and shoulders (middle) and at eye level (high). Data are means with standards errors of the means (*N*=14).

Table 9: Cardiovascular and metabolic strain of static-hold tasks performed in 60-s cycles, with alternating holding and resting stages completed within each cycle. The hold duration increased (20-40 s) while the corresponding recovery stages decreased. Data are averages obtained across three holding heights: below the knees, between knees and shoulders, at eye level.

Subject	Heart rate ($b.min^{-1}$)			Oxygen consumption ($L.min^{-1}$)		
	20 s	30 s	40 s	20 s	30 s	40 s
S1	111	116	114	0.66	0.59	0.59
S2	87	85	86	0.63	0.6	0.62
S3	136	148	141	0.64	0.7	0.75
S4	95	89	101	0.65	0.33	0.56
S5	95	94	106	0.41	0.38	0.53
S6	101	99	102	0.58	0.55	0.58
S7	100	102	116	0.5	0.49	0.67
S8	148	155	156	0.76	0.69	0.8
S9	131	125	135	0.74	0.72	0.84
S10	85	86	87	0.63	0.64	0.71

S11	137	141	155	0.73	0.63	0.72
S12	87	96	89	0.53	0.55	0.5
S13	77	86	88	0.64	0.71	0.72
S14	103	105	126	0.53	0.51	0.59
Mean	107	109	114*†	0.62	0.58*	0.66*†
SD	22	24	25	0.1	0.12	0.1

Notes: * = significantly different from values obtained during the 20-s holding trial ($P < 0.05$). † = significantly different from values obtained during the 30-s holding trial ($P < 0.05$).

2.3.2.2 Simulation 2B: Ladder raise - task performance

Table 10 summarises the performance times of all subjects during the ladder raise task.

2.3.2.3 Simulation 2C: Box lift and place - maximal load

The aim of using the box lift and place task was to evaluate its capacity to predict performance on the ladder-raise evaluation. The logic of this was that the former assessment has been used in many workplace assessments, and, if it could similarly be shown to possess sufficient predictive power, it could be performed instead of the ladder-raise assessment, since this task relies on two people, and the assistance rendered by the second individual would influence the performance score. Data for the box lift test are presented in Table 10, and the predictive capacity of these data is summarised in Figure 15. It is clear, at least with respect to the current data on 14 individuals, that the predictive power of this assessment is too low to warrant its inclusion in a Physical Aptitude Test at this time.

Table 10: Maximal performance for the ladder-raise assessment, the maximal box lift and place task and the maximal endurance box lift and place with 75% of the maximal load (mean load 24.7 kg [SD 7.5]).

Subject	Total time (s)	Rise time (s)	Lower time (s)	Maximal lift (kg)	Endurance lift (repetitions)
S1	30	13	17	35.8	36
S2	35.2	14.2	21	30	22
S3	43.9	15.1	28.8	29.4	10
S4	27.2	15	12.2	37.5	6
S5	29.8	10.6	19.2	35.9	12
S6	50.5	20	30.5	30	12
S7	23.5	11.5	12	40	14
S8	42	19	23	41.3	6

S9	104	47	57	25	11
S10	22.2	9.5	12.7	40	16
S11	53	22	31	15.1	13
S12	36	17	19	33.5	13
S13	43	20	23	51.9	9
S14	60	28	32	15.1	10
Mean	42.9	18.7	24.2	32.9	14
SD	20.9	9.6	11.7	10	8

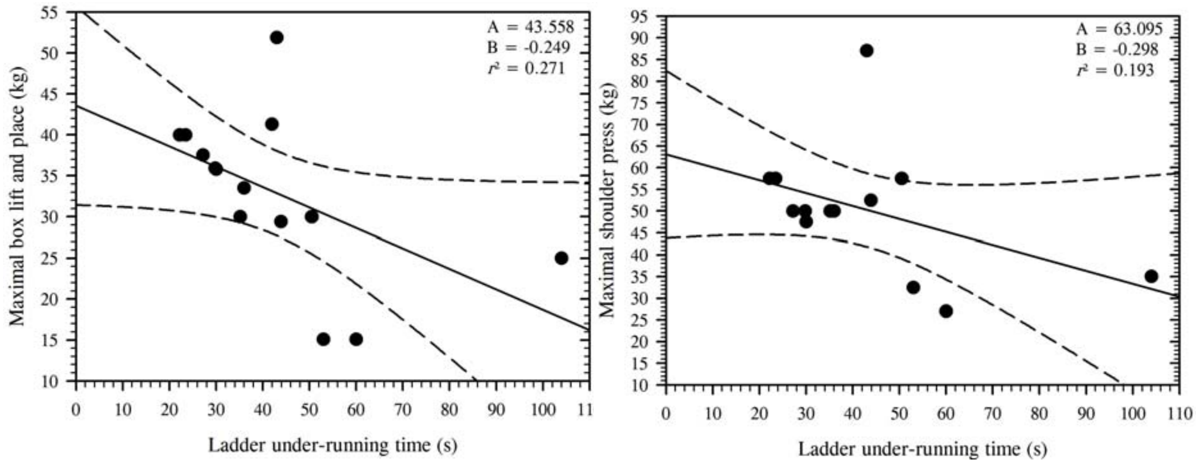

Figure 15: Predicting ladder raise performance from a maximal box lift and place task (left) and a maximal shoulder press test (right). Individual scores are presented with linear modelling and 95% confidence intervals (N=13).

2.3.2.4 Simulation 2D: Box lift and place - endurance

The results from this test are presented in Table 10. This activity also failed to provide a useful prediction of the ladder-raise performance.

2.3.2.5 Simulation 2E: One-repetition maximal shoulder press

The results from the maximal shoulder press test are presented in Figure 15, and these similarly revealed that this test failed to provide a useful prediction of the ladder-raise performance.

2.3.3 Cardiorespiratory dragging tasks

2.3.3.1 Simulation 3A: Hose-drag replication

This trial was performed as a 6-min maximal effort, with time-series data shown in Figure 16. The average distance covered was 596.6 m (SD 69.4), and this equates with an average walking speed of 1.66 m.s^{-1}.

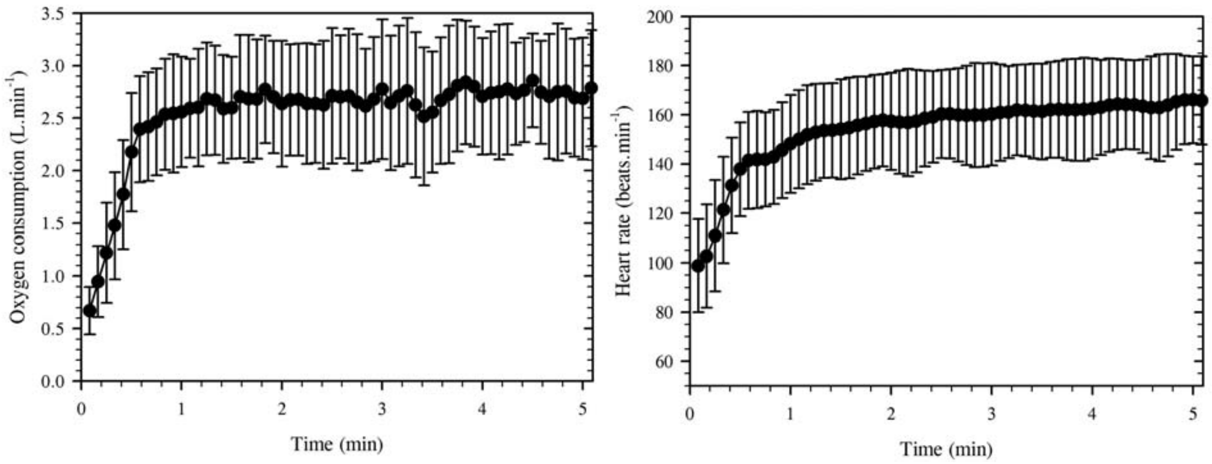

Figure 16: Oxygen consumption and heart rate data during a hose-drag simulation performed at maximal walking speed. Data are means with standard errors of the means ($N=14$).

2.3.3.2 Simulation 3B: Hose-drag replication (two walking speeds)

2.3.3.2.1 Physiological strain

The hose-drag task was performed over 60 m, with participants performing this activity at both slow (0.88 $m.s^{-1}$) and medium (0.93 $m.s^{-1}$) walking speeds and covering walking distances of 317 m and 336 m respectively. Physiological data from these trials are contained within Table 11. When this criterion task was observed in the field (*i.e.* the bushfire hose-drag simulation), the mean absolute oxygen cost was 1.63 $L.min^{-1}$ (Taylor *et al.*, 2012b). This sits between the values obtained at each of these walking speeds. Moreover, the higher of these values is not too far below that observed for the hose drag when ascending stairs: 1.97 $L.min^{-1}$ (Taylor *et al.*, 2012b). However, this range was absolutely covered within the maximal effort hose-drag task (6 min), and these data are illustrated in Figures 16 and 17.

Table 11: Cardiovascular strain and the metabolic cost of hose dragging performed at two walking speeds: slow (0.88 $m.s^{-1}$), medium (0.93 $m.s^{-1}$).

Subject	**Heart rate ($b.min^{-1}$)**		**Oxygen consumption ($L.min^{-1}$)**	
	Slow	Medium	Slow	Medium
S1	117	118	1.823	1.73
S2	111	115	1.659	1.776
S3	166	182	1.867	1.935
S4	118	135	1.579	1.951
S5	143	148	1.461	1.761

S6	126	136	1.331	1.661
S7	146	152	1.433	1.578
S8	165	160	1.534	1.764
S9	144	147	1.461	1.728
S10	117	124	1.775	1.986
S11	155	169	1.638	1.817
S12	106	113	1.177	1.33
S13	106	117	1.458	1.902
S14	148	152	1.417	1.631
Mean	133	141	1.544	1.754
SD	21	21	0.194	0.173

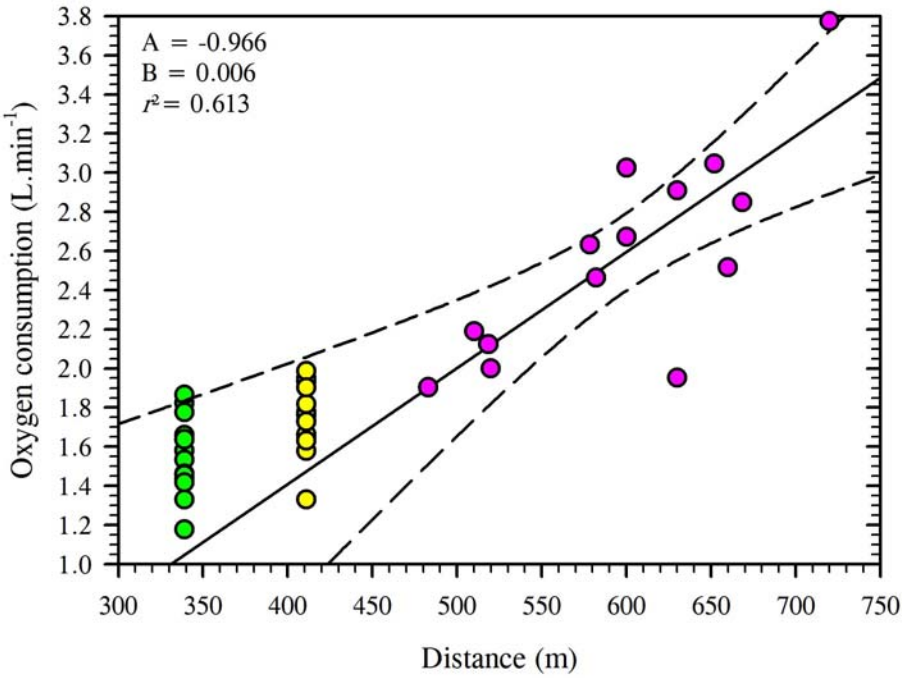

Figure 17: Predicting the oxygen cost of dragging a charged fire hose from distance covered. Individual scores (N=13) are presented (0.88 m.s^{-1} = green, 0.93 m.s^{-1} = yellow, maximal speeds = purple) with linear modelling (solid lines) and 95% confidence intervals (dashed lines).

2.4 CONCLUSIONS

The Research Team proposed that an individual who could satisfactorily complete all of the criterion activities within each class description identified within Table 2 of this report, would possess the minimal physiological attributes necessary to perform the tasks of contemporary fire-fighting within Fire & Rescue NSW (Groeller *et al.*, 2012a, 2012b).

Within this in mind, the isolated task analyses presented within this section of the report were performed such that task replications would be reflective of the demands of contemporary fire fighting. However, the Research Team was also seeking opportunities to maximise testing efficiency, either by performing more easily administered predictive tests instead of task simulations, or by eliminating tests that appeared to be evaluating very similar physiological attributes.

2.4.1 Recommendation one

It is recommended that the two single-sided, load-carriage tasks (the hazmat simulation and loaded box stepping) be considered as necessary items within the Physical Aptitude Test, since the correlation between these tasks was poor. This indicates that these tasks were sufficiently different in their physiological demands, negating the possibility that one task may be used as a substitution for the other.

2.4.2 Recommendation two

Significant upper-body muscular demands were associated with the continuous performance of the single-sided, load-carriage tasks. Since hazmat work is typically interspersed with unloaded activity, and since the ventilation fan carry is a relatively brief task, then it is recommended that the 5-min duration used for these tasks within this evaluation be reduced when performed within the Physical Aptitude Test.

2.4.3 Recommendation three

With respect to the ladder-raising activity, the Research Team had not considered this to be an appropriate task for assessing new recruits due to both the skill component and the possible bias introduced when a second person is used to assist task performance. However, none of the anticipated substitute tasks evaluated in these trials provided meaningful predictive information to warrant their substitution into the Physical Aptitude Test. This could have been due to problems with the sample size and the lack of familiarity of the current subjects with this task. Accordingly, it is recommended that FRNSW undertake an additional exploratory investigation of this critical task. Such an investigation should establish whether or not a satisfactory correlation exists between the ladder-raise task and a strength-based simulation task. A study of this nature is most appropriately conducted on recruits, and with FRNSW content experts used to determine objective standards for satisfactory performance. It is suggested that this work be performed as a maximal lift from the ground to full elbow extension, replicating the ladder-raise movement pattern. It is further suggested that this can be performed using the resistance loader illustrated in Figure 6. These standards may then be correlated with a strength-based simulation of this task.

2.4.4 Recommendation four

Due to the high face validity of the static-hold activity, it is recommended that it be included in the Physical Aptitude Test. Whilst the oxygen demand of this simulation was lower than observed within the field, due to extracting just the holding phase, it is anticipated that this would be elevated, and more reflective of operational performance, should this task be performed following tasks that have greater cardiorespiratory strain. It is therefore recommended that this activity should occur following the single-sided, load-carriage tasks.

2.4.5 Recommendation five

The cardiorespiratory demand observed during the hose-drag simulation was similar to that observed for the bushfire simulation performed by operational firefighters (Taylor *et al.*, 2012b). For this task to be used within a Physical Aptitude Test, it would need to be performed as a repeated dragging task over a restricted course, as it was above. In addition, the metabolic threshold for recruit screening should not be set higher than the level deemed necessary to successfully perform this task under operational conditions. Accordingly, the recommended test work rate for this task was calculated to match the lower 95% confidence interval observed in Research Phase Two (1.41 $L.min^{-1}$; Taylor *et al.*, 2012b).

2.4.6 Verification and approval of the staged screening tests

These screening tests, in the form of a proposed (prototype) Physical Aptitude Test, were submitted to the Project Management Team for consideration, endorsement and validation. Approval to progress to the final Research Phase (development of physiological employment standards for firefighters) was also sought, and this was granted at the Project Management Team meeting held on November 19th (2012: Appendix Six).

