Natural Hazards Research Australia

SES Fit for Task
Final Research Report

Human Performance Science

bushfire&natural HAZARDSCRC

human performance science

afac

ACSES Australian Council of State Emergency Services

Version	Release history	Date
1.0	Initial release of document	24/08/2023

Australian Government

Natural Hazards Research Australia receives grant funding from the Australian Government

We acknowledge the traditional custodians across all the lands on which we live and work, and we pay our respects to Elders both past, present and emerging. We recognise that these lands and waters have always been places of teaching, research and learning.

Publisher:
Natural Hazards Research Australia

ISBN: 978-1-923057-00-5

Report number: 18.2023

August 2023

Citation: Human Performance Science (2023). *SES fit for task – final project report.* Natural Hazards Research Australia. accessible at www.naturalhazards.com.au/fitfortask

Cover photo: South Australian State Emergency Service.

Table of contents

List of abbreviations

ABBREVIATION	DEFINITION
FFT	Fit for Task
HPS	Human Performance Science
AFAC	Australasian Fire and Emergency Service Authorities Council
BNHCRC	Bushfire and Natural Hazards Cooperative Research Centre
SES	State Emergency Service
ACT	Australian Capital Territory
VIC	Victoria
NSW	New South Wales
QLD	Queensland
TAS	Tasmania
SA	South Australia
WA	Western Australia
NT	Northern Territory
SME	Subject matter expert
HR	Heart rate
GPS	Global positioning system
m	Metres
cm	Centimetres
Kg	Kilograms
N	Newtons
CAT	Category
PPE	Personal protective equipment

Operational definitions

TERM	DEFINITION
Fit for Task	Physical and physiological capability to complete job roles.
Programme	The complete integration of newly developed physical assessments.
Member	Person undertaking Fit for Task assessments. Includes current member or new applicant.
Skill Set	Category of SES response.
Core task	Physically demanding tasks that represent the underlying duties of a skill set.
Criterion task	Essential tasks that are physically demanding, frequently occurring and operationally important.
General Rescue	Generic SES role encompassing fundamental SES tasks applicable to all members.
Storm Damage	Skill set – response to all storm and flood events.
Chainsaw Operations	Skill set – management and clearance of fallen trees, branches or debris.
Land Search & Rescue	Skill set – urban or rural walking searches for missing person(s).
Air Search	Skill set – aeroplane or helicopter searches for missing person(s).
Vertical Rescue	Skill set – cliff or steep angle technical assistance for injured or trapped person(s).
Road Crash Rescue	Skill set – response to vehicle accidents, including the removal of personnel from vehicles.
Urban Search & Rescue	Skill set – Searches for person(s) or items amongst rubble.
Boat Operations	Skill set – response to marine incidents including stranded or missing person(s).
Land-based Swiftwater	Skill set – response to marine incidents from shore or riverbank.
In-Water Technician	Skill set – advanced Flood Boat Operations tasks in more treacherous marine environments.
Off Road Driving	Skill set – capacity to drive off-road vehicles during searches of missing person(s).
Firefighting Air Base Support	Skill set – assistance of firefighting agencies during aerial fire operations.
Level 1	Baseline level for a skill set.
Level 2	More advanced level for a skill set.
Level 3 or 4	Most advanced level for a skill set.
Operational Duties	Field-based duties performed by SES members during responses.

Acknowledgements

Human Performance Science acknowledges the widespread and valuable contributions made by all the SES staff and volunteers who assisted with the project over several years. The project has required large scale planning and coordination. Without the SES state and territory representatives, in addition to their regional counterparts, the management of the project would not have been possible. Without the volunteers giving up hours, sometimes days, of their time, no data would have been collected. This project has involved thousands of people from across the country. We express gratitude to each and every person who has contributed.

Executive summary

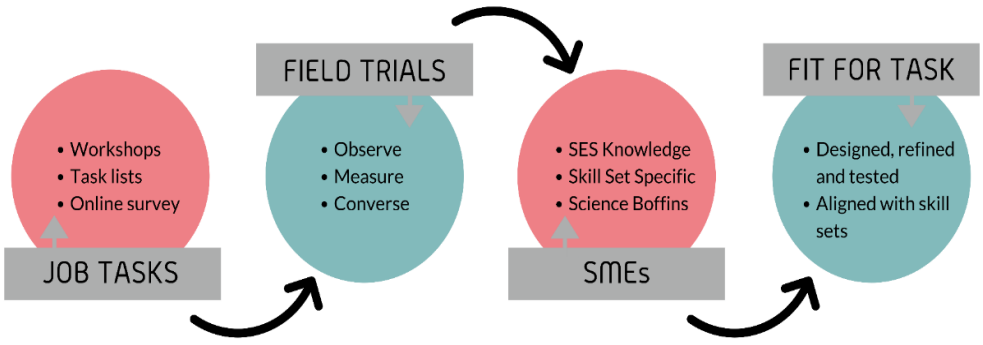
In Australia, State Emergency Service (SES) organisations provide emergency help during and after declared disasters. The SES is also the primary or secondary agency for emergencies, such as storm damage, flood damage, building damage, traffic hazards and road crash rescue. More than 40,000 volunteers are deployed across the state and territory jurisdictions, fulfilling a range of roles.

Given the physically demanding nature of SES roles, there is a need to develop suitable occupational health and safety strategies, which amongst other things, optimises the health of all members. One such strategy is the physical screening of volunteers, matched against the demands of the role. The approach of mitigating injury risk, through physical screening practices, is colloquially referred to as 'Fit for Task'.

The aim of the SES Fit for Task project was to objectively measure physical and physiological demands of a variety of SES roles, to establish physical screening measures, in the form of assessments, for SES personnel. The purpose of the assessments is for SES personnel to demonstrate they have the baseline physical competencies to meet the physical demands of SES tasks.

This report outlines the findings of the SES Fit for Task project, which included national buy-in from all Australian SES agencies, with support from the Australasian Fire and Emergency Service Authorities Council (AFAC), the Bushfire and Natural Hazards Cooperative Research Centre (BNHCRC) and volunteer associations. Over several years, more than 3000 SES volunteers from across Australia, representing a diversity in age, experience, sex, and qualifications, participated in the research.

The SES Fit for Task project employed a robust scientific process, delivered by Human Performance Science. This process reflects gold-standard practices which first required identifying and measuring the physical demands of the various employment roles. This is referred to as a job task analysis. Following the job task analysis, a series of steps were taken to develop prototype physical assessments, before verification steps were utilised amongst subject matter experts from the incumbent population. The following diagram summarises the general phases of the research and the considerations at each stage.



From the early stages of the job task analysis, researchers identified 13 key SES responses that were carried forward for analysis. These responses, which were referred to as 'skill sets', which included (in alphabetical order): Air Search, Boat Operations, Chainsaw Operations, Firefighting Air Base Support, General Rescue, In-water Technician, Land Search & Rescue, Land-based Swiftwater, Off-road Driving, Road Crash Rescue, Storm Damage, Urban Search & Rescue, Vertical Rescue.

Each skill set contained a number of tasks identified through early discussions with subject matter experts. This resulted in a comprehensive task list, inclusive of 209 tasks. Through a series of filtering processes, quantitative and qualitative analyses and verification checkpoints, researchers identified 67 criterion tasks, which represent the tasks most

physically demanding, frequently occurring and operationally important. From the criterion tasks, the research team could group tasks with similar characteristics to arrive at a representative battery of assessments that captured all movement types and physiological demands for all skill sets.

As a result of the research stages, nine new SES Fit for Task assessments have been developed. Each assessment represents an evidence-based screening measure modelled on the functional demands of SES roles. The nine assessments include seven land assessments (LIFT & SHIFT, LADDER CLIMB & LIFT, CARRY, HIKE, DRAG, HAUL, HOLD) and two water assessments (IN-WATER SAFETY, POWER SWIM & RESCUE). All SES Fit for Task assessments are sex and age neutral. That is, the same standard applies regardless of sex or age, rank or seniority. If an SES member undertakes operational activities, they would ordinarily be expected to be able to meet the minimum physical fitness levels required for that operational role.

The new assessments, which represent minimum physical standards, are designed to improve the health and safety of all volunteers, by demonstrating they can meet the inherent demands of the role. Not all volunteers need to complete all nine assessments. Volunteers will only need to complete a subset of the nine assessments, based on which skill sets they currently hold. Similarly, the performance level to which volunteers will need to complete the assessment is matched to the skills sets they hold.

Due to the effect of COVID-19, the later stages of the research, validation and reliability testing, saw fewer than desirable participant numbers. Though SES agencies can remain confident that the assessments represent scientific and functional representations of SES tasks, it is recommended that additional data is collected during 'soft implementation' efforts, to provide additional confidence on the performance standards of assessments which include a time-to-completion standard.

Through successful implementation efforts, SES agencies can be confident in a programme that identifies injury risk and prevents hazards or injuries before they occur. A volunteer workforce that is more functionally fit, that is, more capable of meeting the physical demands of roles, is beneficial to the individual, the organisation, and the community.

This research was completed initially through the Bushfire and Natural Hazards CRC and then Natural Hazards Research Australia.

1 Introduction

1.1 Background

In Australia, State Emergency Service (SES) organisations provide emergency help during and after declared disasters. The SES is also the primary or secondary agency for emergencies, such as storm damage, flood damage, building damage, traffic hazards and road crash rescue. In other scenarios the SES may provide a support role to other agencies, particularly police and fire. Every state and territory in Australia has its own SES organisation, with more than 40,000 volunteers deployed across the country.

Given the physically demanding nature of SES roles, there is a need to develop suitable occupational health and safety strategies (Taylor & Groeller, 2003). One such strategy is the physical screening of volunteers, matched against the demands of the role. The approach of mitigating injury risk, through physical screening practices, is colloquially referred to as 'Fit for Task'.

The SES Fit for Task project aims to objectively measure physical and physiological demands of a variety of SES roles, to establish physical assessments for SES personnel. The purpose of the assessments is for SES personnel to demonstrate they have the baseline physical competencies to meet the physical demands of SES tasks.

The SES Fit for Task project was established in 2014. In an earlier iteration of the Project, researchers from Human Performance Science and Deakin University measured the physical and physiological demands of a subset of SES roles in Victoria and New South Wales. The outcome from this project highlighted the need for further investigation of physical demands across all SES roles, amongst all Australian states and territories. The current SES Fit for Task project builds on this earlier research.

1.2 Purpose of report

The current report outlines the research process undertaken as part of the SES Fit for Task project. The report describes the procedures of the research and the outcomes of each stage. The report does not report on the implementation, policy decisions or change management work that is necessary for the integration of the newly developed SES Fit for Task programme. Details on these are outlined in supporting documentation.

1.3 Fit for Task research

Fundamentally, Fit for Task programmes help identify individuals at risk of injury. A healthy workforce, with minimal injuries, can promote business compliance, reduce lost time injury frequency rate (LTIFR), improve workplace productivity and efficiency, and reduce compensation pay-outs and insurance premiums. Furthermore, the implementation of Fit for Task programmes ensures employers comply with their duty of care and legislative requirements.

The Commonwealth Government of Australia has laws designed to protect workers from workplace discrimination, including during recruitment by prospective employers (Australian Human Rights Commission, 2015). At a federal level, examples of these laws include the *Fair Work Act 2009*, *Age Discrimination Act 2004*, *Australian Human Rights Commission Act 1986*, and the *Disability Discrimination Act 1992* (Australian Human Rights Commission, 2015). In addition, various anti-discrimination and equal opportunity laws operate at a state and territory level.

Physical assessments, such as those within the SES Fit for Task Programme, ensure members demonstrate the ability to perform job tasks, without lawful discrimination. For physical assessments to be lawful, the organisation must be able

to provide evidence the measures accurately reflect the inherent or essential requirements of the job. If properly developed and conducted, physical assessments can be a useful and legally defensible way of determining prospective employees' suitability for a role.

For physical assessments to be evidence-based and legally defensible, they must be developed through a rigorous methodological process. A number of prominent research groups have developed frameworks for this methodology. For an example, please see the 12-step process by Jamnik (2012), illustrated in Table 1.

Step	Components of the Occupational Fitness Standard Research Process
1	Form a project management team including all stakeholders
2	Become familiar with all of the job description and associated requirements
	Conduct a physical demands analysis
4	Establish a representative rank-ordered subset of the critical physically demanding and frequently occurring on-the-job tasks
5	Characterise the subset of the most critical physically demanding and frequently occurring tasks
6	Develop a draft physiological employment standard based on the critical physically demanding and frequently occurring tasks then pilot test and refine the physiological employment standard with job incumbents
7	Establish a standardized, objective assessment procedure for administering the physiological employment standard
8	Establish the scientific accuracy (validity and reliability) of the physiological employment standard
9	Develop employment standards for the physiological employment standard
10	Evaluate the results of applying the physiological employment standard then address any adverse impact and the possibility of accommodation
11	Implement the physiological employment standard
12	Maintain an ongoing review of the effectiveness of the physiological employment standard

TABLE 1: ACCEPTED RESEARCH PROCESS FOR DEVELOPING PHYSICAL SELECTION TESTS (JAMNIK, 2012).

Table 1 highlights the need for a sequential and thorough process. Broadly, the first series of steps requires identifying and measuring the physical demands of the various employment roles. This is referred to as a job task analysis (JTA) (Rayson, 1998). The JTA process involves a series of subjective techniques, including focus groups, panel discussion and questionnaires (Larsen & Aisbett, 2012), and objective techniques, including the quantification of physiological (e.g., heart rate, force output) and physical (e.g., equipment mass, height of lifts, distance equipment is carried) job parameters (Taylor & Groeller, 2003).

From this point, a series of steps are taken to develop prototype physical assessments, before validation and verification of these assessments within the incumbent population (Jamnik, 2012; Payne & Harvey, 2010). These final steps require the careful analysis of data collected during the JTA process, in combination with discussions with the end user, whilst accounting for practical limitations and organisational desires.

The core tenet of the research process is specificity. Organisations should ensure their Fit for Task programme is specific to their workforce. There are no off-the-shelf assessments or standards for any single occupation (Petersen et al., 2016);

organisations should be dissuaded from borrowing or adapting assessments. Fit for Task programmes should demonstrate that the assessments and standards were derived from methodologically sound observation and measuring of a diverse cross-section of employees (or volunteers) capable of performing the job safely and efficiently (Adams, 2016). The closer the nexus between the physical assessments and the inherent requirements of the job, the greater the mitigation of injury risk for current and prospective employees (or volunteers).

Finally, it must be acknowledged that even through rigorous and meticulous research, assessments will never reflect the true demands of occupational roles. There will always be a level of error with the assessments and their associated standards, given the diversity of job tasks, volatile working conditions and subjective processes employed throughout stages of the research. However, the purpose of the research is not to arrive at a perfect battery of assessments and standards. Rather, organisations should aim to implement a programme that reduces the likelihood of false positives (people who can perform the assessment but may be incapable of the job) and false negatives (people who cannot perform the assessment but are capable of performing the job) as far as reasonably practicable. The closer the assessments can align with job demands, the more confident organisations can have in an effective Fit for Task programme.

1.3.1 Assessment standards, sex and age

With every assessment, there is an associated performance standard. Standards can take many forms, for example time to completion, mass, or distance. Regardless of the nature of the standard, standards should reflect the *minimum* requirements of performing a job safely and efficiently (Adams, 2016). Minimum standards reflect the performance that should be expected of all employees (or volunteers), at a minimum.

As demonstrated in Table 1, the development of standards must take place through the robust scientific analysis of job tasks. Only in understanding the inherent levels of physical exertion can suitable standards be conceptualised. Without linking standards to the physical demands of the tasks, assessments cannot be considered legally defensible and may be classified as discriminatory.

The issue of discrimination is relevant to the act of scaling standards to suit demographics. Most commonly, scaling is considered when accommodating for workers of different ages or sexes. When considering age, it is well established there exists a decline in physiological output as adults become older. Cardiovascular fitness, muscular strength and endurance all decrease for every decade an adult ages (Kenny et al., 2016). Consequently, the argument is made that assessment standards should accommodate the differences between younger and older workers.

Such an argument is only valid if there are demonstrated differences in job requirements between older and younger workers. If workers of all ages are required to perform the same job roles, regardless of physical difficulty, the softening of standards for older populations is inappropriate, and discriminatory against younger workers. To avoid ageing workforces that might progressively become incapable of performing physically demanding tasks, assessment standards should remain age-neutral (Petersen et al., 2016).

A similar argument applies to accommodating standards to suit females. Like age-differences, physiological disparities exist between males and females. Males typically have superior physiological output to females, including cardiovascular output, muscular strength and endurance (Roberts et al., 2016), arising from genetic and anatomic differences. However, again the case for scaling standards to accommodate these differences, if not justified by differences in task performance, is pernicious and discriminatory in nature (Roberts et al., 2016). Sex-neutrality should be maintained when the standard is linked to physical demands common to all workers (Petersen et al., 2016). The best way to protect against questions relating to the bias and fairness of the standards is to ensure the test design is a true representation of the critical job-related task requirements (Roberts et al., 2016).

1.4 Supporting documentation

Table 2 illustrates the supporting documentation to the Fit for Task Research Report.

DOCUMENT	DESCRIPTION
SES Fit for Task METHODS_Infographic	A one-page visual summary of the research methodology.
SES Fit for Task ASSESSMENTS_Infographic	A one-page visual summary of the newly developed assessments.
SES Fit for Task Assessment Instructions Booklet	Individual instructions for each of the newly developed assessments, in addition to data recording sheets.
SES Fit for Task data and data processing roadmap	An MS Excel spreadsheet containing the data and data process from the job task analysis.
SES Fit for Task Implementation Strategy	Describes the practical and logistical implication of the SES Fit for Task programme.

TABLE 2: SUPPORTING DOCUMENTATION TO THE FIT FOR TASK RESEARCH REPORT.

1.5 Project personnel

The SES Fit for Task project has been a collaborative effort between multiple parties.

RESEARCH PROVIDER

Organisation: Human Performance Science

Responsibility: Delivery of the scientific process in the development of physical screening measures.

Personnel: Robbie Savage, Aaron Silk, Jared Bailey, Georgia Verry

PROJECT TEAM

Organisations: Australasian Fire and Emergency Service Authorities Council (AFAC). Bushfire and Natural Hazards Cooperative Research Centre (BNHCRC), Victoria State Emergency Service (Vic SES). State Emergency Service National Volunteers Association (SES NVA)

Responsibility: Project management, engagement with all stakeholders, project communication.

Personnel: Terese Howlett (Vic SES), John Bates (BNHCRC), Sandra Lunardi OAM (AFAC), Melissa Peppin (AFAC), Brendan Corboy (Vic SES), Faye Bendrups (SES NVA).

STATE & TERRITORY REPRESENTATIVES

Organisations: Victoria State Emergency Service (Vic SES), New South Wales State Emergency Service (NSWSES), Australian Capital Territory State Emergency Service (ACTSES), South Australia State Emergency Service (SASES), Northern Territory Fire, Rescue and Emergency Services (NTFRES), Tasmania State Emergency Service (TASSES), Department of Fire and Emergency Services (DFES), Queensland Fire & Emergency Services (QFES)

Responsibility: Representation for each state or territory. Coordination of local research. Communication with Research Provider and Project Team.

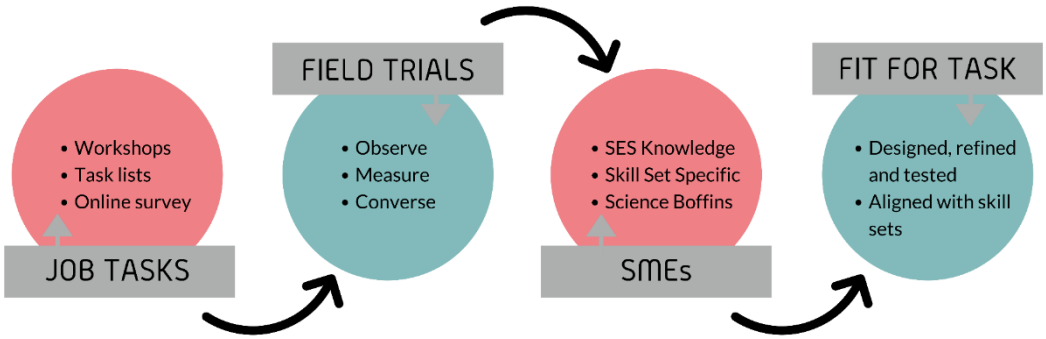
Personnel: Brendan Corboy, (VICSES), Robbie Landon (NSW SES), Lucinda Machelle (NSW SES), Graham Ible (ACT SES), Trevor Arnold (SA SES), Dermot Barry (SA SES), Mark Fishlock (NTFRES), Leon Smith (TAS SES), Chris Kin-Maung (DFES), Eammon Lennon (QFES), Brian Cox (QFES), Glenn Alderton (QFES).

1.6 Research overview

In the development of an evidence-based Fit for Task Programme, there is a need to study the inherent physical demands of job tasks (Taylor & Groeller, 2003), before making informed decisions in the development of assessments. The SES Fit for Task project was designed to follow a series of scientific steps to ensure all SES agencies could be confident in a product that is designed to suitably screen the physical capability of volunteers, as safely as possible. Broadly, the SES Fit for Task project will take place over four broad research stages:

Stage 1 – Job task analysis
Stage 2 – Development of assessments
Stage 3 – Validation
Stage 4 – Implementation

Developing a user-friendly and implementable physical assessment programme, which spans all SES jurisdictions across the nation and encompasses a range of SES responses, requires a comprehensive and collaborative process. At all stages of the research process, engagement was sought amongst SES subject matter experts (SMEs) and volunteers. The flowchart below demonstrates the balance between scientific input from the research provider and input from the end-user (SES) along different steps of the research process.

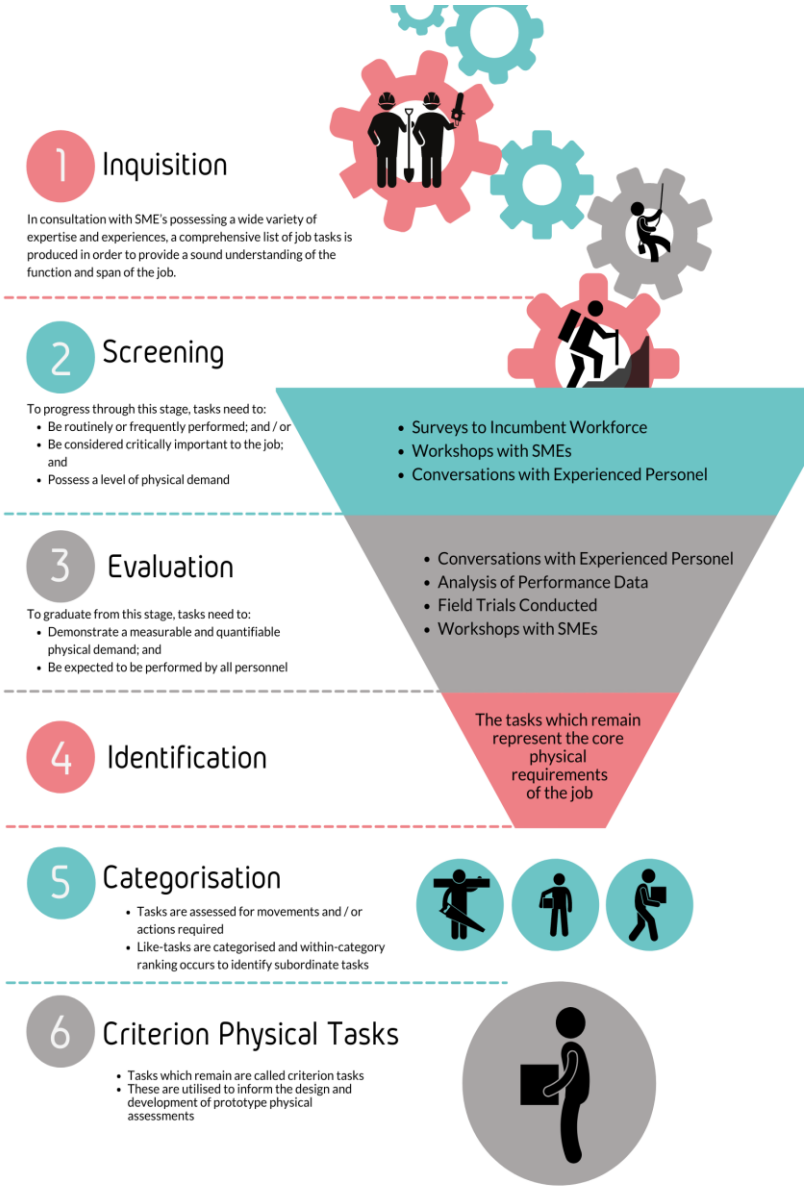


The following sections of the report outline the four stages of the research process, describing the methods that were employed and all salient results.

2 Job task analysis

The job tasks analysis represents the foundation for any scientifically developed Fit for Task programme. A thorough and detailed understanding of job tasks is necessary to make evidence-based decisions on the development or amendment of physical assessments. In this context, a job task analysis is defined as a series of sequential research processes that employ subjective and objective techniques, to identify all physical tasks and filter down to a list of **criterion tasks**. Criterion tasks are the essential tasks that are physically demanding, frequently occurring and operationally important. For a physical assessment to be valid and evidence based it must be explicitly linked to a criterion task. To identify and define a criterion task, a distinct and specialised sequence of events must occur, with evidence to support findings and decision-making collected along the way.

The job task analysis utilised for the SES Fit for Task Project included a thorough process, combining a range of data collection, verification, and analysis techniques. The first phase was the subjective job task analysis, which involved the collection of subjective data from workshops, conversations, and surveys. The second phase was the objective job task analysis, which required the collection of physical and physiological data from job tasks. The full treatment and analysis of data during both phases of the job task analysis was managed through a six-step process, referred to as the **"funnel and filter".** The following page illustrates this process graphically, with the pages that follow elaborating on the funnel and filter process.



The INQUISITION phase of the funnel and filter process required individual workshops with states and territories. A total of 73 SES subject matter experts (SMEs) attended the workshops. Through a series of conversations interrogating the tasks performed by SES volunteers, the workshops resulted in a global task list which included 178 tasks. These tasks covered all states and territories, however, not every state and territory performed all 178 tasks. Figure 1 illustrates the number of tasks relevant to each state and territory. When examining common tasks, 83 of the 178 tasks were performed by every state and territory.

FIGURE 1: NUMBER OF TASKS IDENTIFIED IN EACH STATE/TERRITORY THROUGH THE WORKSHOPS HELD AS PART OF THE INQUISITION PHASE OF THE FUNNEL AND FILTER RESEARCH PROCESS. NOTE: THE ACT HAD THE TASK LIST AS NSW (158 TASKS).

The tasks revealed during the workshops were examined by researchers to determine appropriate skill sets that could be defined to adequately capture all the tasks. In total, 13 individual skill sets were established. The names decided upon were designed to capture as many of the states and territories as possible. They will not necessarily capture every name that a state or territory will use. Additionally, names were left as broad descriptors, which were designed to incorporate specialty roles within that category. For example, Land Search & Rescue will capture specialty roles such as Alpine Rescue. The decision to keep names broad was made since the more specialised a role becomes, the more discrepancies there are in the names assigned to that role between states and territories. The following skill set names were decided upon.

- Air search
- Boat operations
- Chainsaw operations
- Firefighting air base support
- General rescue
- In-water technician
- Land-based swiftwater
- Land search & rescue

- Off-road driving
- Road crash rescue
- Storm damage
- Urban search & rescue
- Vertical rescue

Following the workshops, further conversations amongst SMEs and the Project Team led to the addition of new tasks that were either omitted from the workshops, or it was agreed that the tasks could be duplicated for a different skill set. Following these conversations, the final task list included 209 tasks.

The final task list was imported into **surveys**, which were then distributed to SES volunteers across Australia. The surveys represent the SCREENING phase of the funnel and filter process. The surveys asked volunteers to rate the physical difficulty, frequency and operational importance of all the tasks they perform. This resulted in 2187 partially complete and 866 complete responses. As a secondary step in the screening process, HPS researchers presented the task list to states and territories to verify tasks and seek endorsement.