SECTION 3: PILOT TESTING OF THE PHYSICAL APTITUDE TEST (CIRCUIT) USING LAY PERSONNEL

3.1 INTRODUCTION

3.1.1 Research aim

The aim of this stage of the project was to assemble the individual test items into a single circuit test that would then represent the prototype for the Physical Aptitude Test. Following this assembly, the Research Team aimed to administer the test circuit to a small group of non-firefighters (lay personnel), and in so doing, provide an opportunity to identify problems within the test and to implement immediate solutions. This testing was undertaken in the presence of representatives from Fire & Rescue NSW, who also acted as subjects during pilot test, although no data were collected from these individuals.

3.2 METHODS

3.2.1 Experimental subjects

Eleven male and three female University staff and students (22-59 years old: Table 12) participated, with individuals of tall, medium and short stature again being recruited. All subjects provided written, informed consent and completed a screening questionnaire prior to commencing procedures approved by the Human Research Ethics Committee (University of Wollongong). This screening procedure was aimed at identifying and eliminating individuals for whom the performance of these simulations might be considered an unacceptable health risk.

Table 12: Characteristics of the experimental subjects.

Subject	Gender	Age (y)	Height (cm)	Body mass (kg)
S1	Male	59	184	75
S2	Male	46	189	95
S3	Male	35	172.5	78
S4	Male	24	172.7	82.1
S5	Male	28	181	69
S6	Male	23	173	65.5
S7	Male	32	180	78.4
S8	Male	25	179	69.5
S9	Female	32	175.3	69
S10	Male	40	177	73.5
S11	Female	34	157	56.3
S12	Male	24	174	66.6
S13	Male	22	190	93.7

Subject	Gender	Age (y)	Height (cm)	Body mass (kg)
S14	Female	23	160	61.5
Mean		32	176	73.8
SD		10.6	9.3	11.09

3.2.2 Physiological measurements

3.2.2.1 Oxygen consumption and the data acquisition system

Oxygen consumption was measured using a portable open-circuit, expired gas analysis and ventilation system (Figure 1: Metamax 3B, Cortex Biophysik, Leipzig, Germany: mass = 1.82 kg). This involved separate determination of minute ventilation (turbine), and expired oxygen (electro-chemical cell) and carbon dioxide concentrations (infra-red). Data were recorded on a breath-by-breath basis and reported at 5-s intervals. Equipment was calibrated at the start of each test day, with calibration verification performed throughout each day.

3.2.2.2 Heart rate

Heart rates were monitored continuously from ventricular depolarisation (Polar Electro Sports Tester, Kempele, Finland). These monitors were integrated into the data acquisition system for measuring oxygen consumption, with both sets of data being simultaneously recorded. These data were sampled on a breath-by-breath basis.

3.2.3 Overview of the Physical Aptitude Test

The aim of this testing was to record the total and fractional times taken by each individual to complete the entire Physical Aptitude Test and its six separate components. Subjects were instructed that the test was to be performed as quickly as possible (at walking speed), with subjects chosing their own speed, but with the understanding that the test may last 20-30 min. Running was not permitted, and subjects were instructed that they could stop to rest at any time within or between tasks, and for any duration, except during the static-hold simulations. A member of the Research Team walked through the test with each subject, providing technical advice where necessary, and ensuring that walking speeds were maintained and that the test was performed as safely as possible. The test was performed in the following order to reflect operational sequences and intensities that may be realistically encountered during an incident:

- ***Test item one:*** Unilateral load carriage (distance: 195 m, load: 26 kg): hazmat simulation
- ***Test item two:*** Stepping with a unilateral load (36 steps, 17.5 kg): ventilation fan carriage simulation
- ***Test item three:*** Static hold (load: 19.5 kg, duration: 3 min): motor-vehicle rescue simulation
- ***Test item four:*** Charged hose drag (distance: 150 m): bushfire simulation
- ***Test item five:*** Fire attack (height restriction: 1.25 m, distance: 30 m)
- ***Test item six:*** Firefighter rescue (height restriction: 1.25 m, distance: 10 m).

3.2.3.1 Test items one and two: Unilateral carrying tasks

Two unilateral, load-carriage tasks were performed. The first activity was designed to

replicate the physiological demands encountered at a hazmat incident. This was achieved using a 26-kg mass carried in one hand (subjects could change hands at any time). This mass was chosen since it was equal to 50% of the heaviest object carried during the two-man, loaded carriage component of the hazmat simulation (Taylor *et al.*, 2012b). The total carriage distance was set at 195 m because carrying this mass for 3 min at the target oxygen consumption for this simulation should result in participants covering this distance. Thus, the distance was set using the metabolic screening threshold derived from the previous Section; the lower 95% confidence interval observed in Research Phase Two (Taylor *et al.*, 2012b). A 3-min, rather than 5-min task was used since two single-sided, load-carriage tasks were to be evaluated, and this would reduce the muscular fatigue whilst still permitting each attribute to be assessed. Given this derivation, then an acceptable performance time for the first load-carriage task should be 3 min. However, participants were allowed to choose their own walking pace, so the time to complete this task would vary among subjects. To reduce the space required for the test, this distance was covered by participants completing 6.25 laps of a 30-m course.

The second load-carriage task was aimed at simulating the carriage of a ventilation fan up several flights of stairs. This is typically a two-person task, so the load was set at 17.5 kg, or half the mass of the ventilation fan. The total number of steps to be completed was 36 (3 min), with each step height being that of a standard building (26 cm). The number of steps to be completed was chosen because a stepping cycle rate of 11.25 steps.min^{-1} (44 placements.min^{-1}) with this load was found to equal the lower 95% confidence limit of the metabolic demand (metabolic screening threshold) during the ventilation fan carry performed by operational firefighters (Taylor *et al.*, 2012b).

3.2.3.2 Test item three: Static-hold task

To replicate some critical demands of the motor-vehicle rescue simulation, a static-hold activity was included. This was positioned after two tasks that would elicit a significant metabolic demand so that it would better reflect operational demands. Each of the three holds was for a fixed duration (1 min), and was always performed over this time. That is, this task lasted 3 min for every participant. The load was set at 19.5 kg, and this was designed to match the mass and load distribution of the hydraulic tools used by firefighters (shears and spreaders). Each 60-s hold sequence consisted of a 40-s hold followed by a 20-s resting recovery, during which the load was lowered to the ground. To better reflect the use of hydraulic tools during a motor-vehicle rescue, three hold positions were used: below the knees, between the knees and shoulders, and between shoulder and eye level. Participants held the load at each of these heights, but on their preferred side, with just one sequence of these holds being performed in the following order: (1) hold between shoulder and eye level, (2) hold between the knees and shoulders, (3) hold below knee level.

3.2.3.3 Test item four: Cardiorespiratory hose-dragging task

Immediately following the static-hold, the bushfire hose-drag simulation was performed. This involved dragging a modified, 38-mm hose in a forward direction, whilst holding the branch (nozzle). Each drag was 30 m, which reflects the length of a standard 38-mm hose. On reaching this point, participants handed the hose to one of the Research Team and walked back (without the hose) to the staring point. The hose was returned to the participant

at this point, and the next hose drag commenced immediately. The 38-mm hose was cut to a length of 5.6 m, filled with sand (11 kg), and doubled over in the middle. An anchor point at the centre of the hose was attached to the reel and line resistance device (resistance loader: Figure 6). Thus, this 2.8-m length of hose acted as the interface between the subject and the loader. The resistive load that was applied during this task was determined during pilot testing. This involved dragging a fully charged 38-mm hose across concrete and grass surfaces, with the resistive force being measured and averaged across surfaces. This force was replicated through the combined influence of the hose mass and the resistance applied by the resistance loader. Since this assessment was replicating the demands of the bushfire hose drag, then the total distance covered during this task should be 300 m, consisting of 150 m of dragging 150 m of unloaded walking. Using the logic described above for determining the metabolic screening threshold, this distance was set because dragging this mass for 6 min at the target oxygen consumption (lower 95% confidence interval observed in Research Phase Two [Taylor *et al.*, 2012b]) should result in participants covering a total of 300 m. Thus, the acceptable performance time for this task would have been 6 min. However, participants were allowed to choose their own walking pace, so the time to complete this task would vary among subjects. To ensure the area required to perform these assessments was compact, participants completed this distance through 10 laps of a 30-m course; five loaded (150 m) and five unloaded.

3.2.3.4 Test item five: Fire-attack simulation
The simulated fire attack also involved dragging the same modified, 38-mm hose whilst holding the branch. A height restriction of 1.25 m was used, which covered a distance of 30 m (return trip unloaded). This task was performed in a single effort moving forwards and away from the resistance loader. The resistive force was the same as that used for the bushfire hose drag. By using the height restriction, the overall flexibility of the person was evaluated, since performance in this posture requires a good range of flexibility across several joints, and such a functional evaluation provides a more valid assessment of overall flexibility than do isolated range-of-motion tests (*e.g.* sit and reach test).

3.2.3.5 Test item six: Firefighter-rescue simulation
The final test item was the simulated rescue of a firefighter, which replicates the demands of dragging an 82-kg firefighter wearing full protective clothing and equipment (20 kg). Subjects performed a single effort dragging task, always moving away from the resistance loader, but now in a backwards direction, and over a distance of just 10 m. The resistive load associated with this activity was determined during pilot testing. A 102-kg mass (equal to an 82-kg firefighter wearing 20 kg of personal protective equipment) was dragged across carpet, tiles, grass and linoleum surfaces. The average resistive force was 550 N, and this was replicated using the resistance loader. The distance chosen (10 m) was to replicate the firefighter rescue simulation examined in Phase Two (Taylor *et al.*, 2012b). A height restriction of 1.25 m was used for this task (this was modified for subsequent tests: see Section Four), therefore providing another functional evaluation of overall flexibility, but now with a much greater load and with the participant moving backwards.

3.2.4 Data analysis
Within this section of the report, descriptive statistics including the means (averages), data

ranges (minimal and maximal) and 95% confidence intervals are reported.

3.3 RESULTS AND DISCUSSION

3.3.1 Overall Physical Aptitude Test performance

On average, the time taken to complete the Physical Aptitude Test was 14 min 12 s (range: 10 min 42 s to 22 min 0 s). All participants were exhausted at the end of the test, with most stopping to rest during the course of the test, and some requiring several rest breaks[8]. Raw oxygen consumption data for each subject are provided in Figure 18, and the overall test averages are summarised in Table 13. The physiological data in this Table are simple averages from the start to the end of the test, including rest periods, and show significant cardiorespiratory strain. For instance, the mean heart rate was 158 beats.min^{-1} (95% confidence interval: 150-165 beats.min^{-1}) while the average oxygen consumption was 2.45 L.min^{-1} (95% confidence interval: 2.27-2.64 L.min^{-1}).

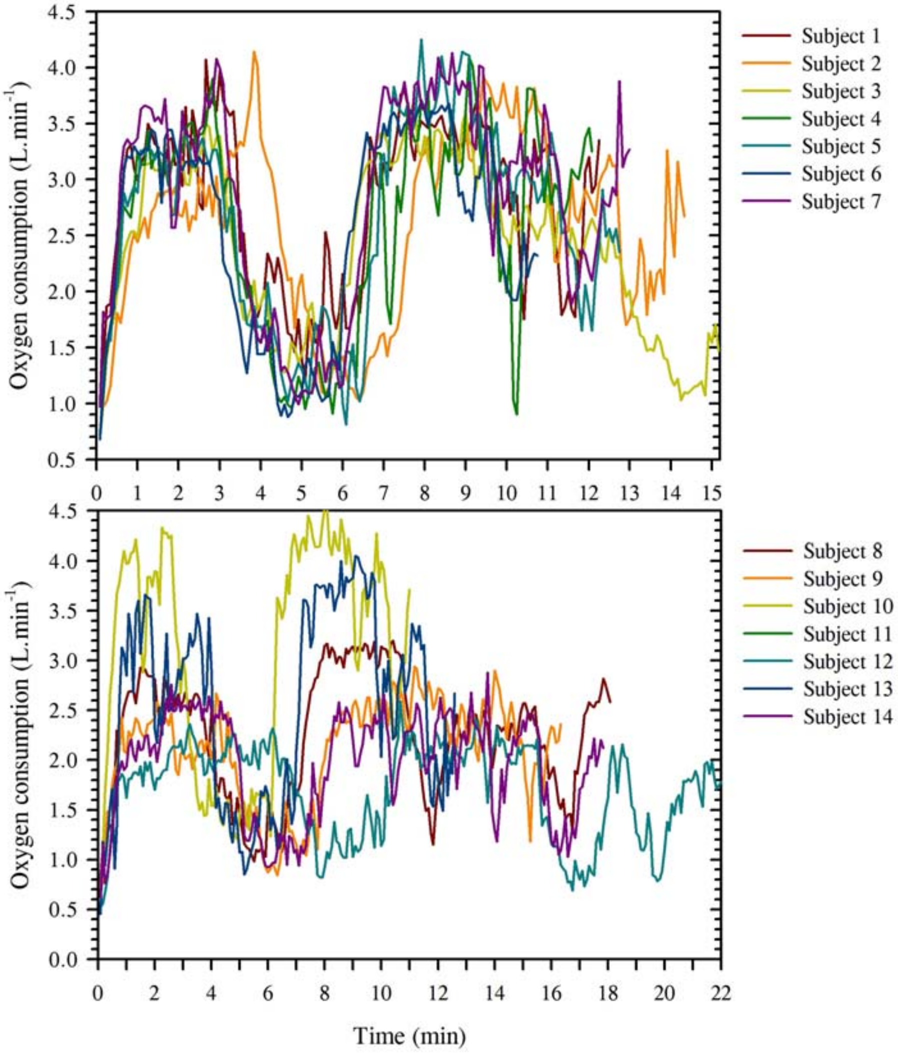

Figure 18: Raw oxygen consumption data of subjects performing the Physical Aptitude Test (N=14).

[8] Since all of these subjects were naive non-firefighters, then they had no idea how difficult the Physical Aptitude Test would be when it was performed as a continuous circuit test. Thus, the pace chosen by most was too high, and could not be sustained, resulting is many volitional pauses.

In the following sections, the discrete fire-fighting simulations have been extracted for closer examination. The start and completion points for each of these simulations was carefully recorded so that rest periods between simulations were not included. However, rests that occurred within a simulation are included within these summary data.

Table 13: Overall performance times for, and physiological strain during the completion of the entire Physical Aptitude Test.

Subject	Time (s)	Heart rate (beats.min^{-1})	Oxygen consumption (L.min^{-1})	Oxygen consumption (mL.kg^{-1}.min^{-1})	Oxygen consumption (mL.kg$^{-0.67}$.min^{-1})
S1	725	154	2.78	29.26	131.52
S2	857	132	2.58	22.45	107.49
S3	900	179	2.4	24.48	111.15
S4	720	154	2.57	25.12	115.63
S5	760	163	2.61	29.35	129.12
S6	640	146	2.54	29.76	129.18
S7	772	157	2.82	28.63	130.17
S8	1080	177	2.26	25.2	111.06
S9	975	163	2.1	23.54	103.56
S10	652	145	3.05	32.66	146.03
S11	1317	172	1.7	22.33	93.34
S12	730	153	2.43	28.03	122.21
S13	750	141	2.56	22.49	107.27
S14	1049	171	1.96	24	102.56
Mean	851.9	158	2.45	26.24	117.16
SD	192.8	14	0.36	3.31	14.56

Note: The mass specific oxygen consumption are provided in two forms to facilitate comparison with data collected on operational firefighters during the characterisation of the critical fire-fighting tasks (Taylor *et al.*, 2012b).

3.3.2 Test item one: Unilateral load carriage: Hazmat simulation

The first task performed within the Physical Aptitude Test was a unilateral, loaded carriage activity, which took 1 min 54 s to complete (range: 1 min 18 s to 3 min 30 s). The heart rate averaged 144 beats.min^{-1} (95% confidence interval: 136-153 beats.min^{-1}) while the mean oxygen consumption was 2.48 L.min^{-1} (95% confidence interval: 2.29-2.68 L.min^{-1}).