This consultation with the incumbent workforce enabled the initial comprehensive task list to be reduced to a list that represented the most important, frequently performed and physically demanding tasks. The process of distilling a comprehensive task list into one which represents a more refined and evidence-based list provided the pathway to progress to **field trials**, which required observing, monitoring and measuring performance of a more appropriate and targeted subset of tasks.

Field trials additionally provided a valuable opportunity to converse with experienced personnel to further understand what was critical to job performance and what, either as an individual or within a small team, they'd expect from each other. Collecting evidence through conversation and developing an understanding of the realities of the job was vital to the research process and the decisions made as the project progressed.

The EVALUATION step allowed researchers to filter tasks, to eliminate any tasks that were low in quantifiable physical demand, were not mandatory for all volunteers to perform, were rarely performed (or being phased out), or that were largely skill-based rather than physical (see Appendix C for this process). At this point, tasks were assigned a 'progress' or 'reject' label. The tasks that were labelled 'progress' were continued to the IDENTIFICATION stages, representing core tasks (Physically demanding tasks that represent the underlying duties of a skill set.)

From this set of core physical tasks, researchers could categorise each task into the "task performance features" that best describe how the task is performed. This represented the CATEGORISATION stage. In situations where there was more than a single task in a task performance feature category, a task comparison was conducted in order to establish those tasks that were superior in nature (and hence their performance can predict that of their subordinates). Consequently, the superior or dominant tasks were termed 'CRITERION TASKS', and it is these tasks that are progressed and considered when developing **Fit for Task** assessments. Figure 2 illustrates the number of tasks at each stage of the funnel and filter process.

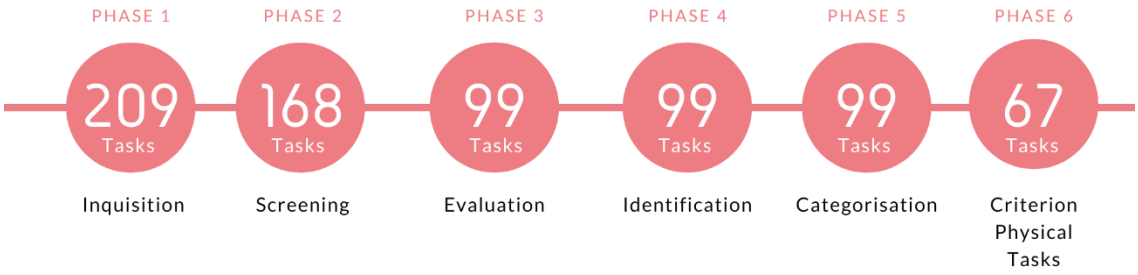


FIGURE 2: PROGRESSION OF TASKS THROUGH THE JOB TASK ANALYSIS, USING A FUNNEL AND FILTER PROCESS.

SKILL SET SUMMARY

The sections that follow describe the funnel and filter research process for each of the 13 skill sets by providing a snapshot of the outcomes at each of the 6 stages within the funnel and filter process. This snapshot highlights the number of tasks identified at each phase and provides a brief description of why tasks may have been omitted. These sections are not intended to provide the evidence for why each and every task is, or is not, progressed through each phase. Rather, these sections help to describe the overall journey each skill set has been on, from beginning with a comprehensive task list to identifying and defining the criterion tasks that guided the development of role-specific assessments. To this end, some of the skill sets will present different data to other skill sets, since there is more applicable data for certain stages of the funnel and filter process. Additionally, some skill sets will include commentary regarding information collected outside of the funnel and filter process. This information relates to information provided by SMEs regarding specific tasks that could not be resolved during the job tasks analysis process.

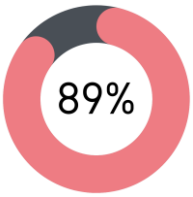
Finally, for many of the objective parameters of tasks which have been described, such as the mass of an item, length of a carry, height of a platform etc, these metrics have been determined by SES staff participating in the research. At the start of each session, researchers requested that common and representative equipment and task parameters be utilised. The Units involved in the research selected the equipment, location, vehicles, platforms etc that were most applicable to common or typical tasks.

2.1 General rescue

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the General Rescue role. In total, there were 11 tasks ranging from “Start petrol engine using pull start” to “Participate in stretcher carry of rescued person(s), carrying one side or corner of the stretcher”.

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 6 tasks remained to progress into the next phase. Examples of tasks that were excluded include “Use a sledge hammer to break concrete (or other materials) into smaller pieces” and “As part of animal rescue, remove an entrapped animal using roping or manual handling techniques”.



89% of qualified personnel agreed that participating in a stretcher carry was the most physically demanding General Rescue task.

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. The 6 tasks progressing from the Screening phase were considered to be performed comparably (e.g., similar tools, distances etc.) in other skill sets, such as Chainsaw, Land Search & Rescue and Storm Damage. Therefore, no dedicated General Rescue field trials were conducted, rather, physiological, observational and conversation data collected throughout field trials from the skill sets that perform these 6 tasks were considered relevant to General Rescue.

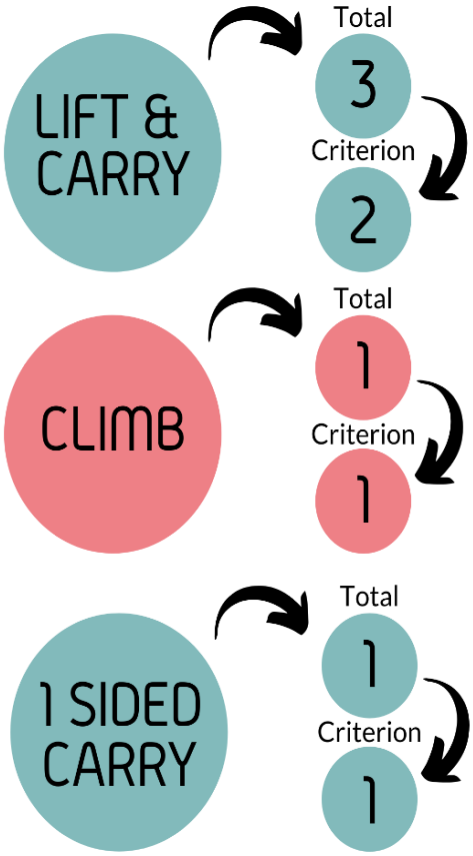
Of the 6 tasks progressing from the Screening phase, 1 task was rejected during the Evaluation phase. Broadly, this task was found to be low in quantifiable physical demand and rarely performed and therefore didn't warrant further progression. This task was "Rescue/remove an entrapped person from an urban site, such as a building, warehouse, or house".

At the conclusion of the Evaluation phase, 5 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION and ⑤ CATEGORISATION

The 5 tasks progressing from the Evaluation phase reflect the core physical requirements of the General Rescue skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart.

The lift and carry task "Carry equipment, such as ladders, pumps, cutting and spreading tools, generators and other equipment" was considered a 'core' physical requirement of the skill set, however, was not categorised as a 'criterion' task. Largely, this task was considered less physically demanding and was performed less frequently than the other lift and carry tasks.



At the conclusion of the Categorisation phase, a total of 4 criterion tasks spanning 3 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

The 4 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified General Rescue volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck and also around the worksite, climb a ladder where the job necessitates and also restow the ladder and participate in a stretcher carry where required.

2.2 Air search

① INQUISITION and ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Air Search role. In total there were 6 tasks ranging from "Assist in loading aircraft with food, water and supplies" to "During a helicopter search, manually extend and retract the winch boom during flight".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, no job tasks remained to progress into the next phase. This was largely attributable to the tasks receiving ratings of low importance, or very low physical demand. Interestingly, the task “Perform long duration air searches, looking for people or artefacts from the window of an aircraft” was the only task rated as being very important/critically important to the Air Search role.

0

③ EVALUATION

Based on conversations with SMEs, there was 1 task identified that required further investigation as a potential core physical requirement. Although the task “Perform long duration air searches, looking for people or artefacts from the window of an aircraft” received a "REJECT" outcome from the survey, based on further investigation it was considered critical to the role, requiring a unique posture (prolonged sitting) and high-level concentration.

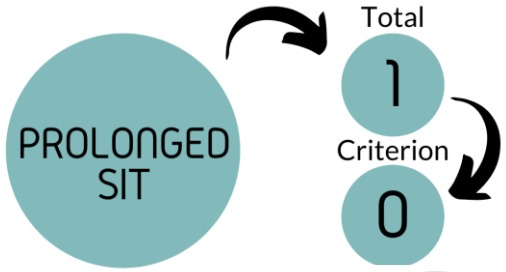
1

After the Evaluation phase, this task progressed to the Identification phase as a core physical requirement of the Air Search skill set.

④ IDENTIFICATION and ⑤ CATEGORISATION

The task progressing from the Evaluation phase reflects the core physical requirement of the Air Search skill set.

Although this task is identified as being core to the physical requirements of the skill set, it is not considered a criterion task as it does not present with quantifiable physical demands (and is undeniably cognitive rather than physical).



⑥ CRITERION PHYSICAL TASKS

This skill set contains no criterion tasks that represent a quantifiable physical demand requiring a physical assessment.

0

2.3 Boat operations

① INQUISITION and ② SCREENING

Through the interrogation of SES documentation and during workshops held in each state and territory with qualified and experienced SMEs, a comprehensive task list was generated encompassing the role. In total, there were 27 tasks ranging from “Conducting a casualty rescue” to “Loading stores onto the vessel”.

27

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 25 job tasks remained to progress into the next phase. The two tasks which were excluded were “Assist members of the public getting into and out of vessels” and “Climb ladders at jetties and marinas”.

25

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of two separate simulation trials is presented here.

During four typical Boat Operations rescue scenarios, including rescue of a 25-kg casualty from a rigid hull boat (S1), a 50-kg casualty from an inflatable boat (S2), a 75-kg casualty from a rigid hull boat (S3) and a 75-kg casualty from an inflatable boat (S4), volunteers were observed working at a moderate to high-intensity, however, this was dependent on the size of the casualty and type of boat (Figure 1-3). Each casualty was a human-sized mannequin utilised by SES in training environments. The 75-kg mannequin represents an adult and the 25-kg represents a child.

When considering the physicality required to complete various essential Boat Operations tasks, pulling a 75-kg casualty into a rigid hull boat requires the greatest amount of strength, followed by lifting the rear of an inflatable boat in a team of six (Figure 4).

In an additional Boat Operations simulation activity, volunteers were observed working at up to 75% of their maximum age-predicted heart rate (220-age; Shookster et al., 2020) when performing general boating tasks, such as manually rowing, dropping and retrieving the anchor and launching/recovering the boat (Figure 5).

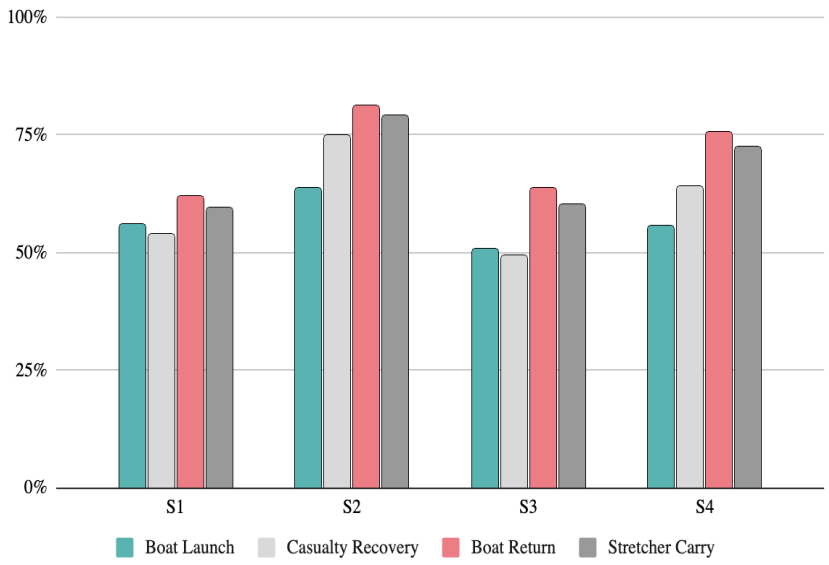


FIgure 1: Average percentage of maximum heart rate attained across four different boat operations scenarios. all scenarios comprised 3 participants and ranged from 10 to 22 minutes in duration.

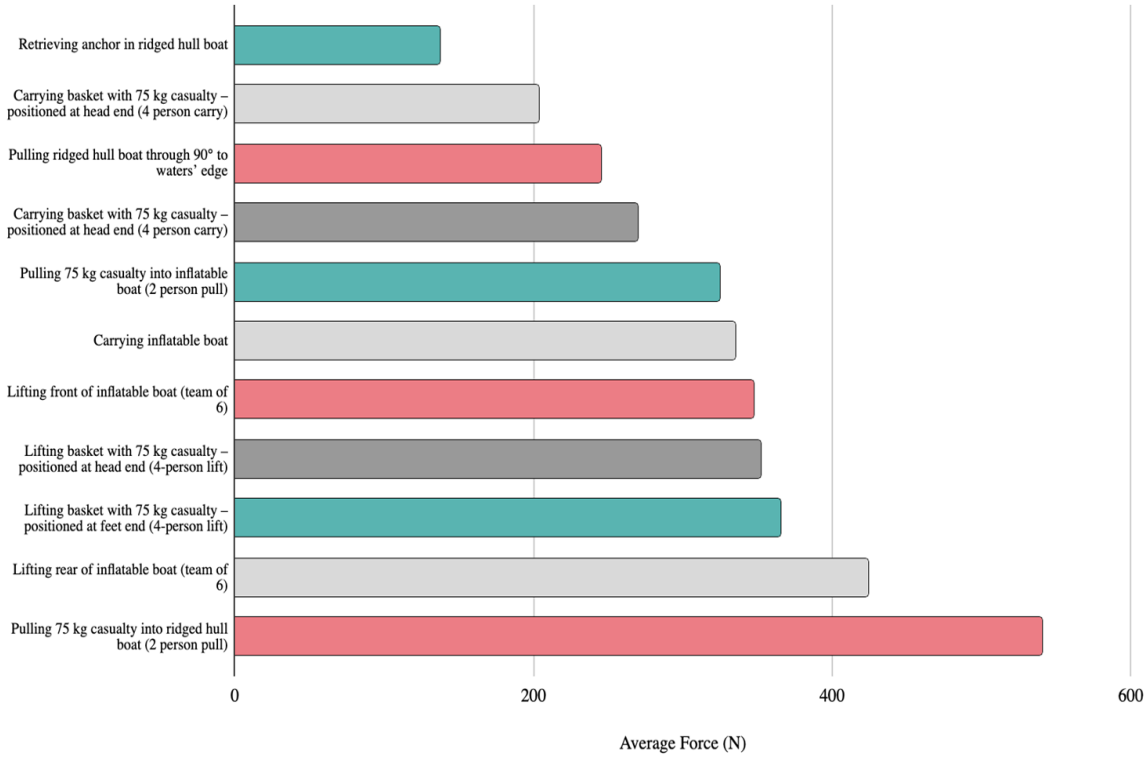


Figure 3: Average amount of force, in newtons, required to complete various boat operations tasks.