3.3.3 Test item two: Loaded stepping task: ventilation fan carry simulation

In the second simulation, the mass was reduced and the emphasis changed from being a simple horizontal ambulatory activity to one that included a vertical component. Because this now included the body mass of each subject, plus the average mass of the personal protective clothing and equipment (mean total mass: 93.8 kg (SD 11.1)), then the cardiorespiratory demand of this activity was significantly greater. The heart rate was 12% higher (18 beats.min^{-1}; $P<0.05$), averaging 162 beats.min^{-1} (95% confidence interval: 155-170 beats.min^{-1}), and the average oxygen consumption was 21% greater (0.53 L.min^{-1}; $P<0.05$): 3.01 L.min^{-1} (95% confidence interval: 2.74-3.28 L.min^{-1}). On average, this task was completed in 1 min 24 s (range: 1 min 0 s to 2 min 18 s).

3.3.4 Test item three: Static hold: holding hydraulic shears simulator

Participants immediately moved to the next simulation; the static hold. This was always performed for the same set duration, and was considered to represent a pass/fail barrier. Any individual unable to complete each of the three positional holds for the required duration would be deemed to have failed the overall Physical Aptitude Test, although such individuals should be allowed to continue the test. However, such people should still be required to pass this component as part of the complete test sequence. One of the female subjects failed this task due to an inability to sustain the holding position above shoulder height. As expected, the cardiovascular load was generally maintained during this activity (mean: 146 beats.min^{-1}, 95% confidence interval: 137-156 beats.min^{-1}). However, the oxygen cost was significantly lower ($P<0.05$), with subjects experiencing a 50% reduction in mean oxygen consumption from the loaded stepping task: 1.54 L.min^{-1} (95% confidence interval: 1.44-1.64 L.min^{-1}).

3.3.5 Test item four: Charged hose drag: bushfire hose-drag simulation

The fourth task was the longest simulation, and it was designed to replicate the physiological demands of dragging a charged hose through the bush. On average, this activity took 3 min 36 s to complete (range: 2 min 30 s to 5 min 48 s). The heart rate averaged 164 beats.min^{-1} (95% confidence interval: 155-173 beats.min^{-1}) while the mean oxygen consumption was 3.07 L.min^{-1} (95% confidence interval: 2.80-3.35 L.min^{-1}).

3.3.6 Test item five: Fire-attack simulation

The last two simulations were performed using a postural control limitation, such that a height restriction was imposed below which each participant was required to remain. This was aimed at replicating the postures that should be adopted by firefighters to stay below smoke layers and to avoid excessive heat during movement within a burning building. This was expected to significantly elevate musculoskeletal strain as well as exposing individuals with inadequate flexibility and ranges of joint movement. On average, this task was completed in 1 min 6 s to complete (range: 0 min 42 s to 2 min 24 s). Relative to the bushfire hose drag, the heart rate was 5% higher (9 beats.min^{-1}; $P<0.05$), now averaging 173 beats.min^{-1} (95% confidence interval: 167-178 beats.min^{-1}). It is clear that subjects were now working towards peak levels of cardiovascular function. However, since this task had its physiological emphasis changed towards that of a strength task, due to the postural constraint, then the average oxygen consumption was significantly reduced, with the average reduction being 0.42 L.min^{-1} ($P<0.05$), and the mean oxygen cost was 2.65 L.min^{-}

[1] (95% confidence interval: 2.38-2.92 $L.min^{-1}$).

3.3.7 Test item six: Firefighter rescue simulation

The final task was the firefighter rescue, also performed with a height restriction, but now with a greater resistive load, and with participants moving backwards. The rescue simulation took, on average, 54 s to complete (range: 18 s to 2 min 6 s). The heart rate averaged 166 $beats.min^{-1}$ (95% confidence interval: 161-171 $beats.min^{-1}$) while the mean oxygen consumption was 2.28 $L.min^{-1}$ (95% confidence interval: 2.01-2.54 $L.min^{-1}$).

3.4 CONCLUSION

Three recommendations are tendered for consideration prior to embarking in further testing.

3.4.1 Recommendation one:

When performing the complete prototype Physical Aptitude Test, participants experienced a significant level of cardiorespiratory strain and fatigue. Indeed, for some individuals, this necessitated prolonged periods of rest, most notably during the final three tasks. This was because this group of subjects was advised to perform the test in the fastest possible time. Thus, it is recommended that all participants should be advised that the final three tasks of the test are likely to promote the greatest level of cardiorespiratory strain, and that they should set a work tempo at the start of the test that would ensure that exhaustion did not occur before the final test item was completed. In addition, it is recommended that firefighters participating in the next testing phase be instructed to complete the test at an operational work intensity. This recommendation is designed to ensure that performance standards that may arise from such trials do not suffer from bias introduced by participants working at unnecessarily high work rates.

3.4.2 Recommendation two:

Technique clearly influenced test performance during the two hose-drag tasks and also during the simulated firefighter rescue. Since the Physical Aptitude Test is designed as a recruit screening aid, then test performance should not be significantly influenced by variations in performance technique, as this would introduce an unacceptable bias into the screening process. It is therefore recommended that all participants should be given sufficient opportunity to attempt these tasks using various dragging techniques, so that each individual will have an opportunity to select and (if necessary) learn the most appropriate technique for himself/herself. Whilst a test briefing will be required on each test day, it is advisable to have a range of information and technical advice available to test candidates prior to arriving at a test venue.

3.4.3 Recommendation three:

The height restriction of 1.25 m during the simulated firefighter rescue forced a number of participants to adopt postures that prevented a rapid response to this emergency, and increased the risk of injury. The observing members of Fire & Rescue NSW considered that, due to the critical nature of this task, this height restriction may have been somewhat unrealistic. It is therefore recommended that this height restriction be further considered, and possibly made less difficult (see Appendices Three and Four).

SECTION 4: FIELD TRIALS OF THE PHYSICAL APTITUDE TEST USING OPERATIONAL FIREFIGHTERS (VOLUNTEERS)

4.1 INTRODUCTION

A critical part in the development of screening tests is establishing the validity of the test. In this instance, one must determine how well the proposed Physical Aptitude Test can identify capable and robust potential firefighters. To this point, four forms of validity for the proposed Physical Aptitude Test have been addressed: construct validity[9], face validity[10], content validity[11] and criterion-related validity[12]. However, there is a special form of construct validity that has not yet been addressed: convergent validity[13]. Therefore, the emphasis of this Section was to evaluate how well the Physical Aptitude Test allows one to evaluate the job performance potential of prospective firefighters, or the job performance fitness of existing firefighters relative to the observed demands of the job (essential or critical tasks) when successfully performed by operational firefighters.

4.1.1 Research aims

In this Research Phase, operational firefighters (both permanent and retained) were exposed to the prototype circuit test that would be recommended to Fire & Rescue NSW as the Physical Aptitude Test. The purpose of this was to evaluate the capacity of the test to identify capable and robust firefighters, and in so doing provide an assessment of its ability to differentiate between the work performance fitness of a range of operational firefighters.

4.2 METHODS

4.2.1 Experimental subjects

Testing took part at the University, and at Fire Stations and centres in NSW (Alexandria, Blacktown, Campbelltown, Goulburn, Katoomba, Kirrawee, Queanbeyan). A total of 148 operational firefighters across ranks up to, and including Superintendent, completed the Physical Aptitude Test (age range 19-64 years: Table 14). Each participant provided written, informed consent and completed a screening questionnaire prior to commencing procedures approved by the Human Research Ethics Committee (University of Wollongong). This screening procedure was aimed at identifying and eliminating individuals for whom the performance of these simulations might be considered to be an unacceptable health risk. Two firefighters were identified with cardiovascular risk factors during this

---

9 Construct validity relates to the ability of a test to measure physiological attributes that are either related to, or that underpin job performance (*e.g.* strength and local-muscle endurance).

10 Face validity refers to the apparent similarity between a screening test (simulation) and a real activity.

11 Content validity relates to using fire-fighting activities or simulations within the Physical Aptitude Test, such that a subject-matter expert would support the test as a fair and reasonable approximation of fire-fighting tasks.

12 Criterion-related validity refers to the capacity of a test to reflect the physiological strain observed during fire-fighting simulations, and it establishes the predictive validity of the screening test.

13 Some constructs cannot easily be measured (*i.e.* non-quantifiable traits or attributes). For instance, while we can easily measure physiological attributes that are strongly related to work performance, we may be unable to quantify performance *per se*. However, one would expect that a Physical Aptitude Test with good construct validity would correlate well with work performance. When such a relationship exists, it is known as convergent validity.

screening, and they were not permitted to participate in this testing.

4.2.2 Measurements

4.2.2.1 Oxygen consumption and the data acquisition system

Oxygen consumption was measured using a portable open-circuit, expired gas analysis and ventilation system (Figure 1: Metamax 3B, Cortex Biophysik, Leipzig, Germany: mass = 1.82 kg). This involved separate determination of minute ventilation (turbine), and expired oxygen (electro-chemical cell) and carbon dioxide concentrations (infra-red). Data were recorded on a breath-by-breath basis and reported at 5-s intervals. Equipment was calibrated at the start of each test day, with calibration verification performed throughout each day.

Table 14: Characteristics of the subjects (means with standard deviations in parenthesis).

Classification	Age (y)	Experience (y)	Height (cm)	Mass (kg)
Overall	39.1 (9.1)	11.8 (8.7)	178.33 (6.82)	84.58 (11.97)
Permanent	40.1 (8.7)	13.0 (8.4)	178.57 (6.71)	92.96 (11.99)
Retained	34.7 (9.6)	5.8 (7.9)	177.19 (7.35)	84.31 (13.33)

4.2.2.2 Heart rate

Heart rates were monitored continuously from ventricular depolarisation (Polar Electro Sports Tester, Kempele, Finland). These monitors were integrated into the data acquisition system for measuring oxygen consumption, with both sets of data being simultaneously recorded. These data were sampled on a breath-by-breath basis.

4.2.2.3 Subjective evaluation of face validity

Immediately after completing the Physical Aptitude Test, each firefighter was asked to rate how well the test replicated the real-world demands of fire fighting. This was performed using a visual-analogue scale, with participants instructed to mark a vertical line through a 100-mm horizontal line, the left end of which indicated "very poor" whilst the opposite end was marked "very good" (Figure 19). This scale provided subjective ratings that could then be assigned a score from 1 ("very poor") to 100 ("very good"), by measuring the distance (mm) from the left-hand end to the point at which the vertical line was marked.

Visual analogue scale:

How well do you think this Physical Aptitude Test replicated the demands of fire fighting?

Please place a vertical line at some point along the horizontal line shown below, and between the two ends points ("very poor" and "very good").

Very poor — Very good

Figure 19: The visual-analogue scale used for firefighter evaluations of the face validity of the Physical Aptitude Test.

4.2.3 Overview of the Physical Aptitude Test

The aim of this testing was to record the time taken for each firefighter to complete the prototype Physical Aptitude Test, which was comprised of six test items. Firefighters were instructed to perform the test at the fastest possible operational pace (*i.e.* at walking speed), with each person setting his/her own task performance rate, but with the understanding that the test may last 20-30 min. Running was not permitted, and firefighters were instructed that they could stop to rest at any time within or between tasks, and for any duration, except during the static-hold simulation. A member of the Research Team walked through the test with each firefighter, providing technical advice where necessary, and ensuring that walking speeds were maintained and that the test was performed as safely as possible. The test was performed in the following order to reflect operational sequences and intensities that may realistically be encountered during an incident:

- ***Test item one:*** Unilateral load carriage (distance: 195 m, load: 26 kg): hazmat simulation
- ***Test item two:*** Stepping with a unilateral load (36 steps [26 cm each], 17.5 kg): ventilation fan carriage simulation
- ***Test item three:*** Static hold (load: 19.5 kg, duration: 3 min): motor-vehicle rescue simulation
- ***Test item four:*** Charged hose drag (distance: 150 m): bushfire simulation
- ***Test item five:*** Fire attack (height restriction: 1.25 m, distance: 30 m)
- ***Test item six:*** Firefighter rescue (height restriction: 1.55 m, distance: 10 m).

4.2.3.1 Test items one and two: Unilateral carrying tasks

Two unilateral, load-carriage tasks were performed. The first activity was designed to replicate the physiological demands encountered at a hazmat incident. This was achieved using a 26-kg mass carried in one hand (subjects could change hands at any time). This mass was chosen since it was equal to 50% of the heaviest object carried during the two-man, loaded carriage component of the hazmat simulation (Taylor *et al.*, 2012b). The total carriage distance was set at 195 m because carrying this mass for 3 min at the target oxygen consumption for this simulation should result in participants covering this distance. Thus, the distance was set using the metabolic screening threshold derived from the previous Section; the lower 95% confidence interval observed in Research Phase Two (Taylor *et al.*, 2012b). A 3-min, rather than 5-min task was used since two single-sided, load-carriage tasks were to be evaluated, and this would reduce the muscular fatigue whilst still permitting each attribute to be assessed. Given this derivation, then an acceptable performance time for the first load-carriage task should be 3 min. However, participants were allowed to choose their own walking pace, so the time to complete this task would vary among subjects. To reduce the space required for the test, this distance was covered by participants completing 6.25 laps of a 30-m course.

The second load-carriage task was aimed at simulating the carriage of a ventilation fan up several flights of stairs. This is typically a two-person task, so the load was set at 17.5 kg, or half the mass of the ventilation fan. The total number of steps to be completed was 36 (3 min), with each step height being that of a standard building (26 cm). The number of steps to be completed was chosen because a stepping cycle rate of 11.25 steps.min^{-1} (44 placements.min^{-1}) with this load was found to equal the lower 95% confidence limit of the

metabolic demand (metabolic screening threshold) during the ventilation fan carry performed by operational firefighters (Taylor *et al.*, 2012b).

4.2.3.2 Test item three: Static-hold task

To replicate some critical demands of the motor-vehicle rescue simulation, a static-hold activity was included. This was positioned after two tasks that would elicit a significant metabolic demand so that it would better reflect operational demands. Each of the three holds was for a fixed duration (1 min), and was always performed over this time. That is, this task lasted 3 min for every participant. The load was set at 19.5 kg, and this was designed to match the mass and load distribution of the hydraulic tools used by firefighters (shears and spreaders). Each 60-s hold sequence consisted of a 40-s hold followed by a 20-s resting recovery, during which the load was lowered to the ground. To better reflect the use of hydraulic tools during a motor-vehicle rescue, three hold positions were used: below the knees, between the knees and shoulders, and between shoulder and eye level. Participants held the load at each of these heights, but on their preferred side, with just one sequence of these holds being performed in the following order: (1) hold between shoulder and eye level, (2) hold between the knees and shoulders, (3) hold below knee level.

4.2.3.3 Test item four: Cardiorespiratory hose-dragging task

Immediately following the static-hold, the bushfire hose-drag simulation was performed. This involved dragging a modified, 38-mm hose in a forward direction, whilst holding the branch (nozzle). Each drag was 30 m, which reflects the length of a standard 38-mm hose. On reaching this point, participants handed the hose to one of the Research Team and walked back (without the hose) to the staring point. The hose was returned to the participant at this point, and the next hose drag commenced immediately. The 38-mm hose was cut to a length of 5.6 m, filled with sand (11 kg), and doubled over in the middle. An anchor point at the centre of the hose was attached to the reel and line resistance device (resistance loader: Figure 6). Thus, this 2.8-m length of hose acted as the interface between the subject and the loader. The resistive load that was applied during this task was determined during pilot testing. This involved dragging a fully charged 38-mm hose across concrete and grass surfaces, with the resistive force being measured and averaged across surfaces. This force was replicated through the combined influence of the hose mass and the resistance applied by the resistance loader (265 N of resistive force including an 11-kg hose mass). Since this assessment was replicating the demands of the bushfire hose drag, then the total distance covered during this task should be 300 m, consisting of 150 m of dragging 150 m of unloaded walking. Using the logic described above for determining the metabolic screening threshold, this distance was set because dragging this mass for 6 min at the target oxygen consumption (lower 95% confidence interval observed in Research Phase Two [Taylor *et al.*, 2012b]) should result in participants covering a total of 300 m. Thus, the acceptable performance time for this task would have been 6 min. However, participants were allowed to choose their own walking pace, so the time to complete this task would vary among subjects. To ensure the area required to perform these assessments was compact, participants completed this distance through 10 laps of a 30-m course; five loaded and five unloaded.