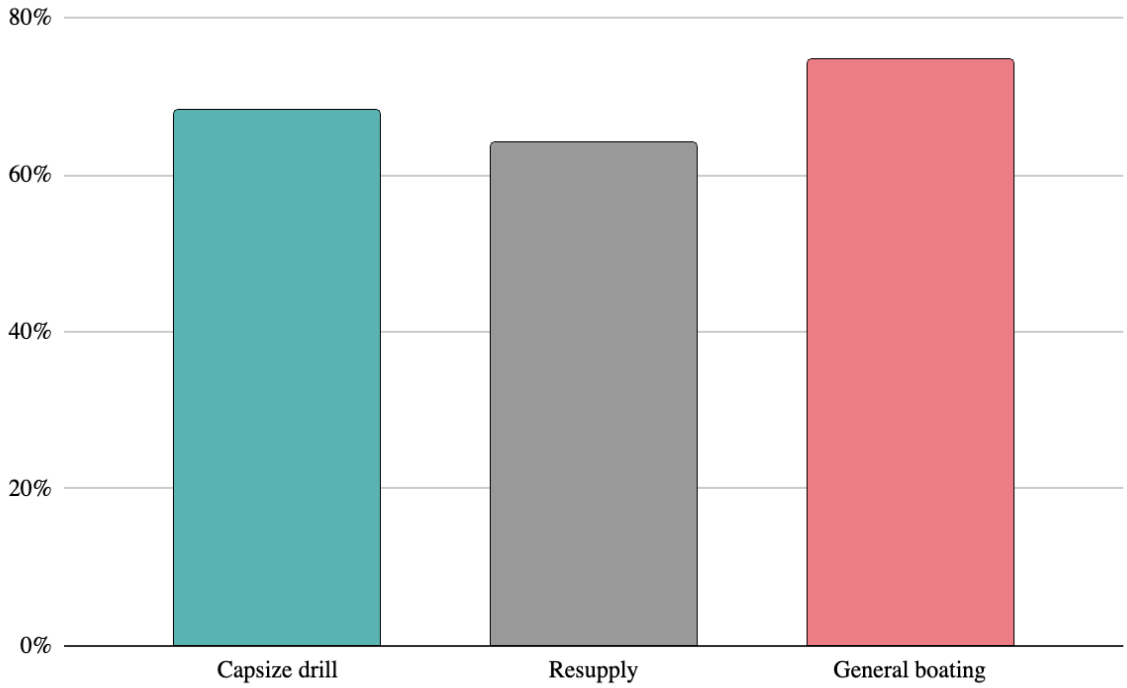


Figure 2: Average percentage of maximum heart rate attained during a capsize drill (n=13), resupply drill (n=18) and general boating activities (n=12). Each activity ranged between 25 and 30 minutes in duration.

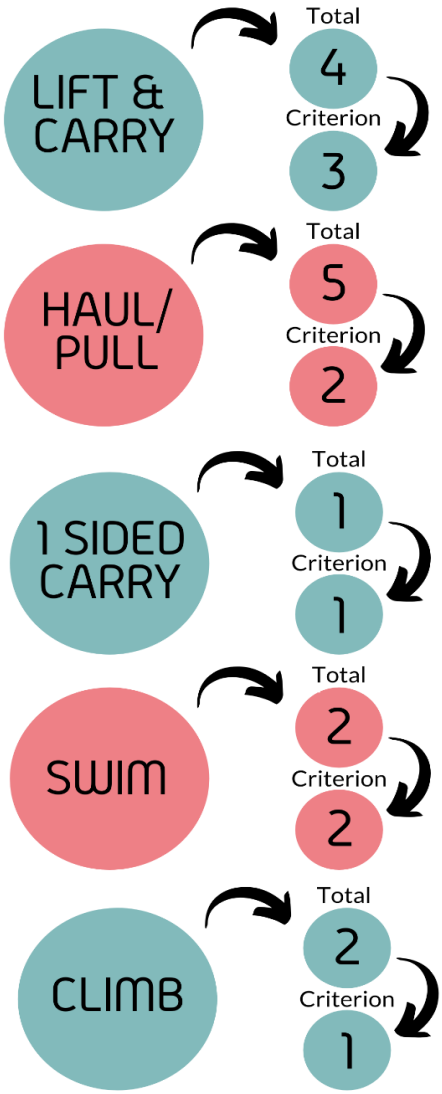
Of the 25 tasks progressing from the Screening phase, 11 were rejected during the Evaluation phase. Broadly, these 11 tasks were found to be low in quantifiable physical demand, were rarely performed, weren't considered mandatory for every volunteer to perform, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "Perform water bailing tasks to remove water from the vessel" and "Lift an outboard motor from rear of vessel during maintenance/cleaning tasks (smaller vessels)".

At the conclusion of the Evaluation phase, 14 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION and ⑤ CATEGORISATION

The 14 tasks progressing from the Evaluation phase reflect the core physical requirements of the Boat Operations skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart below. Considerations were afforded to the environment (Land, On-water, and In-water) in which tasks are conducted and also to the level of physical demand.

Broadly, tasks that were considered as a 'core' physical requirement, but weren't categorised as being 'criterion', were either less physically demanding, represented more of a skill requirement or required the ability to maintain a specific posture rather than a physical demand. This was exemplified by the lift and carry task "Refuel a vessel using jerry cans", which was considered less physically demanding than lift and carry task "Load vessel with food supplies/medical supplies/hay bales/fodder". Similarly, the hauling/pulling task "Help retrieve a colleague from the water", was considered less physically demanding than "Effect a water rescue from a boat using a basket stretcher or backboard".



At the conclusion of the Categorisation phase, a total of 9 criterion tasks spanning four task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

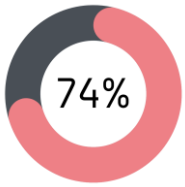
The 9 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Boat Operations volunteer is required to perform to safely fulfil the role. Importantly, the list of criterion tasks includes those which are conducted Land-based, On-water and In-water. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items (including medical supplies or hay bales) to and from the truck and rescue vessel and participate in a stretcher carry. Whilst on the water, the volunteer must be physically capable of effecting a water rescue from a boat using a basket stretcher or backboard and dropping and retrieving the anchor. Lastly, the volunteer must possess a minimum level of physicality to maintain safety in the water using the survival position technique and swim a short distance to make themself safe whilst wearing full personal protective equipment (PPE) and to perform a self-rescue by extricating themself from the water to a point of safety. In this instance, full PPE refers to SES uniform and a personal floatation device (PFD).

2.4 Chainsaw operations

① INQUISITION and ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Chainsaw role. In total, there were 19 tasks ranging from “Use a chainsaw in various postures” to “Clear logs and branches using manual handling techniques”.

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 13 tasks remained to progress into the next phase. Examples of tasks that were excluded include “Use a reach pole to remove debris from high points” and “Use pry bars to manipulate logs”.

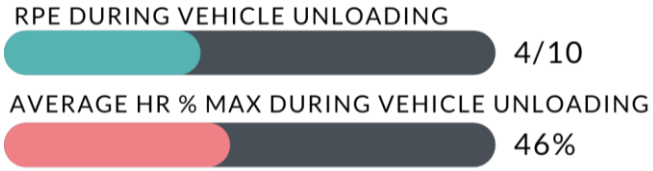


74% of qualified personnel agreed that unloading equipment from the truck is the most frequently performed Chainsaw task

90% of qualified personnel agree that 'clearing logs and branches' is the most physically demanding Chainsaw task.

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here. In a standard Chainsaw simulation activity, volunteers were observed moving equipment from the 4WD to the required location on-site. Both the subjective feedback as well as heart rate data indicate unloading the vehicle is a moderately demanding task.



After unloading the truck of required equipment, volunteers were observed performing static holds with both chainsaws and polesaws. The intensity of these cutting activities was dependent on the type of cutting tool used and the duration of the activity (Figure 6).

Operators were then observed dragging, carrying and lifting the branches they had sawed.

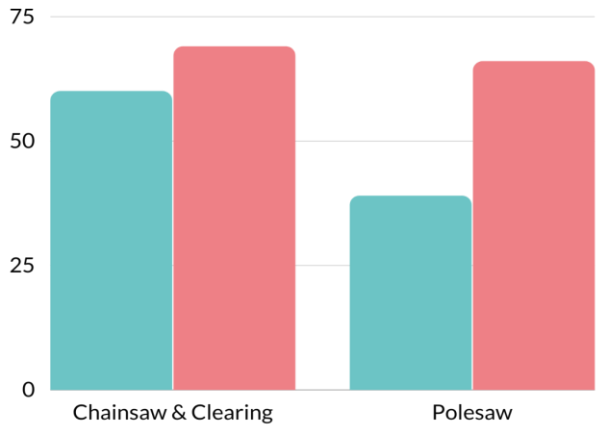


FIgure 4: average heart rate (% max) during chainsaw field trials using medium chainsaws (blue, n=2) and large chainsaws (pink, n=2), for 12 and 6 minutes, respectively. polesaw use was observed once over 6 minutes (blue, n=2) and once over 14 minutes (pink, n=4).

Upon completion of these tasks, operators were observed reloading boxes and equipment onto the 4WD.

Of the 13 tasks progressing from the Screening phase, 5 were rejected during the Evaluation phase. Broadly, these 5 tasks were found to be either low in quantifiable physical demand or were rarely performed and therefore didn't warrant further progression. This included tasks such as "Use a hand saw in various postures" and "Use a bow saw to cut logs/branches".

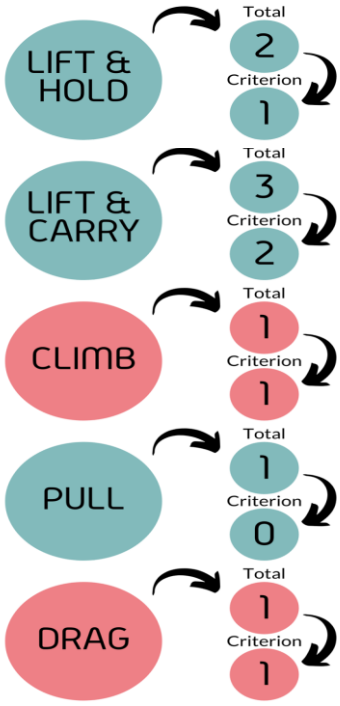
At the conclusion of the Evaluation phase, 8 tasks remained to progress to the Identification and Categorisation phases.

④ IDENTIFICATION and ⑤ CATEGORISATION

The 8 tasks progressing from the Evaluation phase reflect the core physical requirements of the Chainsaw skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart below.

Broadly, tasks that were considered as a 'core' physical requirement, but weren't categorised as being 'criterion', were either less physically demanding, represented more of a skill requirement or required the ability to maintain a specific posture rather than a physical demand. This is exemplified by the lift and hold task "Use a pole saw in various postures", which is considered less physically demanding and is performed less frequently than "Use a chainsaw in various postures".

At the conclusion of the Categorisation phase, a total of 5 criterion tasks spanning 4 task performance feature categories remained.



⑥ CRITERION PHYSICAL TASKS

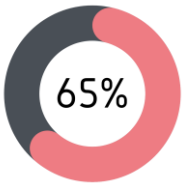
The 5 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Chainsaw volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck, set up and climb a ladder where the job necessitates, operate a chainsaw by performing several static holds of the tool and clear logs and branches from the site where required.

2.5 Firefighting air base support

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Firefighting Air Base Support role. In total there were 4 tasks ranging from "As part of a team set up an airbase" to "Prepare water stations, in assistance for firefighting aircraft".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, all 4 tasks remained to progress into the next phase. This was largely attributable to the tasks receiving ratings of high importance and having a moderate to high physical demand (although no task was reportedly performed very frequently).



65% of qualified personnel agreed that dragging hoses is the most physically demanding Firefighting Air Base Support task.

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. On occasions where field trials could not be conducted, key role-specific information was provided by SMEs. Based on follow-up conversations with SMEs, only 2 of the 4 Firefighting Air Base Support tasks were described to possess some level of physical demand and were also considered essential for every volunteer to complete.

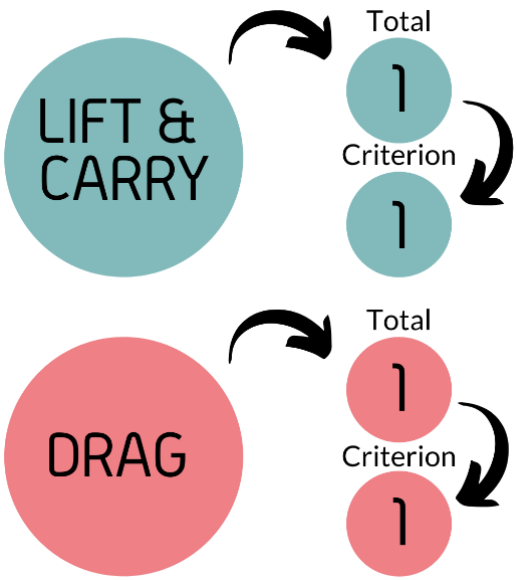
The 2 tasks that were rejected during this phase were done so on the basis of expert guidance. Broadly, these 2 tasks were identified as not being commonly (if at all) performed during typical operations. Additionally, in the case of one of these tasks "Unload boxes and equipment off the aircraft at the end of an operation", the physical demands of performing the task would be subordinate to that of the included task "As part of a team set up an airbase".