4.2.3.4 Test item five: Fire-attack simulation

The simulated fire attack also involved dragging the same modified, 38-mm hose whilst holding the branch. A height restriction of 1.25 m was used, which covered a distance of 30 m. This task was performed in a single effort moving forwards and away from the resistance loader, with an unloaded recovery. The resistive force (265 N of resistive force including an 11-kg hose mass) was the same as that used for the bushfire hose drag. By using the height restriction, the overall flexibility of the person was evaluated, since performance in this posture requires a good range of flexibility across several joints, and such a functional evaluation provides a more valid assessment of overall flexibility than do isolated range-of-motion tests (*e.g.* sit and reach test).

4.2.3.5 Test item six: Firefighter-rescue simulation

The final test item was the simulated rescue of a firefighter, which replicates the demands of dragging an 82-kg firefighter wearing full protective clothing and equipment (20 kg). Subjects performed a single effort dragging task, always moving away from the resistance loader, but now in a backwards direction, and over a distance of just 10 m. The resistive load associated with this activity was determined during pilot testing. A 102-kg mass (equal to an 82-kg firefighter wearing 20 kg of personal protective equipment) was dragged across carpet, tiles, grass and linoleum surfaces. The average resistive force was 550 N, and this was replicated using the resistance loader. The distance chosen (10 m) was to replicate the firefighter rescue simulation examined in Phase Two (Taylor *et al.*, 2012b). A height restriction of 1.55 m was used for this task (Appendices Three and Four), therefore providing another functional evaluation of overall flexibility, but now with a much greater load and with the participant moving backwards.

4.2.3.6 Supplementary test: Ladder-raise task

In Sections 2.3.2.2 to 2.3.2.5 of this document, it was reported that no task simulation had the capacity to predict an individual's ability to successfully raise the 10.5-m (49.5-kg) extension ladder from the ground to a vertical position. Therefore, most firefighters were also asked to perform this task in the following stages. Firstly, each firefighter performed a timed ladder raise. During this, an assistant firefighter provided support at the foot of the ladder, and ladder stabilisation only. No assistance was provided to raise the ladder.

Secondly, the ladder raise was replicated using the resistance loader. A 12.5-kg dumbbell was attached to the line and reel resistance loader, and a three-stage lift was performed: lifting the load from the ground to waist height, moving the load from waist to shoulder height and finally pressing from shoulder height to above the head with full elbow extension. This second lifting sequence was performed to closely reflect the lifting stages (actions) performed during the actual ladder lift, up to the point of the ladder under-run. Firefighters firstly performed this three-stage lift with minimal resistance applied to the reel, allowing for some familiarisation with the movement. The resistance was then progressively adjusted by a member of the Research Team until either the participant reached maximal resistive force on the reel, or was unable to obtain full elbow extension during the third stage of the lift. When testing finished, the final lever position on resistance loader was recorded.

4.2.4 Data analysis

Within this report, descriptive statistics including the means (averages), data ranges (minimal and maximal) and 95% confidence intervals are reported for each task within the Physical Aptitude Test. Two forms of correlation analysis (Pearson's product-moment and Spearman's rank-order correlation) were performed to investigate the predictive power of the simulated reel lift for the ladder-raise task.

4.3 RESULTS AND DISCUSSION

4.3.1 Overall Physical Aptitude Test performance

On average, and across all operational firefighters, the time taken to complete the Physical Aptitude Test (circuit) was 13 min 31 s (range: 9 min 53 s to 25 min 27 s). These performance times were only weakly related to the age of these firefighters (Figure 20) such that age accounted for less than 10% of the variance in test performance. It is noted that two firefighters failed the static-hold task (see below), so their data were excluded from the overall test performances, as well as the separate test-item analyses.

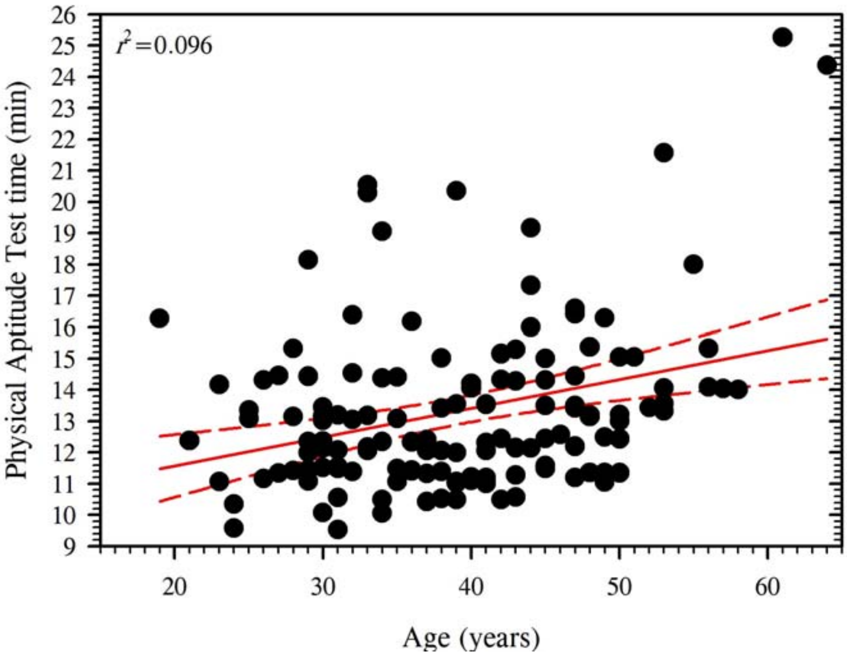

Figure 20: Physical Aptitude Test performance times sorted by firefighter age.

The overall performance time of the firefighters was similar to that recorded by the University subjects who participated in the pilot trials (Section 3), with the latter group being just 62 s slower (on average). However, the retained firefighters performed the Physical Aptitude Test significantly more slowly (15 min 10 s: range: 11 min 5 s to 20 min 29 s) than the permanent firefighters (13 min 10 s: range: 9 min 53 s to 25 min 27 s; $P<0.05$).

In the group of 18 permanent firefighters from whom oxygen consumption and heart rate data were simultaneously recorded, the mean absolute oxygen consumption was 2.36 $L.min^{-1}$ (95% probability range: 2.07-2.65 $L.min^{-1}$), while the mean heart rate was 161 $beats.min^{-1}$ (95% probability range: 156-166 $beats.min^{-1}$). These values were very close to the responses of the University participants (Section 3), differing by only 0.09 $L.min^{-1}$ and 3

beats.min^{-1}. These results provide strong evidence that the physiological impact of the Physical Aptitude Test was not skill dependent, and this is important information, since the test is initially aimed at recruit screening. We will now separately consider the performance of each test item within the complete test circuit.

4.3.2 Test items one and two: Unilateral carrying tasks

The mean time to complete the hazmat incident simulation was 2 min 14 s (range: 1 min 12 s to 3 min 55 s): permanent firefighters 2 min 12 s (range: 1 min 12 s to 3 min 55 s) and retained firefighters 2 min 26 s (range: 1 min 35 s to 3 min 10 s). As previously observed, the slower and less fit individuals required several short rests during the task. Interestingly, the average times of both the permanent and retained firefighters during this task were slower than those observed by the University subjects (Section Three: 1 min 54 s). This most likely reflects the inexperience of the last group, who invariably commenced the test with a work tempo that was far too high.

In the group of 18 firefighters on whom oxygen consumption and heart rate data were recorded, the mean absolute oxygen consumption was 2.25 L.min^{-1} (95% probability range: 1.77-2.73 L.min^{-1}) and the mean heart rate was 148 beats.min^{-1} (95% probability range: 142-154 beats.min^{-1}). The mean heart rate was 4 beats.min^{-1} higher than the University subjects, yet the mean absolute oxygen consumption was 0.23 L.min^{-1} lower. Both the firefighters and the University subjects displayed greater cardiorespiratory strain when compared to the field-based hazmat simulation on operational firefighters (Taylor *et al.*, 2012b). However, an average walking speed of 1.45 m.s^{-1} was observed during the current testing, and this was much quicker than required to elicit the necessary cardiorespiratory demand observed during the hazmat task (0.88 m.s^{-1}). Such an elevation in effort is expected during testing conditions, and it is unavoidable. However, it places a restriction on the utility of such data for determining performance thresholds (standards).

The mean performance time of the loaded stepping (ventilation-fan carriage simulation) task was 1 min 20 s (range: 0 min 18 s to 3 min 36 s): permanent firefighters recorded an average time of 1 min 18 s (range: 0 min 18 s to 2 min 0 s) whilst the retained firefighters averaged 1 min 24 s (range: 1 min 0 s to 2 min 0 s).

For this test item, the mean heart rate was 162 beats.min^{-1} (95% probability range: 156-168 beats.min^{-1}), which was only 5 beats.min^{-1} higher on than recorded during the operational simulations (Taylor *et al.*, 2012b). However, the average oxygen consumption was 2.73 L.min^{-1} (95% probability range: 2.43-3.03 L.min^{-1}), and this was almost double that observed during the simulated ventilation fan carry (Taylor *et al.*, 2012b). This difference is a simple reflection of the performance times which were, on average, less than half that required to elicit the desired oxygen consumption (Figure 13).

4.3.3 Test item three: Static-hold task

This activity was designed to replicate a motor-vehicle rescue. Two operational firefighters (1 permanent female and 1 retained female) failed this task since they were unable to either reach, or to maintain, the static hold at eye height. Since this task was performed for a fixed duration, only the oxygen consumption and heart rate data were recorded: mean oxygen

consumption 1.50 $L.min^{-1}$ (95% probability range: 1.21-1.69 $L.min^{-1}$); heart rate 154 $beats.min^{-1}$ (95% probability range: 148-160 $beats.min^{-1}$). These represent a 45% reduction in oxygen demand following the unilateral stepping task. However, despite the decreased work output, the average heart rate remained relatively unchanged (8 $beats.min^{-1}$ lower). This is quite normal during static-strength tasks which invariably elicit high heart rates, but this is particularly evident when one commences such a task with an already elevated heart rate. During the isolated task analysis, the oxygen demand (0.66 $L.min^{-1}$) and cardiovascular strain (114 $beats.min^{-1}$) associated with this task were both much lower. However, when this task was performed in the Physical Aptitude Test, these values were much closer, being only 0.25 $L.min^{-1}$ and 17 $beats.min^{-1}$ higher (respectively) than the operational simulations (Taylor *et al.*, 2012b)

4.3.4 Test item four: Cardiorespiratory hose-dragging task

It was anticipated that this simulation would be associated with a wide range of performance times, and this expectation was realised: overall average 3 min 57 s (range: 1 min 36 s to 8 min 3 s); permanent firefighters 3 min 50 s (range: 1 min 36 s to 8 min 3 s); retained firefighters 4 min 27 s (range: 2 min 50 s to 6 min 28 s). On average, this task was completed by these firefighters at speeds approximately 30% quicker than was required to meet the oxygen consumption (screening threshold) and the cardiovascular demand observed during the bushfire simulations (1.41 $L.min^{-1}$ and 143 $beats.min^{-1}$; Taylor *et al.*, 2012b). It was therefore not surprising to see a significantly higher average oxygen consumption: 2.90 $L.min^{-1}$ (95% probability range: 2.52-3.28 $L.min^{-1}$) and a mean heart rate of 169 $beats.min^{-1}$ (95% probability range: 164-174 $beats.min^{-1}$).

4.3.5 Test item five: Fire-attack simulation

The average time to complete this simulation was 49 s (95% portability range: 44 s to 54 s). Firefighters completed this task, on average, 17 s quicker than the University subjects, but the permanent firefighters were only 3 s faster than the retained firefighters. Whilst the completion time for this task was short, the very wide performance ranges reflect fatigue that accumulated during this, and the previous activity. In fact, one firefighter chose to terminate the test after completing the fire-attack simulation.

Indeed, this simulation was observed to be the most physically demanding of all tasks within the Physical Aptitude Test, and many firefighters dramatically reduced their walking speed to facilitate recovery during the unloaded walking stage back to the resistance loader. Thus, the average time for this recovery was 53 s (95% probability range 48 s to 58 s). The mean oxygen consumption was 2.69 $L.min^{-1}$ (95% probability range: 2.26-3.12 $L.min^{-1}$), while the heart rate averaged 175 $beats.min^{-1}$ (95% probability range: 169-181 $beats.min^{-1}$).

One objective of this testing was to establish a criterion pass time for all test items up to this task, when performed within the complete Physical Aptitude Test circuit. With this in mind, participants should have reached this stage of the circuit (*i.e.* completed tasks 1-4) in 15 min or less (see Section 2). Ten firefighters (five permanent and five retained) failed to achieve this time threshold. If these ten firefighters were removed from the resulting data, the average completion time for this test item was reduced to 44 s (95% probability range: 41-47 s), followed by a 47-s recovery walk period (95% probability range: 43-51 s).

4.3.6 Test item six: Firefighter-rescue simulation
The final task of the Physical Aptitude Test was the firefighter rescue, for which the mean performance time was 15 s (range: 6 s to 35 s). No difference was observed between the permanent (range: 6 s to 35 s) and the retained firefighters (range: 9 s to 28 s). In the absence of the ten firefighters failing to meet the 15-min time standard for the first four test items, the average firefighter rescue drag time was reduced by just 1 s.

The average oxygen consumption for this test was 2.48 $L.min^{-1}$ (95% probability range: 2.05-2.91 $L.min^{-1}$) whilst the heart rate averaged 173 $beats.min^{-1}$ (95% probability range: 168-179 $beats.min^{-1}$). A common point made by firefighters was that it did not feel as difficult, or as realistic, as the dummy drag. Therefore, the Research Team recommends that this task be modified when it is used in the final Physical Aptitude Test format. Firstly, it is suggested that the harness of a self-contained breathing apparatus unit be used as the interface between the resistance loader and the test subject. Secondly, it is suggested that the mass be primarily dragged rather than carried and dragged. These suggestions will change the mechanics of the drag to more closely replicate the demands of dragging a firefighter during a rescue, and therefore improve the face validity of this test item.

4.3.7 Subjective evaluation of face validity
Subjective ratings of the capacity of the Physical Aptitude Test to reflect the real-world demands of fire fighting were provided by each firefighter upon completion of the final task (visual-analogue scale: Figure 19). On average, the firefighters rated this overall capacity at 82 out of 100 (range: 15-100; standard deviation: 11). On the basis of this evaluation, it is concluded that the face validity of the Physical Aptitude Test was very high, with 66% of the responses falling between 71-93.

4.3.8 Ladder raise task
In addition to performing the Physical Aptitude Test, 108 firefighters within this sample also performed both a standard ladder raise and a simulated ladder raise using the resistance loader. All firefighters successfully completed the ladder raise, with an average performance time of 11.3 s (range: 4.4 s to 31.9 s).

Five firefighters (3.5% of the sample) failed to reach the maximum resistive load applied by the dumbbell and resistance loader during this task. These firefighters were also amongst the slowest to raise the ladder (average time 24 s). Although these numbers are small, a significant ($P<0.05$) correlation was observed, with 40% of the variance in the ladder raise time being explained by differences in the peak load raised using the resistance loader.

4.4 CONCLUSION
This prototype Physical Aptitude Test circuit was devised (Groeller *et al.*, 2012a) following several previous Research Phases (Taylor *et al.*, 2012a, 2012b). In this fourth Research Phase, the prototype test was shown to be a good reflection of physiological and physical capacity, rather than skill, of the test participants.

4.4.1 Recommendation one: Clothing and equipment specifications for testing:
It is recommended that the Physical Aptitude Test be performed whilst carrying a load of 20

kg. This load should be comprised of the following items: underwear, t-shirt, shorts, socks and running shoes (all supplied by the person being tested), and thermal protective clothing, self-contained breathing apparatus (12.5 kg) and ankle weights (1 kg for each leg). The last three items are to be supplied by FRNSW.

4.4.2 Recommendation two: Pass threshold and buffer zone (test items 1-4):
The first four test items of this circuit were designed to match the metabolic cost and the musculoskeletal loading observed during simulations of these tasks as performed by operational firefighters (Taylor *et al.*, 2012b). Based upon the performance speeds and work rates that would elicit the threshold metabolic costs seen in firefighters performing these tasks effectively, and at an acceptable level of operational performance, it was determined that the first four tasks should be completed within 15 min (test item one [hazmat task]: 3 min; test item two [ventilation fan carriage]: 3 min; test item three [motor-vehicle rescue - static hold]: 3 min; test item four [charged hose drag]: 6 min). It is therefore recommended that this 15-min time be considered as a screening threshold (performance standard) for the first four items of the Physical Aptitude Test.

In this testing stage, the average performance time of operational firefighters for the first four test items was 11 min 16 s. Such a fast average performance time was not surprising as volunteers for such testing tend to be amongst the fittest individuals within any organisation. However, ten firefighters failed to complete these tasks under 15 min, and so it was considered that, whilst many firefighters of a high fitness level were tested, a significant range of physical fitness levels were present within the firefighter sample tested.

Of course, consideration must also be given to the use of either an absolute screening threshold, or a threshold with a zone of uncertainty. It is recommended that the latter be adopted, since there will always exist a range of scores around such screening thresholds for which one can never have absolute certainty that these individuals have either passed or failed the Physical Aptitude Test. It is therefore recommended that a 5% buffer (45 s) be used for this purpose. Such a buffer could be used in two ways. Firstly, those falling outside this buffer would be deemed to have absolutely failed the test at this point, and such individuals would not be permitted to continue. Secondly, those slower than the 15-min threshold, but inside the buffer zone, would be allowed to continue the test, with an opportunity to make up time on the last two test items.

The justification for this 5% buffer is two-fold. Firstly, the recommended acceptable work rates for some tasks were set to simulate values at the lower end of the 95% probability range, as this corresponded with values observed in operational firefighters who could still successfully complete the task, but at the lower end of the range of these physiological responses. Thus, by definition, the recommended performance standard could be passed by 95% of the operational firefighters tested. Secondly, since it is the widely accepted convention in science to set statistical thresholds for physiological research at the 5% level, then this value was also considered appropriate in this context.

4.4.3 Recommendation three: Pass threshold and buffer zone (test items 5-6):
A second aim of this testing was to work towards setting an appropriate time standard for

the final two test items of the Physical Aptitude Test: the fire-attack simulation and the firefighter rescue. It is recommended a combined time of these two tasks (the pass threshold) be set at 1 min 53 s (the lower 95% confidence interval for these test items: Sections 4.3.5 and 4.3.6). It is again recommended that a 5% buffer (6 s) be used. In this instance, the buffer would apply just to these two test items, and this would necessitate discrete timing for these tests. Thus, one could satisfy the two pass thresholds for items one-four and five-six. However, whilst various combinations of threshold achievement are possible, and these are addressed in recommendation four, it is recommended that a failing score that falls outside the buffer zone for these last two criterion tasks should constitute an absolute test failure.

4.4.4 Recommendation four: Physical Aptitude Test pass and failure criteria:
It is recommended that any person who satisfies the three static-position holds (item three), as well as both of the threshold times (items one-four and items five-six), should be deemed to have passed the Physical Aptitude Test.

However, it is recommended that four ways should exist for one to absolutely fail the Physical Aptitude Test, and that failure on any one of these criteria should represent test failure, independently of the performance on any of the other test components:

- failure to complete the three positional holds for the required duration: such individuals could be allowed to continue the test, but should still be required to pass this component as part of the complete test sequence
- failure to satisfy the time threshold for test items one-four, and with a score outside the buffer zone
- failure to satisfy the time threshold for test items five-six, and with a score outside the buffer zone
- the attainment of two test times beyond the threshold targets for test items one-four and five-six, both of which fall within the respective buffer zones.

It is further recommended that individuals completing the test in a time frame that falls within the 5% buffer zones for either of the test-item groups (*i.e.* items 1-4 or 5-6) should be identified as candidates for whom retesting might be appropriate.

4.4.5 Recommendation five:
It is advised that all candidates be given a full briefing and are walked through the Physical Aptitude Test before being tested. It is further recommended that all candidates are given a full opportunity to practise each task, and be offered information and advice concerning optimal performance techniques. Such familiarisation periods can improve performance by approximately 5% in strength-based tests (Mendez-Villanueva *et al.*, 2007; Ploutz-Snyder and Giamis, 2001).

SECTION 5: REFERENCES AND RECOMMENDED READING

Constable, S.H., and Palmer, B. (2000). *The process of physical fitness standards development*. Human Systems Information Analysis Center Program Office, Wright Patterson Air Force Base, OH, USA.

Di Prampero, P.E., Davies, C.T., Cerretelli, P., and Margaria, R. (1970). An analysis of O_2 debt contracted in submaximal exercise. *Journal of Applied Physiology*. 29:547-551.

Fullagar, H.H.K., Groeller, H., Sampson, J.A., Notley, S.R., Burley, S.D., Lee, D.S., and Taylor, N.A.S. (2012). The physiological and physical demands of contemporary fire fighting: simulations performed by operational fire fighters. In: Taylor, N.A.S., and Billing, D.C. (Editors). *Physiological and physical employment standards I*. Proceedings of the First Australian Conference on Physiological and Physical Employment Standards (November 27th-28th, 2012. Canberra, Australia). University Of Wollongong, Wollongong, Australia. Pp. 42-43. ***ISBN:*** 978-1-74128-220-7

Gledhill, N., Bonneau, J., and Salmon, A. (2001). *Bona fide occupational requirements*. Proceedings of the consensus forum on establishing *bona fide* requirements for physically demanding occupations. York University, Toronto, Canada. September 13th-16th, 2000.

Gledhill, N., and Jamnik, V.K. (1992a). Characterisation of the physical demands of fire fighting. *Canadian Journal of Sport and Science*. 17(3):207-213.

Gledhill, N., and Jamnik, V.K. (1992b). Development and validation of a fitness screening protocol for firefighter applicants. *Canadian Journal of Sport and Science*. 17(3):199-206.

Groeller, H., Fullagar, H.H.K.F., Sampson, J.A., and Taylor, N.A.S. (2012a). Physiological employment standards for firefighters: *Report 3:* Physical aptitude tests for contemporary firefighters. *UOW-CHAP-HPL-Report-048*. Human Performance Laboratories, University of Wollongong, Australia. For: Fire & Rescue NSW, Sydney, Australia. Pp. 1-32.

Groeller, H., Fullagar, H.H.K., Sampson, J.A., and Taylor, N.A.S. (2012b). Recommended screening tests for contemporary firefighters. In: Taylor, N.A.S., and Billing, D.C. (Editors). *Physiological and physical employment standards I*. Proceedings of the First Australian Conference on Physiological and Physical Employment Standards (November 27th-28th, 2012. Canberra, Australia). University Of Wollongong, Wollongong, Australia. Pp. 44-45. ***ISBN:*** 978-1-74128-220-7

Henry, F.M., and DeMoor, J.C. (1956). Lactic and alactic oxygen consumption in moderate exercise of graded intensity. *Journal of Applied Physiology*. 8:608-614.

Hill, D.W., and Smith, J.C. (1999). Determination of critical power by pulmonary gas exchange. *Canadian Journal of Applied Physiology*. 24:74-86.

Jamnik, V.K., Gumienak, R., and Gledhill, N. (2012). Physiological employment standards II: Developing and implementing physical employment standards for safety-related occupations. *European Journal of Applied Physiology*. *In Press*.

Mendez-Villanueva, A., Bishop, D., and Hamer, P. (2007). Reproducibility of a 6-s maximal cycling sprint test. *Journal of Science and Medicine in Sport*. 10:323-326.

Micklewright, D., Papadopoulou, E., Swart, J., and Noakes, T. (2010). Previous

experience influences pacing during 20 km time trial cycling. *British Journal of Sports Medicine*. 44:952-960.
Payne, W., and Harvey, J. (2010). A framework for the design and development of physical employment tests and standards. *Ergonomics*. 53:858-871.
Ploutz-Snyder, L.L., and Giamis, E. (2001). Orientation and familiarization to 1RM strength testing in old and young women. *Journal of Strength and Conditioning Research*. 15:519-523.
Smither, J.W., Reilly, R.R., Millsap, R.E., At, T.K.P., and Stoffey, R.W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*. 46:49-76.
Taylor, N.A.S., and Billing, D.C. (Editors). (2012). *Physiological and physical employment standards I*. University Of Wollongong, Wollongong, Australia. Pp. 1-109. ***ISBN:*** 978-1-74128-220-7
Taylor, N.A.S., Fullagar, H.H.K.F., Sampson, J.A., and Groeller, H. (2012a). Physiological employment standards for firefighters: *Report 1:* The essential, physically demanding tasks of contemporary fire fighting. *UOW-CHAP-HPL-Report-043*. Human Performance Laboratories, University of Wollongong, Australia. Pp. 1-88.
Taylor, N.A.S., Fullagar, H.H.K.F., Sampson, J.A., Lee, D.S., Notley, S.R., Burley, S.D., and Groeller, H. (2012b). Physiological employment standards for firefighters: *Report 2:* The physiological demands of performing physically demanding fire-fighting duties. *UOW-CHAP-HPL-Report-046*. Human Performance Laboratories, University of Wollongong, Australia. Pp. 1-149.
Taylor, N.A.S., Fullagar, H.H.K., Sampson, J.A., and Groeller, H. (2012c). The physically demanding and critical tasks performed by permanent and retained firefighters. In: Taylor, N.A.S., and Billing, D.C. (Editors). *Physiological and physical employment standards I*. Proceedings of the First Australian Conference on Physiological and Physical Employment Standards (November 27th-28th, 2012. Canberra, Australia). University Of Wollongong, Wollongong, Australia. Pp. 40-41. ***ISBN:*** 978-1-74128-220-7
Taylor, N.A.S., and Groeller, H. (2003). Work-based assessments of physically-demanding jobs: a methodological overview. *Journal of Physiological Anthropology*. 22:73-81.
Tipton, M.J., Milligan, G.S., and Reilly, T.J. (2012). Physiological employment standards I: Occupational fitness standards: objectively subjective? *European Journal of Applied Physiology*. *In Press*.
Whipp, B.J., and Wasserman, K. (1972). Oxygen uptake kinetics for various intensities of constant-load work. *Journal of Applied Physiology*. 33:351-356.

SECTION 6: APPENDICES

APPENDIX ONE: Meeting to discuss possible Physical Aptitude Test options for the approved criterion trade tasks arising from Phase Two

Date: 21/05/12
Location: Deans Meeting Room (University of Wollongong).
Present: Brendan Mott (FRNSW), Nigel Taylor (UOW), Herb Groeller (UOW), John Sampson (UOW), Hugh Fullagar (UOW).
Summary:
The first purpose of this meeting was to present and evaluate a preliminary proposal for the classification of the criterion tasks. Since there were movement pattern similarities across these tasks, then each criterion task was classified into one of four different movement categories.

During these discussions, it was decided to sub-divide the ladder task into two parts (carrying and under-running) as it was considered to fall within each of two different movement classes. The second aim of this meeting was to identify operational constraints that may subsequently influence test selection, design and implementation. Finally, criteria were discussed upon which individual tasks may be eliminated from the list, since it was considered that some duplication may exist among these criterion tasks, and task culling would increase testing efficiency.

Six constraints of practical or logistical significance were identified and discussed in detail:

- environmental constraints: locations and facilities for test administration, and climatic variations
- equipment constraints: personal protective clothing and equipment
- the height of the operating posture (below the neutral plane) for firefighters to avoid excessive heat and smoke exposure
- the structures and surfaces used during ambulatory and load-carriage tasks
- the mass that would be used within the crucial strength task (patient mass)
- the correlation of a “lift and place” activity with critical tasks that require the holding of variously sized objects.

Each was carefully considered when determining the suitability of any criterion task, or a modification thereof, for inclusion within an occupation-specific Physical Aptitude Test.

Three exclusion criteria were identified:

- a low relative, whole-body physiological demand
- movement task duplication
- the availability of suitable substitution tasks.

Should any task satisfy one or more of these characteristics, it would be considered for elimination by Research Team. In this way, efficiencies within the proposed fire-fighter Physical Aptitude Test could be found, without compromising either the sensitivity or specificity of the proposed task battery.

APPENDIX TWO: Meeting to approve Physical Aptitude Test options arising from Research Phase Three

MEETING: Project Management Team:
Date: 22/6/12
Location: Board Room Head Office (FRNSW).
Present: Chair: Brendan Mott (FRNSW), Alison Donohoe (teleconference connection: FRNSW), Darren Husdell (FRNSW), Jim Hamilton (FRNSW), Jim Smith (FRNSW), Rick Griffiths (FRNSW), Brendan Mott (FRNSW), Megan Smith (FRNSW), Herbert Groeller (UOW), Nigel Taylor (UOW).
Summary:
(1) Previous Minutes of 21 May 2012 were accepted by all.

(2) BM indicated since last meeting further discussion had occurred between himself and UoW Research Team regarding some of the recommendations for the testings, location of where testing would be held, who would conduct assessments etc.

(3) NT reiterated how the Research Team came about identifying the 15 tasks that were then assessed. The project is now up to completing report number 3 of 4 reports and needs the committees input to determine if the tests created are practical, transportable and have face validity.

(4) HG went through the attached pages and explained the information detailed on each page. Minutes are to be read in conjunction with documents handed out at the meeting and attached to these minutes.

Table 1: This outlined the 15 tasks that were assessed in Phase 2 of the project.

Table 2: This outlined those tasks that were no longer formally considered in Phase 4 of the project. More specifically:
Simulation 3: this task had very low oxygen consumption, was short duration compared to other tasks tested.
Simulation 6: low physiological demand compared to other tasks which would require 40-50% greater demand.
Simulation 10: the same physiological demands can be replicated by other tasks.
Simulation 11: low demand physiologically and can be substituted.
Simulation 12: this can be substituted with Simulation no 14.

JH raised concerns that some hose drag tasks were removed, but was reassured by HG other tasks remaining met the same physiological demands.

Table 3: Outlined the tasks now remaining which are recommended as part of the fitness assessment.

Table 4 and 5: Breaks down the tasks into those that have a holding and those with a carry characteristic.

Table 6: Breaks down all tasks into 4 classes, with a few tasks generally under each class. HG believes this is the minimum number of classes that needs to be included in the assessment battery. It is felt this number will assist in keeping it simple but also test appropriately to the correct level. BM reinforced any battery of tests needed to also show face validity.

AD asked whether tasks that may have a skill component will be altered to a generic task replicating the movement required. For example, if the ladder raise is still to be included, that it may need to be performed in a simulated capacity e.g. unilateral shoulder raise. HG advised may be able to use other equipment to simulate the ladder being raised to and above shoulder height. NT advised they may utilise university students to test different ways of testing tasks such as the ladder raise task, and to determine the load given by assistance of the person stabilising the ladder vs underrunning the ladder.