At the conclusion of the Evaluation phase, 2 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION AND ⑤ CATEGORISATION

The 2 tasks progressing from the Evaluation phase reflect the core physical requirements of the Firefighting Air Base Support skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart below.

The 2 tasks that remained were identified as being 'criterion', as each possessed different task performance features.



At the conclusion of the Categorisation phase, the 2 criterion tasks were placed into 2 task performance feature categories.

⑥ CRITERION PHYSICAL TASKS

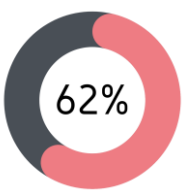
The 2 criterion tasks become the foundation for what a qualified Firefighting Air Base Support volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to set up an air base and drag hoses for aircraft refilling tasks.

2.6 In-water technician

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the In-water Technician role. In total there were 24 tasks ranging from "Help retrieve a colleague from the water (e.g., man overboard)" to "Participate in a stretcher carry of rescued person(s), carrying one side or corner of the stretcher".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, all 24 job tasks remained to progress into the next phase.



62% of qualified personnel agreed that loading equipment onto the truck is the most frequently performed In-water Technician task.

80% of qualified personnel agreed that participating in a stretcher carry is the most physically demanding In-water Technician task

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. On occasions where field trials could not be conducted, key role-specific information was provided by SMEs. Based on conversations with SMEs, 13 tasks were rejected during the Evaluation phase. Broadly, these 13 tasks were found to be low in quantifiable physical demand, were rarely performed, weren't considered mandatory for every volunteer to perform, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "Perform water bailing tasks to remove water from the vessel" and "Perform belaying roles in-water".

11

At the conclusion of the Evaluation phase, 11 tasks remained to progress to the Identification and Categorisation phase.

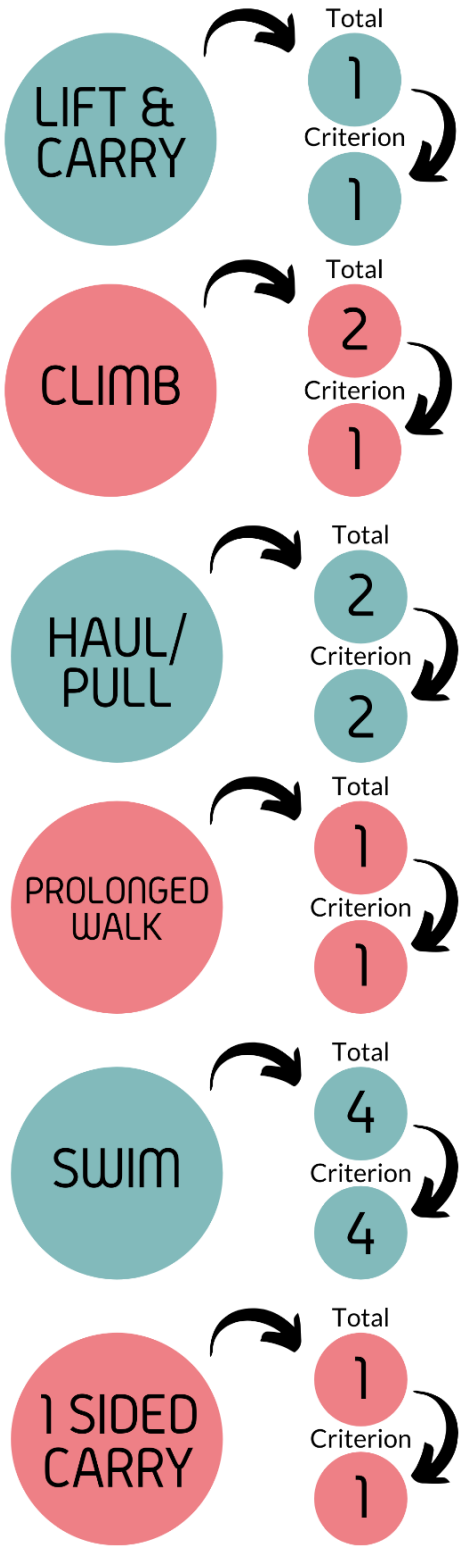
④ IDENTIFICATION AND ⑤ CATEGORISATION

The 11 tasks progressing from the Evaluation phase reflect the core physical requirements of the In-water Technician skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart.

The climbing task "Climb onto capsized vessel & right the vessel in water, wearing full PPE" was considered a 'core' physical requirement of the skill set, however, was not categorised as being a 'criterion' task. Largely, this task was considered less physically demanding than the other climbing task "Perform a self-rescue - extricate yourself from the water to a point of safety".

10

At the conclusion of the Categorisation phase, a total of 10 criterion tasks spanning 6 task performance feature categories remained.



⑥ CRITERION PHYSICAL TASKS

The 10 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified In-water Technician volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry often awkward and heavy items, such as rafts or roping systems, from the truck and hike it to the launch site, haul a rope system and participate in a stretcher carry. The volunteer must also possess a minimum level of physicality to maintain safety in the water using the survival position technique and swim a short distance to make themself safe whilst wearing full PPE, safely perform a swimming rescue of a casualty, swim for short intense bursts whilst wearing full PPE in moving water and to perform a self-rescue by extricating themself from the water to a point of safety. In this instance, full PPE refers to wetsuit, booties, helmet and PFD.

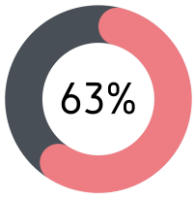
2.7 Land search and rescue

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Land Search & Rescue role. In total, there were 18 tasks ranging from "Conduct contact search activities (e.g., on hands and knees to locate a small object)" to "As part of any search, negotiate obstacles, such as logs, rocks, bush, or shrub".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 13 job tasks remained to progress into the next phase. Examples of tasks that were excluded include "Perform prolonged searches on motor bikes" and "As part of a cave rescue, walk, crawl and push your body through small, dark and confined spaces".

90% of qualified personnel agreed that participating in a stretcher carry is the most physically demanding Land Search task.



63% of qualified personnel agreed that loading equipment onto the truck is the most frequently performed Land Search task

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here. Based on conversations with SMEs, there was 1 task identified that didn't appear in the Inquisition or Screening phases but required further investigation as a potential core physical requirement. After being explored during the Evaluation phase, the task "Unload boxes and equipment off the trailer/truck/4WD to prepare for operations" progressed to the Identification phase as a core physical requirement of the Land Search & Rescue skill set.

During observations of typical Land Search & Rescue training exercises, volunteers were physically active for 1 to 2 hours, with alpine search and rescue tasks lasting in excess of 3 hours (Figure 7). Additionally, during these activities volunteers covered between 1.5 km to 6 km, with alpine search and rescue tasks covering in excess of 15 km. These rescue activities often included a stretcher carry of an injured civilian (Figure 8).

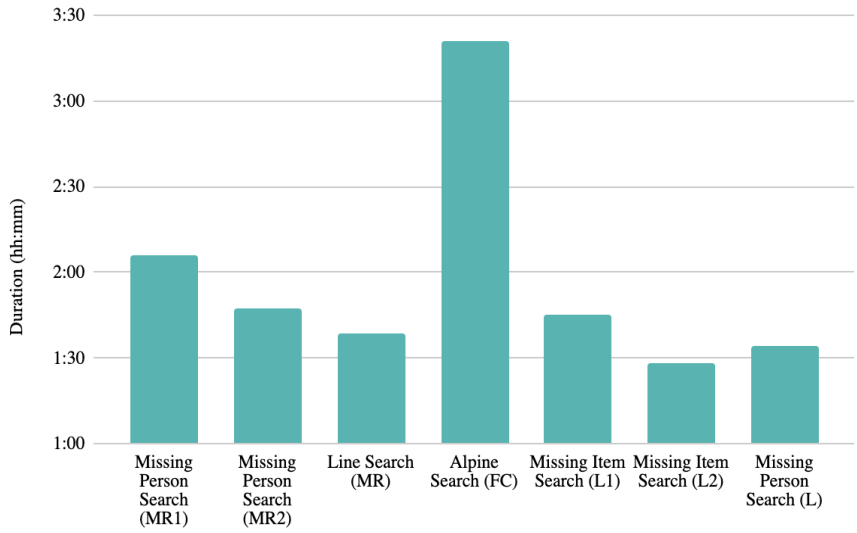


Figure 7: Average duration of physical activity for typical Land Search & Rescue training exercises. FC, Falls Creek; L, Lysterfield; MR, Margaret River.

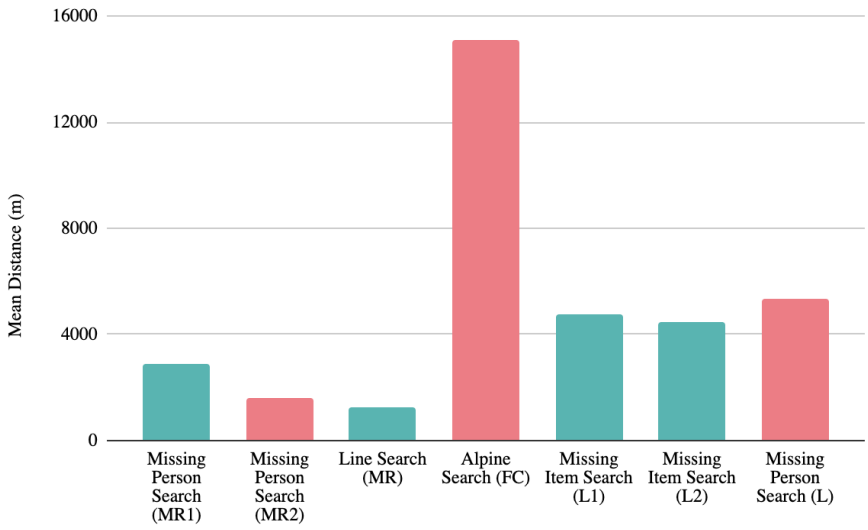


Figure 8: Average distance covered during various Land Search & Rescue training exercises. FC, Falls Creek; L, Lysterfield; MR, Margaret River.

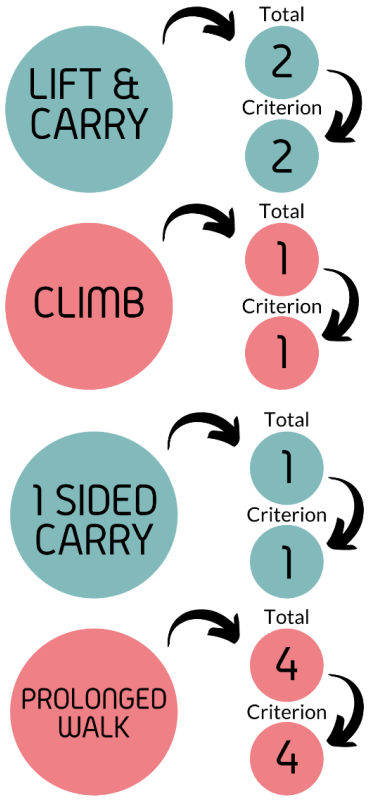
Of the 14 tasks progressing from the Screening phase, 6 were rejected during the Evaluation phase. Broadly, these 6 tasks were found to be low in quantifiable physical demand, were rarely performed, weren't considered mandatory for every volunteer to perform, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "As part of an alpine search, lift and lower packs onto and off snowmobiles" and "Conduct contact search activities (e.g., on hands and knees to locate a small object)".

At the conclusion of the Evaluation phase, 8 tasks remained to progress to the Identification and Categorisation phase.

8

④ IDENTIFICATION AND ⑤ CATEGORISATION

The 8 tasks progressing from the Evaluation phase reflect the core physical requirements of the Land Search & Rescue skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart. All tasks identified as being core to the physical requirements of the skill set are also considered criterion tasks. This means that for these particular tasks, there are no subordinate tasks and any tasks categorised similarly are considered equally physically demanding or contain unique physical characteristics that represent the minimum physical requirements.



8

At the conclusion of the Categorisation phase, a total of 8 criterion tasks spanning 4 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

The 8 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Land Search & Rescue volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck, participate in a stretcher carry and to negotiate obstacles whilst performing a prolonged walk at operational tempo, in conditions/terrain relevant to the level of qualification.

2.8 Land-based swiftwater

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Land-based Swiftwater role. In total there were 11 tasks ranging from "Load vessel with food supplies/medical supplies/hay bales/fodder" to "Perform a throw-bag rescue".

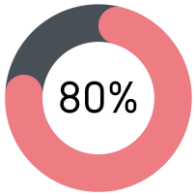
11

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, all 11 job tasks remained to progress into the next phase.

11

62% of qualified personnel agreed that loading and unloading equipment from the truck is the most frequently performed Land-based Swiftwater task.

80% of qualified personnel agreed that participating in a stretcher carry is the most physically demanding Land-based Swiftwater task.



③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here.

During five typical throw-bag rescue scenarios, where volunteers rotated through the role as primary and secondary throw-bag rescuer and upstream and downstream observer, the primary throw-bag rescuer recorded the highest average percentage of maximum age-predicted heart rate (220-age; Shookster et al., 2020) (Figure 9). Each scenario ranged from 1 to 4 minutes in duration.