The assessment constraints page outlines the guidelines the researchers are taking into consideration when developing the battery of tests. Any equipment that will be required to perform testing: will need to consider its weight (*e.g.* if a dummy is used), cost and portability.

(5) HG and NT explained the recommendations:

Recommendation 1: To determine what would be the average mass of firefighters in the Australian population (taking into account the same gender mix in FRNSW). HG advised this could be 82 kg. With the addition of PPE and boots, this would equate to 102 kg for firefighter rescue. Although this task may be rarely performed, it is highly critical. Of the 55 firefighters who performed the simulations as part of the research, the average weight of the participants was 83 kg.

To determine the load required for simulation 8 (firefighter down- rescue) UOW need to determine how to create a load that can be consistent. The Research Team initially thought of a sled, but this would not work due to different surfaces test may be performed on *e.g.* grass versus vs concrete. It was agreed that an alternate to using dummies would be ideal.

Recommendation 2: All agreed that the test would need to be portable across the state, and hence an email was circulated out of session identifying some of the possible constraints that may need to be considered.

Recommendation 3: It was discussed that the approximate weight of PPE is 20 kg (minus boots) but this could be added with a weight vest or belt etc. to mimic the mass of boots.

Recommendation 4: A picture of a mock frictional force device (Fire Sim Loader) is on Figure 1. Creating a device like this will make the test transportable and reduce manual handling risks to the assessors, as the device will be lightweight.

Recommendation 5: Researchers want to know what height will be considered the 'neutral plane' to maintain a crouch posture with certain tests *e.g.* firefighter rescue. See also figure

2. Needs to be determined how this height can be consistent *e.g.* laser level, poles at a certain height placed by assessor.

Recommendation 6: To make the test portable, would consider use of a step activity to mimic negotiation of a stairwell.

Recommendation 7: The committee will need to determine whether they want to use an absolute or relative height for lifting tasks. There will also have to be tasks that assess ability to lift to shoulder height and above shoulder height as single or repetitive lift.

Recommendation 8: This relates to Figure 2. No time frame has been determined as yet as to how long the course will run for, or individual tasks will run for. Some participants may be able to pass individual elements *e.g.* the carry and hold tasks, but with all tasks combined in a sequence may not be able to complete the course. It will require FRNSW to make a determination of the duration of individual tasks and whole assessment time.

(6) ACTIONS:
UOW will send out a formal letter of request asking for answers to key questions raised from their recommendations. This will be done within the next week. Answers are to be provided back from the Committee to help the UOW further streamline the battery of tests that make up the PAT.

APPENDIX THREE: Summary of the electronic communications and discussions relating to the development of a formal response by Fire & Rescue NSW concerning request three from Appendix Four of Report Three: Safe operating height

Request three:
Within the fire-attack and firefighter-rescue tasks, it is recommended that all participants be required to adopt safe operating postures (below the neutral plane) to simulate the avoidance of excessive heat and smoke exposure. The resulting semi-crouched position will be determined so that each participant keeps his/her helmet below the same fixed height. This will also facilitate an assessment of each participant's ability to operate within somewhat confined spaces.

The electronic mail correspondence below was provided to the researchers by Fire & Rescue NSW for inclusion within this report[14]. The purpose of this inclusion is to show how this height was determined.

> > > Brendan Mott 17/07/2012 4:49 PM > > >

Hi John,
Thank you for your time this afternoon.

As discussed, the University Of Wollongong has requested some additional information following the simulations which you and your team helped us out with earlier in the year.

As the fire attack and firefighter rescue tasks have now progressed to the stage where we are developing the physical screening test the University has requested information relating to the heights below which it would be expected that firefighters would be expected to perform 1) a fire attack and 2) firefighter rescue (these may be one in the same).

If you could discuss this with your team and provide a recommendation (along with an outline of the methodology/reasoning applied) that would be great.

Thanks
Brendan Mott AEP
Team Leader Health and Fitness
Health & Safety

> > > Brendan Mott 30/07/12 12:37 > > >

Hi John,
Just wondering if you and your team had any discussions surrounding these heights?
Cheers
Brendan Mott AEP

[14] To facilitate reading, these electronic communications have been re-arranged in chronological order, as they were originally provided in the reverse order, as is the case for most such electronic records. However, the text has not been modified.

> > > John Mcdonough 2/08/12 8:59 > > >

Can I have any comments asap. My view is, ‘it depends’. I spoke to Brendan about this and he is still adamant that he needs some rule of thumb height to base the research on.
Cheers,
John

> > > Jason Collits 2/08/12 10:34 > > >

My suggestion would be this…

I would suggest that a firefighter should be capable of operating at a low level height. I would suggest a ‘ceiling’ height for this exercise of 1.25 metres.

The firefighter should be capable of moving forwards and backwards, perhaps for a distance of 5 metres or so, ensuring that their head did not at any point encounter that ceiling height.

Just how the firefighter moved, whether it was on one knee, two knees, or even a ‘side saddle’ movement with the weight on the hip and bent leg, I would suggest is not a crucial part of any such testing.

What I would suggest is essential is that the firefighter can hold a branch in one hand that does not come in contact with the ground (*i.e.* not reliant on moving on both knees and both hands) and can maintain situational awareness as they move.

This may require them being capable of looking forward the bulk of the time of the exercise, but still being capable of looking up without again breaking the ‘ceiling’ height of the test.

The key aspects would be:

The ability to maintain a low body position
The ability to display mobility at that low body position
The ability to have at least one hand free at all times to manage a branch, haul hose, support / move a casualty etc.
The ability to keep looking in front, above and around and still maintain that low body position.

That’s my suggestion.
Cheers,
Jason

> > > John Mcdonough 7/08/2012 10:09 AM > > >

Hi Brenden,
Sorry about the delay mate. Here are some ideas as aticulated by Jason after we talked about it as a group. I think this would be a good starting point.

Cheers,
John

> > > Brendan Mott 7/08/12 11:19 > > >

How does this look?

Recommendation 3 (Response to request 3)
The Firefighter Operational Training Team (Team Leader John McDonough) has made the following recommendations relating to Critical Task 7 - Fire Attack and Critical Task 8 - Firefighter Rescue:

- The height of a neutral plane which would be reasonable expected to be encountered in the event of Fire Attack and Firefighter Rescue would be 1.25m.
- Firefighter should therefore be capable of operating below a 'ceiling' height of 1.25 metres in the event of Fire Attack and Firefighter Rescue.
- Movement could on one knee, two knees, or even a 'side saddle' movement with the weight on the hip and bent leg.
- The applicant should hold a branch in one hand that does not come in contact with the ground during fire attack (i.e. not reliant on moving on both knees and both hands)
- This may require them being capable of looking forward the bulk of the time of the exercise, but still being capable of looking up without again breaking the 'ceiling' height of the test.

Brendan Mott AEP
Team Leader Health and Fitness
Health & Safety

> > > John Mcdonough 8/08/12 3:36 > > >

Hi Brendan,
I think that's a good start. I think 1.25 is reasonable. I'm 184cm, and can do it easily.
Cheers - John

< < < Brendan Mott 20/11/12 12:35 > > >

Hi John and Jason,
Jason, Thanks for your earlier email. While the sling methods you pointed out definitely seem to improve the ease of the rescues at low levels, the task analysis of the rescues the University conducted at Londonderry identified the use of the BA harnesses as the attachment point. As such the screening tests they have developed are based on the 2 handed backwards drag observed during the simulations.

I will be travelling down to Wollongong in the morning to view the testing myself for the first time. However early indications are that generating the required force to drag the equivalent of 82 kg backwards whilst staying below 1.25m is proving very difficult and requires very awkward postures.

In discussions with the project management team yesterday it was discussed whether the 1.25m height is a realistic expectation for this task and technique.

See attached preliminary trial photos - please note that a weighted BA harness/backplate will replace the dumbbell seen in the current photos of the rescue and a section of weighted hose will replace the weights on the fire attack simulation.

Your thoughts on the suitability of the 1.25m for this style of firefighter rescue would be appreciated.

Regards
Brendan Mott

< < < Brendan Mott 20/11/12 12:35 > > >

Hi Brendan,

Thanks for the email and for the photos. I wasn't present at the testing conducted at Londonderry, so was not aware that there had been previous tests done on the casualty drags. The bulk of the rescues done by FRNSW personnel are done on civilians, (I can't recall when one of our people was last rescued in an unconscious or non-ambulatory state from within a structure) but I can understand why a firefighter removal is being used as the test benchmark.

We do instruct a technique that could be used by many firefighters to move an 80kg patient without equipment that would see them remain under the 1.25 metre 'ceiling', but it is not an easy thing for most to do and it would be very difficult to replicate using a backplate and resistance line. And this technique would not work for the rescue of a firefighter in SCBA as it involves cradling the casualty against your body as you shuffle back, and the SCBA on the firefighter being rescued would prevent the close cradling necessitated.

There is absolutely no question that maneuvering a casualty out of a building whilst maintaining a low body positioning is a difficult thing to do and it does require awkward postures.

John and I were just discussing this and perhaps the threshold of 1.25 metres could be used as a mobility height for the purposes of fire attack. Perhaps when a casualty is being removed the threshold is removed, the person is simply told to stay as low as they can, or an increased height of perhaps 1.5 or even 1.6 metres or there abouts is used.

The reality should be that when a casualty is being removed from the structure the firefighter is retreating from the danger, so the heat impacting on the body should theoretically be lessening. As a result the need to stay as low is hopefully not as great.

If the internal conditions within the structure were such that a very low body positioning had to be maintained during the removal, perhaps the casualty removal becomes a job for two

firefighters (as is often the case for 100+kg casualties).

As you can see from this video, rescues from within a structure are not graceful nor necessarily ergonomically sound. They are rough and urgent.

http://www.youtube.com/watch?v=3E6WIZ36E2s

I hope this helps in some way.

Best of luck with the testing.

Cheers,

Jason

APPENDIX FOUR: Formal responses from Fire & Rescue NSW to the requests for supplementary information contained within Appendix Four of Report Three

NSW GOVERNMENT | Fire & Rescue NSW

File Ref. No: NFB/08711
TRIM Ref. No:

Associate Professor Nigel Taylor

Centre for Human and Applied Physiology
School of Health Sciences
University Of Wollongong
Northfields Avenue
Wollongong, NSW, 2522

7 August 2012

Dear Nigel,

The following information has been provided by the Physical Employment Standard Project Management Team in response to Report 3 - Appendix Four: Request for additional information from Fire & Rescue NSW:

Recommendation 1 (Response to request 1)

The project management team recommends that participants wear the following when participating in physical aptitude testing:

1) Structural Firefighting Helmet
2) Flash hood
3) Structural Firefighting Tunic (over own clothing)
4) Structural Firefighting Pants (over own clothing)
4) Gloves
5) SCBA Cylinder*
6) Weight vest accounting for Radio (1.1kg), torch (400grams) and Boots (2.6kg)**

**Assessments should account for SCBA not being warn during rescue tool lift/hold assessment*

***The listed weight for the radio and boots are mean data collected during simulations of critical tasks 1-15 performed in phase 2 of this research. Boot mean weight was also verified within reasonable limits in a report by Dr Kinchington (Podiatric Physician & Musculoskeletal Injury Consultant) – "Fire & Rescue NSW Structural Fireboot Report For Clinical Applications December 2011"*

Fire & Rescue NSW
Executive Directorate

ABN 12 593 473 110
PO Box A249,
Sydney South NSW 1232

www.fire.nsw.gov.au
T (02) 9265 2999
F (02) 9265 2988

Recommendation 2 (Response to request 2)

The Team Leader of Rescue Training Anthony NorthbrookHine in consultation with his team has provided the following recommendations relating to Critical Task 2 - Motor Vehicle Rescue:

- For the Super Spreaders, a user of this tool would reasonably be expected to use the tool from ground level to their shoulder height. Any higher and safety dictates using a step or ladder to reach the task.
- For the Shears a user of this tool would reasonably be expected to use the tool from ground level to head height (tool at eye level) with arms below the tool. Any higher and safety dictates using a step or ladder to reach the task.

The technique used to complete Critical Task 13 - Underrunning Ladder governs that the lift heights are relative to the participant i.e. ground to overhead height of the applicant during underrunning.

Recommendation 3 (Response to request 3)

The Firefighter Operational Training Team (Team Leader John McDonough) has made the following recommendations relating to Critical Task 7 - Fire Attack and Critical Task 8 - Firefighter Rescue:

- The height of a neutral plane which would be reasonable expected to be encountered in the event of Fire Attack and Firefighter Rescue would be 1.25m.
- Firefighter should therefore be capable of operating below a 'ceiling' height of 1.25 metres in the event of Fire Attack and Firefighter Rescue.
- Movement could on one knee, two knees, or even a 'side saddle' movement with the weight on the hip and bent leg.
- The applicant should hold a branch in one hand that does not come in contact with the ground during fire attack (i.e. not reliant on moving on both knees and both hands)
- This may require them being capable of looking forward the majority of the time, but still being capable of looking up without again breaking the 'ceiling' height.

Recommendation 4 (Response to request 4)

The project management team after considering future workforce planning and current workforce demographics, endorses the following methodology relating to setting the firefighter mass for Critical Task 8 Firefighter Rescue.

The current gender distribution of FRNSW firefighters is 4.5% female and 95.5% male. Accounting for the current trend towards an increasing numbers of successful applications from females for firefighting positions the project management team considers it reasonable to utilise a gender distribution of one female (10%) to every

nine males (90%) when setting the firefighter mass for the assessment arising from Critical Task 8 Firefighter Rescue.

Utilising the most current Australian Statistics, the average mass of firefighters could be determined using 90% of the average Australian male mass and 10% of the average Australian female mass. This method provides an indicative average mass of 82.01 kg (Australian Bureau of Statistics, 2005).

The project management team endorses this mass (82.01kg) for use in developing the assessment arising from Critical Task 8 Firefighter Rescue.

Yours sincerely

Alison Donohoe
Assistant Director Health & Safety

APPENDIX FIVE: International scientific conference abstracts: First Australian Conference on Physiological and Physical Employment Standards

The following abstracts were communicated at the *First Australian Conference on Physiological and Physical Employment Standards* (November 27th-28th, 2012, Canberra, Australia). This was the first international, scientific conference dedicated to the development of physiological and physical employment standards. The conference was organised by the Research Team in collaboration with the Defence Science and Technology Organisation (DSTO). Guest speakers included:

Alexander Zelinsky (Chief Defence Scientist, DSTO)
Warren Snowdon (Minister for Defence Science & Personnel)
Paul Wellings (Vice Chancellor, University of Wollongong)
David Morrison (Chief of Army, Australian Defence Force)
Leanne Close (Assistant Commissioner, Australian Federal Police)
Greg Mullins (Commissioner, Fire & Rescue NSW)

Keynote speakers included:

Daryl Allard (Canadian Defence Forces, Canada)
Alison Donohoe (Fire & Rescue NSW, Australia)
Yoram Epstein (Heller Institute of Medical Research, Israel)
Veronica Jamnik (York University, Canada)
Bradley Nindl (US Army, U.S.A.)
Stephan Rudzki (Director General Strategic Health Coordination, Australia)
Michael Tipton (University of Portsmouth, U.K.).

This conference provided a first-level peer review and critique of the research undertaken across all stages of this project. Indeed, one session was dedicated to the presentation of the methods and observations that led to the proposed physiological employment standards for Fire & Rescue NSW. The Chair of that session was Veronica Jamnik (York University, Canada) who was instrumental in establishing the endurance fitness standard for firefighters used by fire brigades around the world, including Fire & Rescue NSW.

Abstracts from this meeting were published within the following proceedings document:

Taylor, N.A.S., and Billing, D.C. (Editors). *Physiological and physical employment standards I.* Proceedings of the First Australian Conference on Physiological and Physical Employment Standards (November 27th-28th, 2012. Canberra, Australia). University Of Wollongong, Wollongong, Australia. ***ISBN:*** 978-1-74128-220-7.

- *Contemporary physical and medical standards for Fire & Rescue NSW*. Pp. 38-39.
- *The physically demanding and critical tasks performed by permanent and retained firefighters*. Pp. 40-41.
- *The physiological and physical demands of contemporary fire fighting: simulations performed by operational fire fighters*. Pp. 42-43.
- *Recommended screening tests for contemporary firefighters*. Pp. 44-45.
- *A commentary on endurance fitness standards applied to occupations that involve load carriage and manual handling*. Pp. 49-50.
- *How much money could an Australian fire brigade save annually by increasing firefighter fitness?* Pp. 82-83.

A5.1 Abstract one: Contemporary physical and medical standards for Fire & Rescue NSW

Alison Donohoe
Health & Safety Branch, Fire & Rescue New South Wales, Australia
Corresponding author: Alison.Donohoe@fire.nsw.gov.au

Fire & Rescue NSW (FRNSW) is one of the world's largest urban fire and rescue services, and the largest and busiest in Australia. The physical capability and psychological resilience of firefighters is critical to safely achieving the organisations overriding purposes of enhancing community safety, quality of life and public confidence by minimising the effects of hazards and emergencies on the population it serves (7.2 M), and on the property, environment and economy of NSW.

The equipment for, and techniques of contemporary fire fighting have evolved over time, along with the demands of this essential job. Indeed, the exposure of firefighters to stresses associated with material handling, load carriage (including personal protective equipment), operating in confined spaces and hot environments, are known to impose unique physical and physiological demands upon firefighters. These demands often lead to injuries, which are skewed towards older firefighters. For instance, male firefighters >40-years old represent 61% of the workforce, but suffer 75% of the sprain and strain injuries and account for about 90% of the lost work time (1998-2007). The sudden changes in work tempo from rest to high intensity work, place significant demands upon the cardiovascular system. Thus, firefighters are 14 times more likely than similarly aged members of the community to suffer sudden cardiac death following an alarm, and 136 times more likely after strenuous fire fighting. This too is more evident in older individuals.

To maximise operational readiness, and the health and safety of the workforce, FRNSW has engaged appropriate occupational medicine and scientific experts to ensure that its medical and physiological employment standards reflect this evolution. The aim of this research is two-fold. Firstly, it seeks to facilitate the identification and recruitment of individuals who are medically and physiologically capable of tolerating the work-related stress of fire fighting, without succumbing to injury or endangering colleagues or community members. Secondly, it is aimed at identifying individuals who are ill-suited to fire fighting, and who would be prone to medical complications or injuries when performing the more demanding aspect of fire fighting.

Once in the workforce, FRNSW promotes the maintenance of a functional and medical status consistent with these standards. This requires extensive support programmes supported by Exercise Scientists within FRNSW, and other allied health professionals. Additionally, FRNSW has established robust support programmes for the implementation of mandatory standards.

A5.2 Abstract two: The physically demanding and critical tasks performed by permanent and retained firefighters

Nigel A.S. Taylor, Hugh H.K. Fullagar, John A. Sampson and Herbert Groeller
Centre for Human and Applied Physiology, University of Wollongong, Australia
Corresponding author: ntaylor@uow.edu.au

INTRODUCTION

During fire emergencies and rescue operations, firefighters experience a high-level of physical stress, and are expected to possess the physiological attributes necessary to tolerate these demands. Thus, recruitment practises leading to the identification of stress-resistant individuals would increase unit capability and minimise the risk of firefighter injury. One strategy for this is to develop valid, physiological employment standards. In this, and two subsequent presentations, the research stages leading to the development of predictive screening tools for firefighters will be described. The first phase in achieving this outcome was to conduct a comprehensive review of the physical demands of fire fighting, and this is focus of our first communication.

METHODS

To identify the essential tasks of fire fighting, a four-stage process was undertaken. Firstly, researchers were familiarised with the trade, visiting 11 Fire Stations and interviewing 106 permanent and retained firefighters. Secondly, many tasks were demonstrated to the researchers, who performed preliminary movement and task analyses. Thirdly, in consultation with senior subject-matter experts, 31 tasks were identified for further investigation. Finally, 1,011 firefighters (717 permanent, 272 retained, 22 incomplete) participated in a confidential survey concerning task importance, frequency, duration and difficulty, thereby facilitating identification of the essential trade tasks.

RESULTS AND CONCLUSION

Since the next phase involved quantifying the demands of these tasks, then it would be inefficient to study all 31 tasks. A five-stage filtration process was therefore applied. (1) Tasks with a subjective effort rating below the two calibration tasks were eliminated. (2) If a sub-threshold task was performed more frequently or was more important than the calibration tasks, or was identified by >20% of firefighters to cause a physical limitation, then it was retained. (3) Where tasks were sufficiently similar, the more difficult task was retained. (4) Two-person, skilled tasks introduce uncontrolled performance variability, reducing measurement precision. Most were eliminated, unless the task was unskilled or individual contributions could be easily measured. (5) Tasks that were difficult to define, due to their nature or duration, were hard to reduce into discrete and reproducible tasks. These too were excluded. The resulting 15 tasks were: rolling out hoses, locating and connecting to hydrant, coupling hoses, dragging charged hoses, stair climb with hose, prolonged use of hoses (38 and 70 mm), fire attack, ladder use, firefighter rescue, using spreaders and shears, using sledge hammer to gain entry, carrying ventilation fan (stairs), hazmat operations and dragging charged hose (bush: hilly and uneven terrain). These tasks represent a valid subset of physically demanding activities performed by firefighters in metropolitan and regional NSW.

A5.3 Abstract three: The physiological and physical demands of contemporary fire fighting: simulations performed by operational fire fighters

Hugh H.K. Fullagar, Herbert Groeller, John A. Sampson, Sean R. Notley, Simon D. Burley, Daniel S. Lee, and Nigel A.S. Taylor
Centre for Human and Applied Physiology, University of Wollongong, Australia
Corresponding author: ntaylor@uow.edu.au

INTRODUCTION
In the preceding communication, 15 tasks, representing a valid subset of the physically demanding activities performed firefighters in metropolitan and regional NSW, were identified. The second phase in developing legally defensible physiological employment standards was to quantify and evaluate the physiological demands of these tasks, when performed by firefighters.

METHODS
Experienced, operational firefighters (*N*=51) with a range of skills, ages, body sizes and fitness levels were recruited (37.3 y [range 23-57]; operational experience 9.2 y [range 1-29]; mass 55.3-113.6 kg); female representation reflected the workforce. Simulations were designed by subject-matter experts, then set out and controlled by Training Officers. Firefighters participated as platoons, under direction of their Station Officer, ensuring realistic operational efficiency and tempo. All simulations used contemporary tools and equipment, with firefighters wearing the appropriate personal protective clothing, equipment and breathing apparatus. The following variables were monitored continuously: heart rate, oxygen consumption, minute ventilation, tidal volume and breathing frequency. Tasks were also analysed using still and video photography.

RESULTS AND CONCLUSION
Simulation durations ranged from 1.14-52.33 min, with eight being <5 min, two were 5-10 min, one was 10-15 min and five tasks lasted >15 min. Since the next phase involved developing a subset of screening tests that could represent the observed tasks and physiological strain, algorithms were used to separate strength- and endurance-related activities, and to classify tasks according to the body region involved, primary movement patterns and loads carried. No tasks were unloaded endurance activities. No strength or muscular-endurance activities involved lifting and placing, or twisting and turning. Only 30% of these tasks were deemed to be endurance-dependent activities (cardiorespiratory or muscular). The 15 tasks were dominated by actions involving the pushing, pulling or dragging of objects >20 kg. The algorithm also culled the least demanding activities, and grouped tasks sharing common movements and physiological attributes, thereby identifying other activities that could assess these capacities under more stressful conditions. It was evident that these essential fire-fighting tasks primarily involved holding and carrying actions, and were largely reliant upon upper-body strength or muscular endurance. Through this analysis, identification of the criterion fire-fighting tasks was facilitated, providing a valid prescription for designing recruit screening tests. This is the focus of the next presentation.

A5.4 Abstract four: Recommended screening tests for contemporary firefighters

Herbert Groeller, Hugh H.K. Fullagar, John A. Sampson and Nigel A.S. Taylor
Centre for Human and Applied Physiology, University of Wollongong, Australia
Corresponding author: hgroell@uow.edu.au

INTRODUCTION

The previous presentation described the detailed analysis of 15 essential fire-fighting tasks, leading to the generation of a list of criterion tasks and a valid prescription for the design of physiological screening tests for potential firefighters. The third phase of this research project, and the focus of this communication, was centred upon the development of these screening tests.

METHODS

The criterion tasks were carefully examined. Each task was classified into a movement category (Table), and evaluated against exclusion criteria: low metabolic demand, movement duplication and possible substitution tasks. Six operational constraints were considered: environmental, equipment, operating posture, structures and surfaces used during ambulatory and load-carriage tasks, load masses, and the existence of possible generic lifting and locomotor activities.

Table: Criterion task movement classifications.

Class	Class descriptions	Criterion tasks
1	Single-sided carrying tasks	Hazmat task
		Rolling out hose (70 mm)
		Locating and connecting to hydrant
		Drag charged 70-mm hose (lateral)
		Ladder carriage (10.5 m)
		Prolonged use of hose (38 mm)
		Prolonged use of hose (70 mm)
		Stair climb with ventilation fan
2	Overhead push and holding tasks	Motor-vehicle rescue
		Ladder under-run (10.5 m)
		Using a sledge axe to gain entry
3	Cardiorespiratory dragging tasks	Fire attack
		Dragging charged hose (38 mm)
4	Crucial strength tasks	Stair climb with charged hose
		Firefighter rescue

RESULTS AND CONCLUSION

From this distillation, it was proposed that the screening tests be conducted as an uninterrupted sequence of activities (circuit) involving: single-sided carriage task(s), holding task(s), hose drag, fire attack and firefighter rescue. It was recommended that satisfactory performance be determined from the time required to complete the circuit, with the standard being dependent upon the minimal acceptable work tempo for these tasks, as performed by existing operational firefighters.

Human Performance Laboratories (Australia)

A5.5 Abstract five: A commentary on endurance fitness standards applied to occupations that involve load carriage and manual handling

Nigel A.S. Taylor
Centre for Human and Applied Physiology, University of Wollongong, Australia
Corresponding author: ntaylor@uow.edu.au

INTRODUCTION
Cardiorespiratory (endurance) fitness standards for physically demanding occupations are often referenced to a measured or predicted peak aerobic power (peak oxygen consumption). Most commonly, this is expressed in the form of a mass-normalised or specific power ($mL.kg^{-1}.min^{-1}$), which is then predicted from a maximal test (*e.g.* shuttle-run test) performed in sports attire. In this communication, the validity of this approach is examined, with respect to its application to endurance fitness standards for occupations that involve load carriage and manual handling, and with a view to removing mass bias from physiological employment standards.

COMMENTARY
A good working example for this discussion is the fitness standard used by most fire-fighting organisations (45 $mL.kg^{-1}.min^{-1}$), which is invariable predicted as described above. Notwithstanding its popular use, this overall approach is frequently inappropriate for the following reasons. (1) A one-to-one relationship between oxygen consumption and body mass does not exist. (2) Linear (arithmetic) normalisation fails to fully account for the inter-individual variability in oxygen consumption. (3) The coefficient of variation for oxygen consumption often exceeds that for body mass. (4) There is a positive relationship between the peak absolute oxygen consumption and body mass, but a negative relationship between peak specific oxygen consumption and mass. (5) The affect of these artefacts increases as individuals approach body-size extremes. (6) Normalisation must reflect total mass, including protective clothing and equipment masses, otherwise the standard will be artificially inflated. (7) The fitness standard should be derive as a power ($mL.kg^{-0.67}.min^{-1}$), and not as a linear function. This will ensure that inter-individual variations in the masses of those investigated were not responsible for determining the employment standard. (8) Unloaded endurance tests are unreliable screening methods for occupations in which load carriage is an integral requirement, since absolute oxygen cost changes is proportional to the specific load (5% increase in relative load produces a 5% elevation in oxygen consumption). Thus, constant loads carried by workers of different masses, represent a greater metabolic demand for lighter individuals.

CONCLUSION
If one accepts these points, then one arrives at two conclusions. Firstly, the minimal endurance standard for firefighters (45 $mL.kg^{-1}.min^{-1}$) may not have been correctly derived, and it may even be artificially inflated. Secondly, it is quite possibly invalid to use an unloaded endurance test to evaluate this physiological capacity for workers in occupations that involve load carriage and manual handling.

A5.6 Abstract six: How much money could an Australian fire brigade save annually by increasing firefighter fitness?

Nigel A.S. Taylor and Elizabeth A. Taylor
Centre for Human and Applied Physiology, University of Wollongong, Australia
Corresponding author: ntaylor@uow.edu.au

INTRODUCTION
At this conference, several organisations have described injury rates in stressful occupations. In some cases, injuries are associated with non-work activities, often most prevalent in younger people. Indeed, across the manual-trade industries, workplace injuries occur most frequently within young, inexperienced workers. However, in some jobs, the reverse is observed, and this tendency, we suspect, is linked with habitual exercise behaviours. Herein, the case is presented that fire brigades, and perhaps other emergency services, can reduce both the rate and financial burden of workplace injuries, simply by increasing the health and fitness of employees.

METHODS
Firefighter injury data were extracted for 1998-2007 (Australia) from a database maintained by a Workers' Compensation insurer. Within an organisation employing 3,514 permanent and 3,390 on-call firefighters, injuries and near-misses sustained at work that led to lost work time or medical treatment entered this database. Data were normalised to 1,000 full-time employees. The savings identified below include only the compensation claims. However, the potential savings, as reflected within insurance premiums, can be several times greater.

RESULTS
Annual averages were: 170.5 injuries, 41.4 years of lost time and $2,829,682 in claims per 1,000 firefighters. Individuals aged 40-50 years accounted for 38.4% of all injuries, but there was a significant under-representation of firefighters <30 years, with 9.3% of the workforce suffering just 2.4% of all injuries. Injury rates were skewed towards the older ages: 40-50 y = 17.6%, 50-60 y = 23.9%, 60-70 y = 77.1%. Firefighters over 40 y had an injury probability 4-18 times higher than those <30 y. Thus, if the injury rate of 30-40-y-old males (11.7%) could be retained as firefighters aged, then 237 fewer injuries would occur each year in males >40 y, and the age- and gender-specific saving would be $5,687,934 *per annum*. The single most common injury types were joint and muscle strains and sprains (65.6% of all injuries). These incurred a net cost of $69,290,798 over this decade. There was also a very strong age dependency for these injuries, and this was linked very powerfully to the net cost of claims.

CONCLUSIONS
With the exception of emergencies, fire fighting is a somewhat sedentary occupation. Since, habitual inactivity leads to many degenerative states that predispose to injury, then implementing structured health, fitness and rehabilitation programmes that target older firefighters will reduce their high injury rates, and yield significant personal and organisational savings.

APPENDIX SIX: Meeting to report on, and approve the completion of Phase Three (Section Two) research activities

Date: 19/11/12
Location: Board Room Head Office (FRNSW).
Present: Jim Hamilton (FRNSW), Megan Smith (FRNSW), Brendan Mott (FRNSW), Daniel Willoughby (FRNSW), Herb Groeller (UOW), John Sampson (UOW)
Apologies: Alison Donohoe (FRNSW), Rob McNeil (FRNSW), Darren Husdell (FRNSW), Jim Smith (FRNSW).
Summary:
(1) BM provided a brief overview of the phases of the project that had been completed so far. Outlined committee provided recommendations on assessment protocol at the end of June and these have been taken on board. The research team has now put together a circuit of assessments in line with both these recommendations and the data arising from the task simulations.
(2) **Proposed Physical Aptitude Test Components:**
UOW provided an overview of the protocol of each assessment included in the proposed Physical Aptitude Test. The following notes should be read in conjunction with the attached handout.