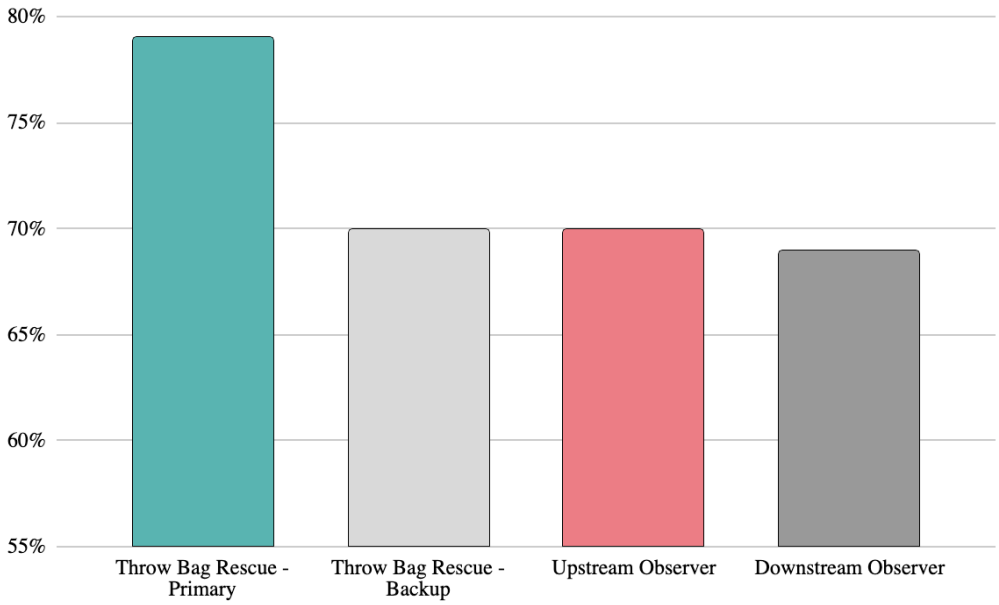


Figure 9: Average percentage of maximum heart rate attained across five different throw-bag rescue scenarios by four Land-based Swiftwater roles (n=4).

When exploring the intensity of specific Land-based Swiftwater scenarios, the rating of perceived exertion (RPE) for rescue scenarios that required the victim to be stretcher carried from the water's edge to an awaiting vehicle was much higher than for scenarios that did not require a stretcher carry (Figure 10).

Figure 10: Rating of perceived exertion (RPE) for scenarios requiring, or not requiring, a stretcher carry.

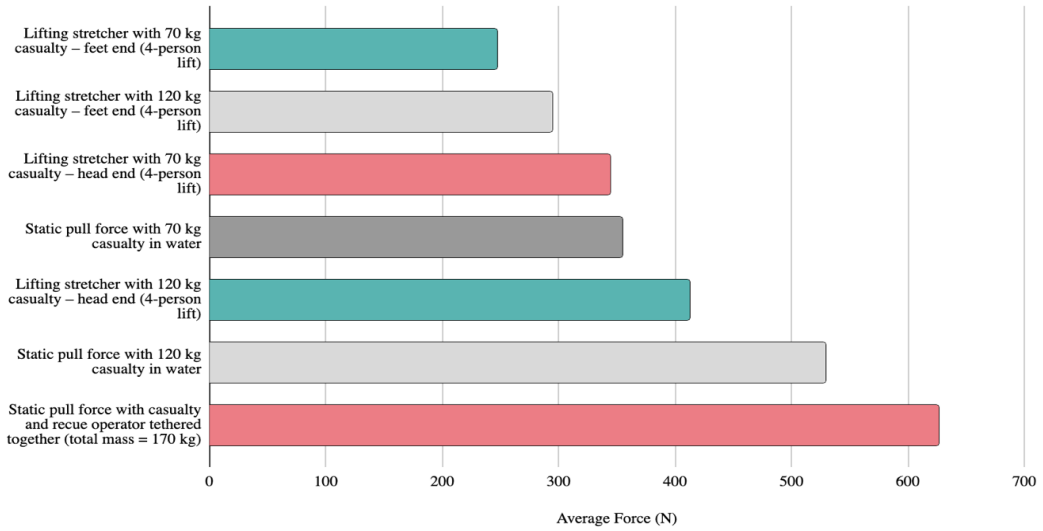


Figure 11: Average amount of force, in Newtons, required to complete various Land-based Swiftwater tasks.

When considering the physicality required to complete various essential Land-based Swiftwater tasks, hauling tasks associated with rescuing casualties from the water typically required a greater level of force production than stretcher carry tasks (Figure 11).

Of the 11 tasks progressing from the Screening phase, 3 were rejected during the Evaluation phase. Broadly, these 3 tasks were found to be low in quantifiable physical demand and were rarely performed and therefore didn't warrant further progression. This included tasks such as "Perform a hauling pole rescue - single person rescue" and "Stabilise a casualty by placing spinal board/stretcher underneath them".

At the conclusion of the Evaluation phase, 8 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION AND ⑤ CATEGORISATION

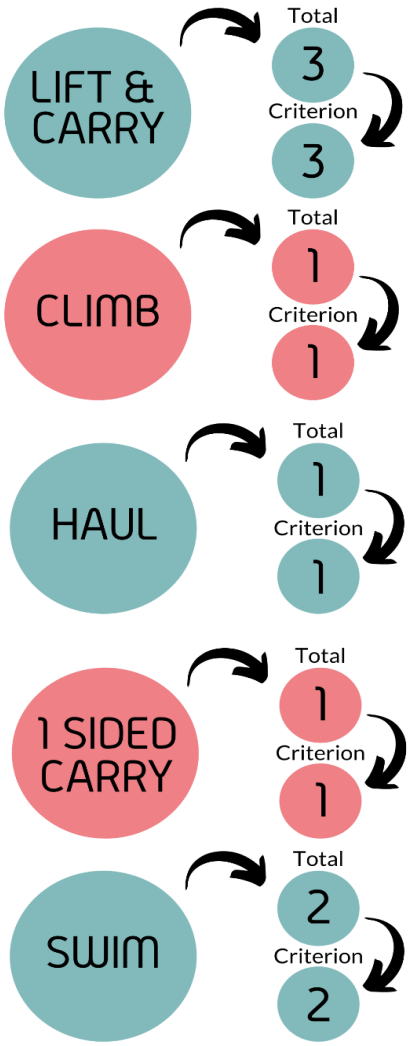
After the Evaluation phase there were 8 tasks identified that reflected the core physical requirements of the skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart below.

All tasks identified as being core to the physical requirements of the skill set are also considered criterion tasks. This means that for these particular tasks, there are no subordinate tasks and any tasks categorised similarly are considered equally physically demanding or contain unique physical characteristics that represent the minimum physical requirements.

At the conclusion of the Categorisation phase, a total of 8 criterion tasks spanning 5 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

The 8 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Land-based Swiftwater volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry equipment from the support vehicle and load the vessel, as well as participating in a stretcher carry and performing a throw-bag rescue. The volunteer must also possess a minimum level of physicality to maintain safety in the water using the survival position technique and swim a short distance to make themself safe whilst wearing full PPE, and to perform a self-rescue by extricating themself from the water to a point of safety.



2.9 Off-road driving

① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Off-road Driving role. In total there were 9 tasks ranging from "As part of a winching activity, pull a winch cable out from vehicle" to "Affix snow chains to a 4WD vehicle".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 3 job tasks remained to progress into the next phase. Examples of tasks that were excluded include "Manoeuvre motorbike (dirt bikes) onto and off trailers" and "Right an upturned quad bike or ATV".

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials. Based on conversations with SMEs, there was 1 task identified that didn't appear in the Inquisition or Screening phases but required further investigation as a potential core physical requirement. This task was "Load the vehicle with all required recovery equipment". After being explored further, this new task satisfied the requirements to progress to the Identification phase as a core physical requirement of the Off-road Driving skill set.

During a typical Off-road Driving training exercise, volunteers were observed covering up to 450 m whilst completing role-specific activities, such as loading and unloading the truck and setting up a tirfor winch (Figure 12).

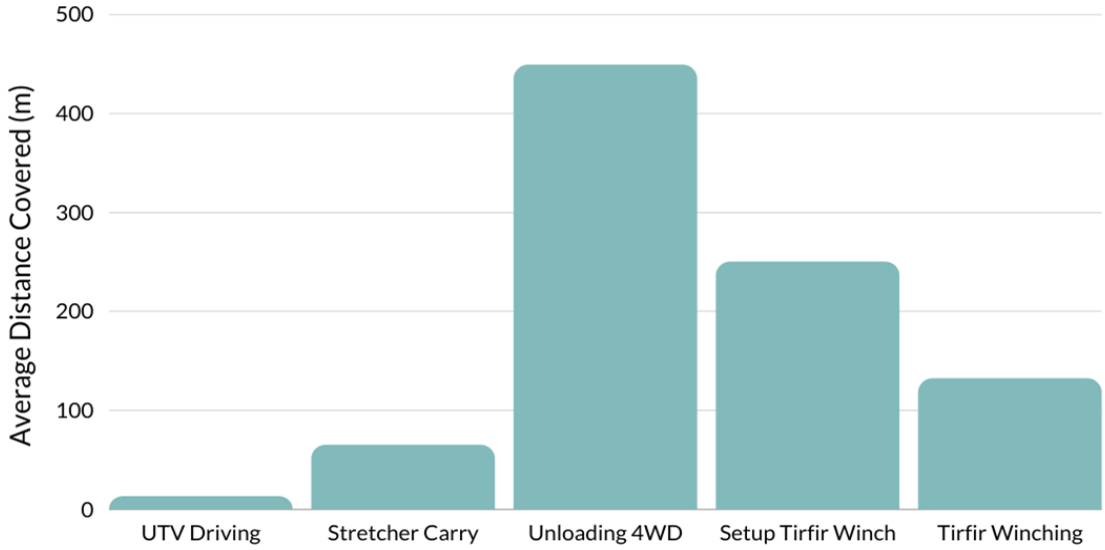


Figure 11: Average amount of force, in Newtons, required to complete various Land-based Swiftwater tasks.

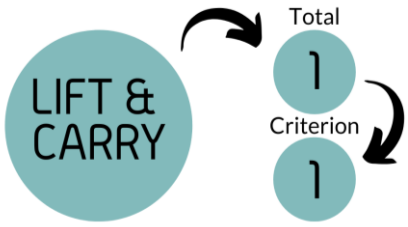
Of the 3 tasks progressing from the Screening phase and the additional task that was identified, 3 were ultimately rejected. Broadly, these 3 rejected tasks were found to be low in physical demand, were rarely performed or weren't considered mandatory for every volunteer to perform and therefore didn't warrant further progression. This included tasks such as "As part of a winching activity, pull a winch cable out from the vehicle" and "Change the tyre of a 4WD vehicle".

At the conclusion of the Evaluation phase, 1 task remained to progress to the Identification and Categorisation phase

④ IDENTIFICATION AND ⑤ CATEGORISATION

The task progressing from the Evaluation phase reflects the core physical requirement of the Off-road Driving skill set.

This task identified as being core to the physical requirements of the skill set is also considered a criterion task. This means that for this lift and carry task, there are no subordinate tasks, and this task represents the most important, physical and frequently occurring task.



At the conclusion of the Categorisation phase, a total of 1 criterion task spanning 1 task performance feature category remained.

⑥ CRITERION PHYSICAL TASKS

The 1 criterion task that progressed from the Categorisation phase became the foundation for what a qualified Off-road Driving volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck whilst on a typical call out.

2.10 Road crash rescue

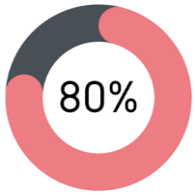
① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Road Crash Rescue (RCR) role. In total, there were 15 tasks ranging from "Perform several static holds with equipment" to "Remove victim from vehicle".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 14 job tasks remained to progress into the next phase. The task that was excluded was "Set up lights around the crash site".

93% of qualified personnel agreed that carrying equipment is the most physically demanding RCR task.

80% of qualified personnel agreed that loading equipment into and out of the truck was the most frequently performed RCR task.



③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here. In a standard RCR simulation activity, volunteers were observed moving equipment from the vehicle to the required location on-site, as well as from location to location throughout the field trial. Most scenarios within the simulation commenced with a stabilization task. Due to the time-sensitive nature of stabilisation, this task was characterised as a mostly moderate intensity task based on rating of perceived exertion (RPE).

During five typical RCR scenarios, including casualty removal from a car on its side (S1), on its roof (S2) and on its wheels (S3), as well as the removal of a casualty requiring a long haul (S4), or a rescue with a suspected spinal injury (S5), volunteers were observed working at or above 70% of their maximum age-predicted heart rate (220-age; Shookster et al., 2020) (Figure 13). S1 and S2 lasted 20 minutes in duration, S4 lasted 9 minutes and S3 and S5 lasted 35 minutes.

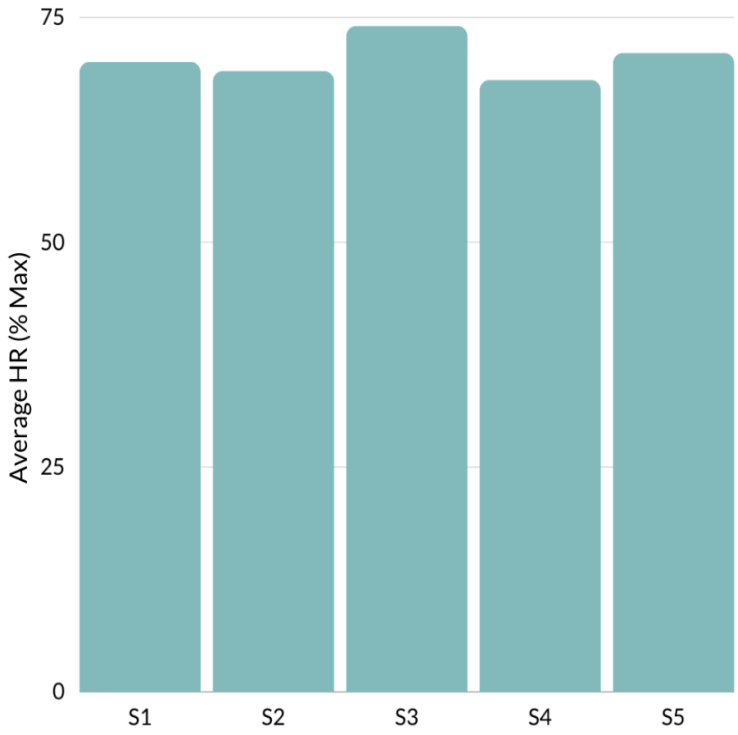


Figure 13: Average percentage of maximum heart rate attained across five different Road Crash Rescue scenarios (n=7). S = scenario.