1. Box lift and place
This is designed to assess the physical aptitude of the ladder raise. Main concern with the ladder raise is still the skill component required to raise the ladder. Looked at varying options including the box lift and seated shoulder press as a measure of upper body strength. Most correlated was determined to be the box lift. Further investigation of potential assessments will occur during the validation trial to be performed with incumbent firefighters. This assessment will be performed as an isolated assessment and not as part of the circuit.

2. Unilateral load carry (flat)
Refer to page 2 of the handout. Assessed 8 loaded laps and 8 unloaded laps of the course alternating one at a time. Then correlated and tested pure load carrying demands. Lower range recorded VO2 was 1.4L/min which correlated to 330 meters in 5 minutes with the load carried the entire time. Due to the likelihood of failure of this assessment due to grip strength rather the cardiovascular demands decided to reduce the test time to 3 minutes, particularly considering other load carry assessments. This equated to 195 metres in 3 minutes.

3. Loaded step task.
Refer to page 3 of handout. Single step 36 times. Similarly 3 minutes duration to complete

4. Motor vehicle rescue simulation.
Refer to pages 4-6. Three holds of 40 seconds each with 20 seconds rest in between. One lift above shoulder height, one at waist height and one at knee height. As requested by FRNSW committee all heights are relative to applicant. Tool is 19kg in weight. Handle is held in line with eye, waist and knee, therefore weight is forward of these points to simulate

the mechanics of the rescue tool. BM suggested giving some guidelines for heights to applicant to take away subjectivity. HG suggested using an individuals anatomical points to distinguish this and provide a warning system approach. BM felt this is open to challenge from an applicant and difficult for assessors without an anatomical understanding. HG outlined that the main focus is to ensure that you distinguish the 3 different lift styles. Committee agreed with this recommendation.

5. Walk hose drag.
Refer to page 11. Hose drag has both a lift and drag force that is required. Reel system represents the drag force and weight represents the lift force. Finals numbers for each of these forces are still to be validated. Thirty metre distance resisted walk followed by returning 30 metres unloaded (load carried back by helper), then repeated until accumulated desired length of 270 metres. Assessment requires 2 assessors, one to wind up reel and one to carry load. Cannot have applicant carry load back.

6. Simulated fire attack.
Refer to page 12-13. Time to complete this task not determined as yet. Height restriction of 1.25 metres came from operational recommendations.

7. Simulated Fire fighter rescue.
Refer to page 14-15. 45kg drag resistance and 20kg lift over 10 metres. Official time to be determined and finalised. Looking at using BA vest to make interface with the load resistance reel more realistic

HG outlined that validation of some loads is still required using a load cell and subjective feedback.
HG outlined that the drag reel device works really well and nullifies and previous limitations of sled dragging test on differing terrains such as concrete vs grass.

(3) **<u>Project Management Team Feedback:</u>**
BM raised the concerns around the logistics of setting up a box lift assessment for the ladder raise. All committee agreed that it would be better if the reel was used and would also mean that there was less equipment to be carried by the assessor. HG outlined that they would need to correlate the forces required with the reel resistance. Will need to correlate data at FF assessment stage.

JH questioned the FF rescue and 1.25metre height restriction. Suggested that this was a very subjective measure and would be based on many environmental factors that could alter this. JH feels that most FF would stand more upright and simply deal with the heat for the short period of 10 metres, rather then kneel and crawl under this level. JH feels that the hose drag under 1.25 metres is a realistic figure. BM outlined that he would discuss this with subject matter experts and report back (see action items).

BM discussed the possibility of limiting the reel wind up phase of hose drag assessment to remove the requirement of having 2 assessors present. HG outlined that an automatic function could be designed using a power source such as a car battery. This would take

some time to design and implement. BM suggested that this could be an improvement made moving forward. Next permanent recruit class is likely to be March 2013 with basic reel version implemented by this time and improvement to follow.

BM outlined to committee that no significant differences in FF tasks were noted in review of the survey results between regional and metropolitan areas. Discussed specifically the tower climb assessment and that with 36 steps assessed this would be representative of negotiating flights of stairs in many regional hospitals. As a result assessment battery likely to representative of FF tasks all areas.

(4) **Proposed Validation Trial with Incumbent Firefighters:**
BM outlined the preliminary discussions for proposed FF validation trials. Discussion amongst committee of what locations to conduct these trials at. BM asked the UOW of the most important factor with regards to the subject group, whether it was age and gender or regional and metropolitan cohorts. HG outlined that most importantly it would be age and gender representative of FRNSW personnel. JH (uow) outlined that the goal was to have 200 FF undertake the validation trial because this creates a sample size that is legally defensible. BM to put together schedule for validation trials (see action items).

(5) **ACTIONS:**
1. BM to liaise with subject matter experts regarding FF rescue under 1.25 metres and clarify.
2. BM to put together schedule for validation trials and communicate with area commanders to organise.
3. BM to liaise with FBEU to confirm current phase of research and invite to attend validation trials.

APPENDIX SEVEN: Formal responses from Fire & Rescue NSW concerning the first draft of Report Four

NSW GOVERNMENT | Fire & Rescue NSW

File Ref. No: NFB/08711
TRIM Ref. No: D13/43910

Associate Professor Nigel Taylor

Centre for Human and Applied Physiology
School of Health Sciences
University Of Wollongong
Northfields Avenue
Wollongong, NSW, 2522

13 September 2013

Dear Professor Taylor,

Fire & Rescue NSW representatives of the PAT Project Steering Committee met recently to discuss the recommendations of report four "Recommended Physiological Employment Standards for the Firefighters of Fire & Rescue NSW". As a result of discussions during this meeting, it would be appreciated if you could please provide clarification on the following aspects of the report's recommendations:

1) The application of a 5% buffer to the time standard for tasks 1-4 and 5-6 has been recommended (recommendation 2 and 3 of report pg 50-51). It is unclear from your report how a margin of 5% has been selected.

   Q: Could you please explain the methodology used to support the recommended application of a 5% buffer to the performance standard?

2) Recommendations 2-4 (pg 50-51) utilises "pass/fail" and "absolute failure" terminology. Such terminology seems incongruous with other wording provided in your recommendations: '*possible* to determine that the tasks *should* be completed within 15 minutes. It is therefore recommended that this 15-min time be *considered* as a *possible* screening threshold.....'

   Q: Could you please explain how the university has recommended the use of such definitive terminology, when the research indicates a less precise outcome?

3) Recommendation 2 (pg 50) states "Firstly those falling outside the buffer would be deemed to have absolutely failed the test at this point, and as such individuals would not be permitted to continue". The steering committee is of the opinion that with additional controls, there is minimal risk to the safety of the applicant if they are permitted to continue with the assessment. Allowing the applicant to continue the assessment would then serve to provide further familiarisation for any future participation in assessments.

   Q: Would the University agree that FRNSW could conduct a risk assessment, stipulating additional controls (e.g. pre-assessment medical clearance, assessment of ambient temperature, monitoring of fatigue levels and availability of advanced first aid) that may

enable an applicant to safely continue with the assessment despite their performance being outside the 5% buffer? If not, why not?

4) Recommendation 4 (pg 51) states that failure to complete three positional holds represents test failure, independently of the performance on any other test component. FRNSW's steering committee representatives are of the opinion that applicants who are unable to complete this strength based test, should be permitted to continue with the subsequent stages of the assessment. It has been stated earlier in your report that familiarization of the assessment may assist in successful completion in subsequent attempts.

   Q: Considering the above, could the university please provide clarification regarding the basis on which the recommendation to 'absolutely fail' is based.

5) Recommendation 4 (pg 51) states that "individuals completing the test with a failing score that falls within one of the buffer zones, should be allowed one (and only one) further attempt to pass the test".

   Q: Could the research team provide justification as to why this opportunity is limited to "one (and only one)" further attempt?

To allow this project to progress, the steering committee would appreciate a response to the above questions within two weeks of the date of this letter.

Should you require any additional information please contact Alison Donohoe, Assistant Director Health & Safety on (02) 9265 2642.

Yours sincerely

ADonohoe

Alison Donohoe
On behalf of FRNSW PAT Steering Committee

APPENDIX EIGHT: Formal responses to Fire & Rescue NSW concerning comments on the first draft of Report Four

Dear Alison,

Please excuse my delayed reply to your letter of September 13th. As you are aware, this letter arrived just a few days before I left for a trip to Germany, from which I have just returned.

Pending further discussion, I will incorporate the points noted below within the final version of our latest report: "***Physiological employment standards for firefighters: Report 4: Recommended physiological employment standards for the firefighters of Fire & Rescue NSW***".

Item 1: Clarification of the recommended 5% buffer
All testing is associated with uncertainty. For instance, whether evaluating performance on a written examination or a practical test, pass and fail criteria, though established in good faith and with all possible information, may still lack the necessary precision to allow examiners to claim with absolute certainty that an individual achieving 49% should fail, whilst another scoring 50% should pass. This uncertainty leaves the assessment tool and the examiner vulnerable to challenge, and some such challenges may be entirely justified. As a consequence, it is both fair and reasonable to those being examined for a zone of uncertainty to be recognised and applied to the assessment outcome. This approach was undertaken with regard to the pass/fail recommendations for the two groups of test items identified in our report (items 1-4 and 5-6).

In these instances, our recommendation that a 5% buffer should be applied was based upon two primary points. Firstly, the recommended acceptable work rates for some tasks were set to simulate values at the lower end of the 95% probability range, as this corresponded with values observed in operational firefighters who could still successfully complete the task, but at the lower end of the range of these physiological responses. Thus, by definition, we had recommended a performance standard that could be passed by 95% of the operational firefighters we had tested. It was, therefore, internally consistent to recommend this threshold be set at the 5% level. Secondly, since it is the widely accepted convention in science to set statistical thresholds for physiological research at the 5% level, then this value was considered appropriate in this context also.

Text to explain this position will be inserted into P 50 of the report, pending further discussion and clarification (if required).

Item 2: Pass/fail terminology
These terms were deliberately used within our report to provide FRNSW with unambiguous recommendations. This was motivated by our desire and obligation to help protect the health and safety of prospective recruits during testing. Since University staff would not be administering any of these tests, then no mechanism would exist for the University to oversee testing or participant health and safety, then we were obliged to make such

recommendations. Moreover, since we had recommended acceptable work rates for some tasks at the lower end of the 95% probability range for the relevant physiological responses, then this would mean that individuals struggling to satisfy the recommended performance standards would be those most likely to suffer adverse health outcomes during, or after, performing these tests. In this case, we had a strong ethical obligation to work towards preventing adverse outcomes for all participants, but those high-risk individuals in particular. In addition, there existed a legal obligation, for it was deemed that some legal responsibility would transfer to the University in the case of an adverse health and safety outcome, if such a recommended was not issued.

However, we agree with the comment that our use of less affirmative language in other parts of the report (recommendation ten of the executive summary and sub-section 4.2.2) could lead to confusion in this regard. We apologise for this. Accordingly, the word "possible" has now been removed from these locations of the report. Furthermore, we again highlight that it is not our role to set performance standards. Instead, it is our role to provide justifiable recommendations to FRNSW upon which it may, or may not, choose to act. Thus, the words "*recommend*" and "*considered*" have been retained.

In making these recommendations, it is acknowledged that FRNSW employs suitably qualified professionals who provide that organisation with a capacity to evaluate and implement a wide range of preventative strategies. Accordingly, the organisation may choose an alternative strategy for risk management. Notwithstanding, this capacity does not remove the need for the University to exercise caution. This is the approach that the University expects all of its scientists to take in such matters, and this has been followed in this report.

Item 3: Recommendation two: absolute test failure
We absolutely agree that test familiarisation would be useful for all participants, as noted in sub-section 4.4.5. We have fully explained (above) our position with respect to the health and safety of prospective high-risk recruits during testing. We further recognise that it is the prerogative of FRNSW to choose not to act upon our recommendation for the "*absolute test failure*" of recruits falling outside the buffer zone. However, we are ethically and legally bound retain our affirmative wording within the existing report.

Item 4: Recommendation four: absolute test failure
We absolutely agree with the position taken by the Steering Committee in this regard. Indeed, the lack of sufficient strength, or muscular endurance, to complete this testing requirement does not imply the absence of the necessary cardiorespiratory endurance to complete the rest of the test. Accordingly, the report has been revised (as follows) to reflect this change.

> "*failure to complete the three positional holds for the required duration: such individuals could be allowed to continue the test, but should still be required to pass this component as part of the complete test sequence*"

To pass the PAT, it is recommended that all individuals must complete the entire assessment whilst meeting each of the separate criteria. Such an approach is very important, since the test was not designed to be performed in isolation, but as a continuous series of activities.

Thus, some residual tiredness is carried from one test item to the next, as would occur under operational conditions, and this is a fundamental design principle of the recommended PAT and its recommended performance standards. Therefore, it is not possible to claim that individuals who satisfactorily perform test items in isolation will be able to perform operational fire fighting activities to the satisfaction of FRNSW, at least as determined during the course of this research.

However, the phrase "*absolute test failure*" has been retained. Extensive feedback from the Fire-Station visits and the firefighter survey convinced the research team that this test item was of critical importance to effective operational fire fighting.

Item 5: Recommendation two: one further attempt
There is no scientific justification for this recommendation. Indeed, it was a purely pragmatic recommendation to help FRNSW. Within the University environment, we find that such clauses are beneficial and cost effective when students are wanting to re-sit examinations. Accordingly, this suggestion was inserted with a view to helping FRNSW in a similar situation. However, should FRNSW agree, then we could modify this text to read as follows: "*It is recommended that individuals completing the tests in a timeframe that falls within the 5% buffer zones for either of the test items groups (i.e. 1-4 or 5-6) should be identified as candidates for whom retesting might be appropriate.*"

I trust that these responses provide full and adequate information to the Steering Committee.

Your sincerely,
Nigel.

APPENDIX NINE: Formal responses from Fire & Rescue NSW concerning responses to comments on the first draft of Report Four

Hi Nigel,

As usual it was great to talk with you this morning.

Report 4
As discussed the FRNSW PAT Project management team met last week to discuss the UOW response to the 5 questions relating to report 4.

Each of the responses were considered and the following agreed upon (bold UOW actions):
1) All accepted the University's recommended use of a 5% buffer and agree for the text to be inserted into the report explaining this as per Professor Nigel Taylors' letter.
2) All accepted UOW will retain 'pass' and 'fail' terminology in this report, however it is recommended the application of the standard by FRNSW will use the terminology 'achieved' and 'not yet achieved'.
3) All accepted UOW will use absolute fail terminology in the final report. However it is recommended the application of the standard by FRNSW will allow applicants to complete the entire PAT despite not yet achieving the standard for particular tasks. Applicant suitability for participation will be assessed prior to the PAT and fatigue monitored throughout the assessment with appropriate safety controls implemented at the assessment site.
4) All accepted the changes UOW will make to the report regarding applicants being able to continue with the rest of the PAT if they do not yet achieve the standard for the shear holds.
5) All accepted the concept of the buffer as a means to identify individuals who are more likely to achieve the standard if invited back to have another attempt at the PAT. They agreed that the recommended modifications to the text be made.

Once these changes are made report 4 will be finalised.

Report 5
As per our discussions earlier today FRNSW Health and Safety have already performed a body of work to develop both a PAT Applicant Preparation Guide and a PAT Assessment Manual. From discussions with you I understand that report number 5 was going to include both of these items. I have spoken with Alison and she agreed that if you (herb and John) could review these documents and provide us with any recommendations then report 5 would not be required.

Due to file size I will send these documents in separate emails.

Regards
Brendan Mott AEP
Manager Health Promotion
Health & Safety Branch