An important aspect of field observations and trials is the opportunity to gain valuable insight from SME and experienced volunteers. Arising from such opportunities were 2 tasks that required further consideration. The first task was written quite ambiguously and was in need of clarification and the second was a new task not previously captured in the process. These tasks were 1) "Remove victim from vehicle." and 2) "Pull out hydraulic hoses from truck-mounted reels." The clarification gained regarding the first task resulted in the inclusion of the requirement to transport (i.e., carry) the casualty up to 100 m to get further medical assistance.

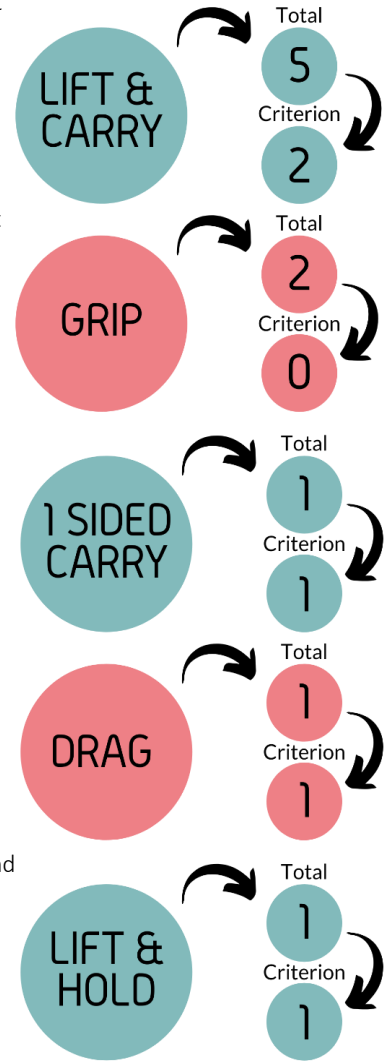
Of the 14 tasks progressing from the Screening phase and the additional task arising from SME guidance, 4 were rejected. Broadly, these 4 tasks were found to be low in quantifiable physical demand, were rarely performed, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "In the event of a road crash down an embankment, or steep drop off, lower equipment down to the site using a hauling system" and "Perform a technical rescue – slow and deliberate rescue".

At the conclusion of the Evaluation phase, 10 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION AND ⑤ CATEGORISATION

The 10 tasks progressing from the Evaluation phase reflect the core physical requirements of the Road Crash Rescue skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart.

Broadly, tasks that were considered as a 'core' physical requirement, but weren't categorised as being 'criterion', were either less physically demanding, represented more of a skill requirement or required the ability to maintain a specific posture rather than a physical demand. This is exemplified by the gripping task "Use hand tools to remove vehicle interior - e.g., dashboard, console, seats etc", which is considered to be low in physical demand. Similarly, the crawling task "Manoeuvre around inside vehicle – tight spaces" is identified as requiring the ability to perform a specialised movement pattern rather than a quantifiable physical demand.



5

At the conclusion of the Categorisation phase, a total of 5 criterion tasks spanning 4 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

The 5 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Road Crash Rescue volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift, carry and drag various items to and from the vehicle and also around the worksite, extract casualties from vehicles and carry them to further medical assistance and hold doors, cutters, spreaders and other heavy objects.

2.11 Storm damage

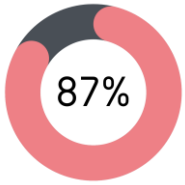
① INQUISITION AND ② SCREENING

35

29

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Storm Damage (comprising both Weatherproofing and Sandbagging) role. In total, there were 35 tasks ranging from "Sweep water using brooms and squeegees" to "Lift and carry sandbags to secure tarpaulins/plastics".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 29 tasks remained to progress into the next phase. Examples of tasks that were excluded include "Pull out plastic sheeting, attached to a roller, from the side of vehicle" and "Establish external lights on stands or establish portable LED lights".



87% of qualified personnel agreed that lifting and carrying sandbags was the most physically demanding Storm Damage task.

75% of qualified personnel agree that loading and unloading equipment from the truck is the most frequently performed Storm Damage task.

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here. In a standard Storm Damage simulation activity, volunteers were observed moving equipment such as tool boxes, traffic management bags, ratchet straps, acro props, cutting tools and sandbags, from the truck to the required location on-site. This equipment weighed anywhere from 11 to 25 kg and it was often necessary to complete multiple carries of a sandbag (Figure 14).

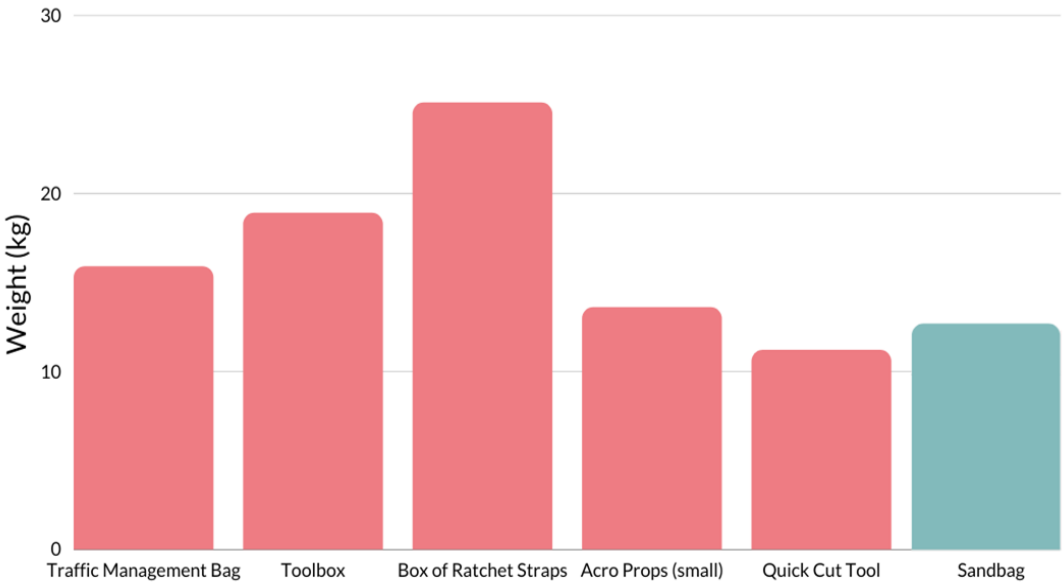


Figure 14: Typical mass of equipment required for a Storm Damage debris-clearing activity. Green = requires multiple lifts and carries.

After unloading the truck of required equipment, volunteers were observed clearing debris either from the ground or from a roof. This required some members to operate tools and haul ropes, whilst other members were completing repeat movements to clear debris. This involved clearing debris from the required site and then returning back to the clearing site to remove subsequent pieces of debris. In these debris-clearing activities, volunteers covered between 480m and 1150m whilst onsite.

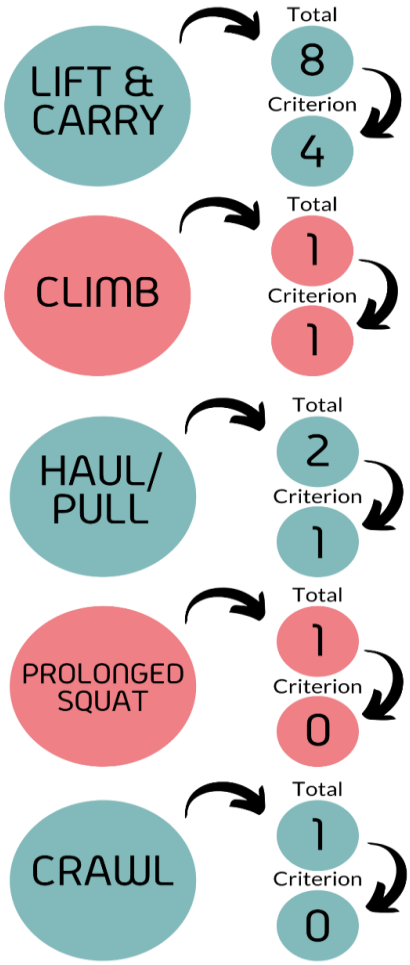
Of the 29 tasks progressing from the Screening phase, 16 were rejected during the Evaluation phase. Broadly, these 16 tasks were found to be low in quantifiable physical demand, were rarely performed, weren't considered mandatory for every volunteer to perform, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "Holding empty sandbags whilst they are being filled" and "Sweeping water using brooms and squeegees".

At the conclusion of the Evaluation phase,13 tasks remained to progress to the Identification and Categorisation phase.

13

④ IDENTIFICATION AND ⑤ CATEGORISATION

The 13 tasks progressing from the Evaluation phase reflect the core physical requirements of the Storm Damage skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart below.



Broadly, tasks that were considered as a 'core' physical requirement, but weren't categorised as being 'criterion', were either less physically demanding, represented more of a skill requirement or required the ability to maintain a specific posture rather than a physical demand. This is exemplified by the crawling task "Move around inside a roof or on top a roof for prolonged periods", which is considered a task requiring the ability to perform a specialised movement pattern rather than a quantifiable physical demand. Similarly, the lift and carry task "Lift and carry ladders to various locations on site", was considered less physically demanding than lift and carry task "Lift and carry sandbags for various distances".

6

At the conclusion of the Categorisation phase, a total of 6 criterion tasks spanning 3 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

The 6 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Storm Damage volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items (including sandbags) to and from the truck and also around the worksite, climb a ladder where the job necessitates, haul equipment (such as tools and tarpaulins) onto the roof and clear debris from the site where required.

2.12 Urban search and rescue

① INQUISITION AND ② SCREENING

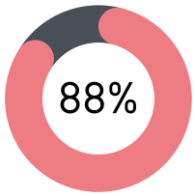
Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Urban Search & Rescue skill set. In total, there were 13 tasks ranging from "Secure star pickets into the ground using either a sledge-hammer or picket dolly" to "Perform repetitive scrambling over rubble and unstable debris".

13

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 12 tasks remained to progress into the next phase. The task excluded was "As part of a CAT 2 rescue, move debris using your hands or equipment in confined spaces".

54% of qualified personnel agreed that loading equipment onto the truck is the most frequently performed USAR task.

88% of qualified personnel agree that participating in a stretcher carry is the most physically demanding USAR task.

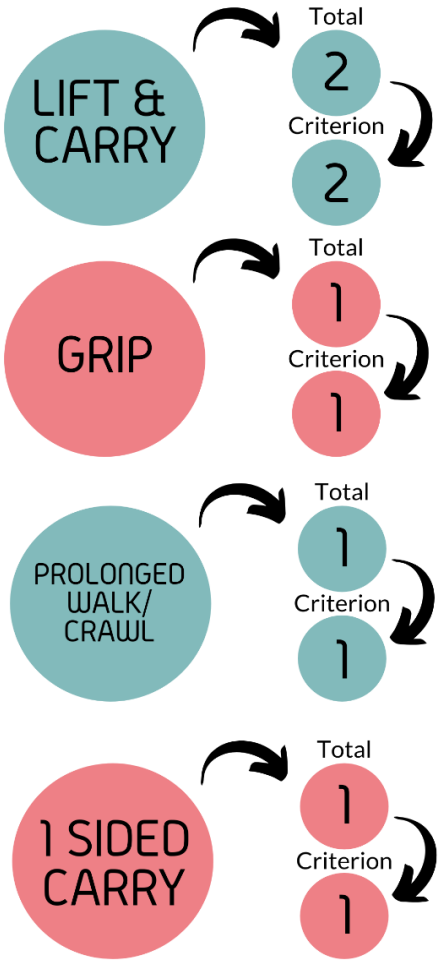


③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of one simulation trial is presented here. When conducting a typical USAR training scenario, where teams crawled over rubble to find an object with a team member becoming injured and requiring stretcher carrying out, volunteers were observed working at above 60% of their maximum HR.

Of the 12 tasks progressing from the Screening phase, 7 were rejected during the Evaluation phase. Broadly, these 7 tasks were found to be low in quantifiable physical demand, were rarely performed, weren't considered mandatory for every volunteer to perform, or were largely skill-based rather than physical and therefore didn't warrant further progression. This included tasks such as "Perform a ground search - looking for people in the affected area. Use a line and hail search technique" and "As part of a CAT 2 rescue, operate tools in a range of positions in confined spaces including laying on back".

At the conclusion of the Evaluation phase, 8 tasks remained to progress to the Identification and Categorisation phase.



④ IDENTIFICATION AND ⑤ CATEGORISATION

The 8 tasks progressing from the Evaluation phase reflect the core physical requirements of the Urban Search & Rescue skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart.

All tasks identified as being core to the physical requirements of the skill set are also considered criterion tasks. This means that for these particular tasks, there are no subordinate tasks and any tasks categorised similarly are considered equally physically demanding or contain unique physical characteristics that represent the minimum physical requirements.

At the conclusion of the Categorisation phase, a total of 5 criterion tasks spanning 4 task performance feature categories remained.

⑥ CRITERION PHYSICAL TASKS

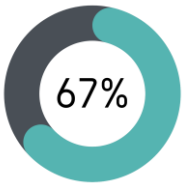
The 5 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Urban Search & Rescue volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck, have the capacity to perform repetitive scrambling over rubble, as well as the strength required to pass and receive a victim in a stretcher over rubble and participate in carrying the stretcher across stable ground.

2.13 Vertical rescue

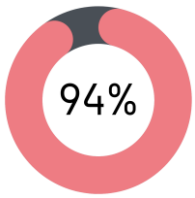
① INQUISITION AND ② SCREENING

Through the interrogation of SES documentation and from workshops held with qualified and experienced SMEs in each state and territory, a comprehensive task list was generated encompassing the Vertical Rescue (VR) role. In total there were 14 tasks ranging from "Abseiling to a victim" to "Tying and untying knots".

During the Screening phase, the identified tasks were then entered into an online survey, where qualified personnel rated each task on performance frequency, physical demand and task importance. Once the responses were analysed, 10 tasks remained to progress into the next phase. Tasks that were excluded include "Perform an edge manager role" and "Climb a 25 m ladder during an industrial rescue".



67% of qualified personnel agreed that loading and unloading equipment from the truck is the most frequently performed VR task.



94% of qualified personnel agree that participating in a stretcher carry is the most physically demanding VR task.

③ EVALUATION

A critical step in the Evaluation phase is the observation, monitoring and measuring of task performance demands during field trials. Both physiological and observational data are collected throughout field trials, of which a small example of multiple different simulation trials is presented here.

During a typical VR simulation activity, volunteers were first observed unloading boxes and equipment off the trailer/truck/4WD to prepare for operations. Subsequently, volunteers then carried equipment over long distances, where the average heart rate was above 70% of maximum capacity, indicating a high physical intensity (Figure 15 and Figure 16). The walking activities ranged from 5 to 9 minutes in duration.

A variety of "on-rope" tasks were observed, including the first and second responder roles, where volunteers were typically working between 62% and 74% of maximum age-predicted heart rate capacity (220-age; Shookster et al., 2020). The initial survey data indicated that 'hauling' is a physically demanding task, which was affirmed by field trial

data that indicated the average heart rate while hauling was around 70% of maximum capacity. Finally, operators reload boxes and equipment onto the trailer/truck/4WD at the end of an operation.

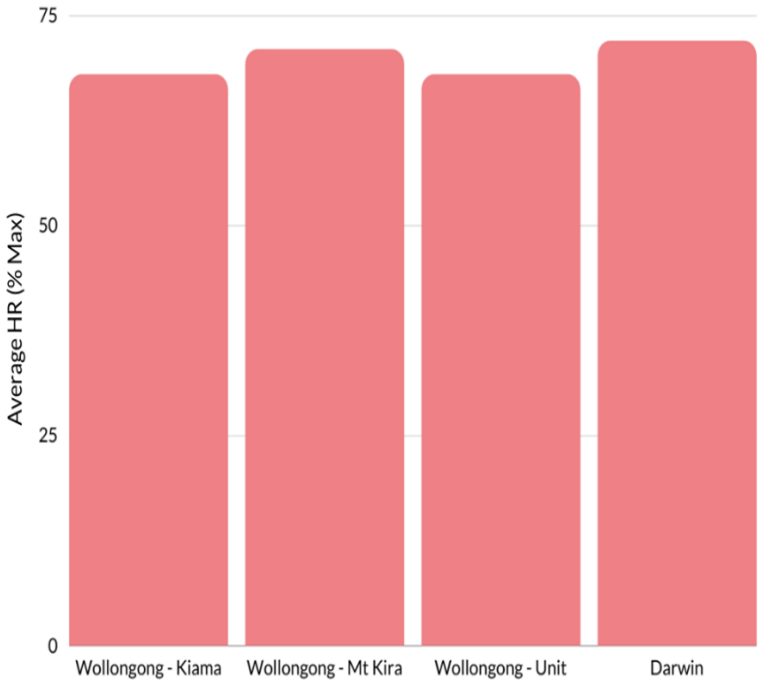


Figure 15: Percentage of maximum heart rate attained whilst walking with equipment during typical Vertical Rescue simulation activities at Kiama (n=7), Mt. Kira (n=6), Wollongong Unit (n=6) and Darwin (n=3).

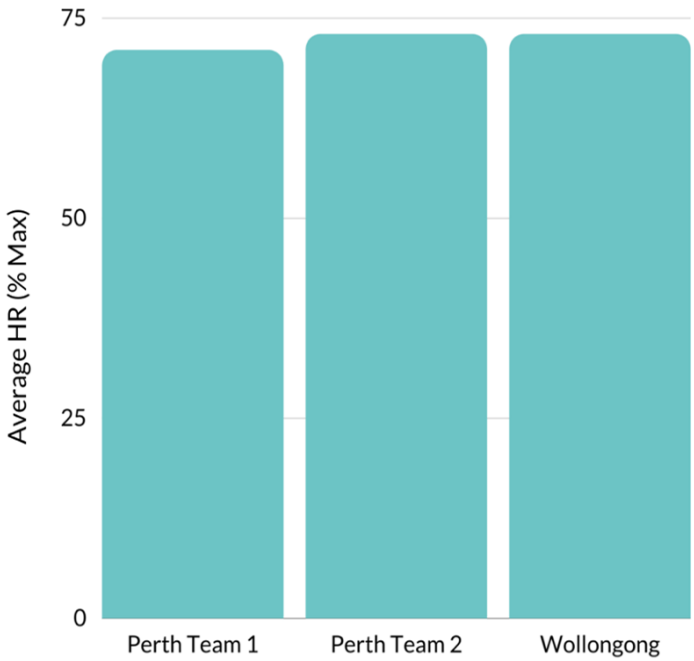


Figure 16: Percentage of maximum heart rate attained whilst walking with equipment during typical Vertical Rescue simulation activities for three different teams (Perth 1, n=9; Perth 2, n=10; Wollongong, n=3).

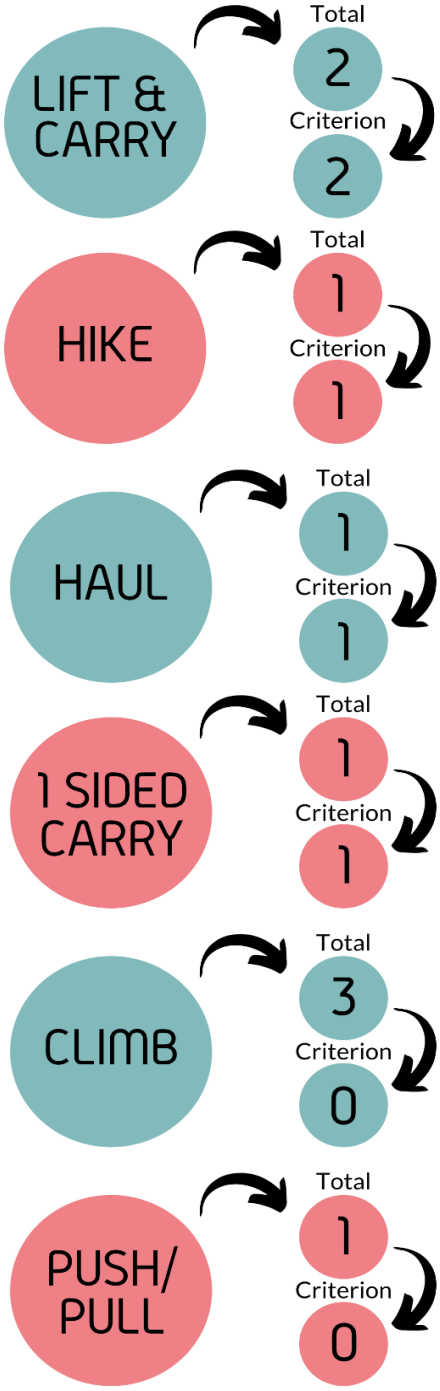
Of the 10 tasks progressing from the Screening phase, 1 was rejected during the Evaluation phase. This task was "Perform a skull drag manoeuvre - where the victim is hauled with no mechanical advantage used" and was rejected due to being rarely performed. At the conclusion of the Evaluation phase, 9 tasks remained to progress to the Identification and Categorisation phase.

④ IDENTIFICATION AND ⑤ CATEGORISATION

The 9 tasks progressing from the Evaluation phase reflect the core physical requirements of the Vertical Rescue skill set. These tasks have been categorised based on their task performance features and are presented in the flowchart.

The tasks that were considered as a 'core' physical requirement, but weren't categorised as being 'criterion', were the on-rope tasks. Whilst there is a large component of skill involved in these on-rope tasks, there are also physical demands associated with their performance. The spread of physical demands covered by the 5 criterion tasks provides confidence that those who meet the standards for the criterion tasks will have the requisite physicality to safely perform these core on-rope tasks. It is highly recommended that a skill-based on-rope assessment is included for the Vertical Rescue skill set.

At the conclusion of the Categorisation phase, a total of 5 criterion tasks, spanning 4 task performance feature categories, remained.



⑥ CRITERION PHYSICAL TASKS

The 5 criterion tasks that progressed from the Categorisation phase become the foundation for what a qualified Vertical Rescue volunteer is required to perform to safely fulfil the role. In this instance, the 'funnel and filter' process has identified that each and every qualified volunteer needs to possess the physical capability to lift and carry various items to and from the truck, hike over long distances carrying equipment, possess the unilateral strength to complete a stretcher carry and haul ropes to hoist victims up a cliff.

3 Developing assessments

This section of the report aims to provide a transparent and detailed understanding of the process undertaken to craft the raw ingredients (i.e., the criterion tasks) arising from all relevant skill sets, into a product that is best placed to address and assess the physical underpinnings of safely performing each of the job roles. To achieve this, the approach comprised three phases: 1) conceptualisation, 2) creation and 3) evaluation. This multi-phase process is presented in Figure 17, with each phase explained in greater detail below.

The first phase is necessary for conceptualising the approach to the problem and needs to include identifying and understanding the important considerations that inform the creative boundaries and guide the development of the design framework. The second phase focuses on the creation of a product which not only possesses the requisite scientific rigour, but also correlates with organisational desires and real-world implementation capability. Importantly, this phase includes opportunities for structured pilot testing to occur. The third phase is responsible for ensuring that the previous two phases are held accountable and have met their aspirations and obligations, resulting in the development of a product that meets both scientific and organisational requirements. Importantly, the third phase ensures the product is ready to progress to the next stage in the research design (i.e., reliability and validity).

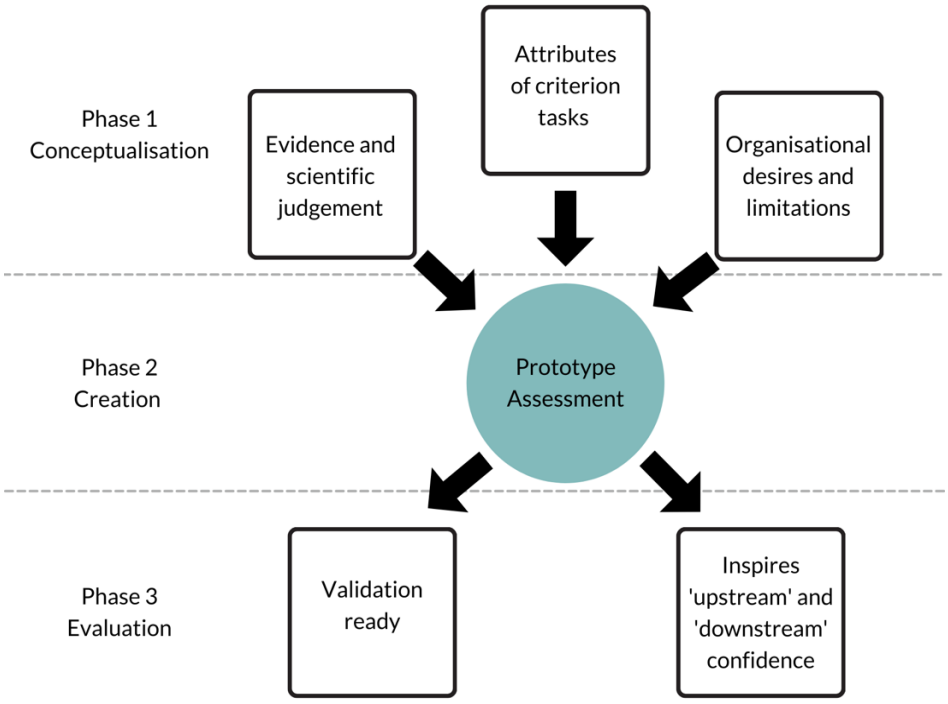


Figure 17: The three-phase creative process employed for the design and development of SES Fit for Task prototype assessments.

PHASE 1: CONCEPTUALISATION

The conceptualisation comprises the INPUTS (or considerations) which aid in setting creative boundaries:

1. Evidence and scientific judgement: what was seen, heard, measured, analysed, what's current best practice, and the research team's expert experience and opinion.
2. Attributes of criterion tasks: the metrics that result in successful task performance.
3. Organisational desires and limitations: direction and guidance received from end-users about how they envisage implementing individual assessments.

PHASE 2: CREATION
These are the elements of programme design, product development and testing and refinement:

1. Programme design: this is the framework that the prototype assessments nest within. The programme guides how all the pieces fit together and how it is to be used within the organisation.
2. Prototype assessment development: this should aim to include as many assessment performance (e.g., time to complete an assessment) and design characteristics (e.g., using appropriate equipment) as possible.
3. Pilot testing: there should be opportunities afforded to the programme and the prototype assessments undergoing a real-world 'test and adjust' procedure. This includes the prototype assessments being exposed to the scrutiny of SMEs and also conducting a pilot trial to gauge feedback from end-users. Programme and prototype refinement and modification to be made, if necessary and justified.

PHASE 3: EVALUATION
These are the OUTPUTS (or KPIs) of the process and form the basis for accountability:

1. Validation ready: the prototype is (at least) the 80% solution (in reference to design and performance characteristics) and capable of progressing to the next research stage, validation trials.
2. Inspires 'upstream' and 'downstream' confidence: the prototype must be able to inspire confidence that it captures the requisite physicality of what it has been designed to represent (i.e., the 'upstream' criterion tasks). In addition to this, it must be able to meet the expressed desires of end-users and be able to be implemented and supported ('downstream' confidence).

3.1 Conceptualisation phase

As should be expected in a highly applied research project such as this, reliance purely on the raw task performance data will not lend itself to achieving an optimal or desired outcome. Whilst this type of evidence is important, conversations had with SMEs and experienced personnel (especially during Workshops and Field Trials) as well as specific organisational knowledge also comprise key components to shaping outcomes and making (and validating) decisions. Without the careful synthesis of the objective evidence, SME input and organisational desires, there is a real risk of developing a highly accurate product that is impossible to implement, or conversely a product that is simple to implement but provides no insight into the physical capability of the population of interest.

Task performance data collected during each stage of the project contributes to the overall job (and task) performance knowledge bank. SMEs and experienced personnel provide important commentary and context to findings and experiences from the previous research phases. Organisation-specific knowledge is a crucial consideration here as the desires for the end-product (i.e., the physical assessment programme) and the reality of implementation capability need to be clearly articulated, understood, and acknowledged. The research team at the centre of it all are charged with pulling all those components together and shaping them into an assessment programme which meets not only the requisite scientific rigour but also organisational expectations (including any expressed limitations).

During field trial data collection, it was common for the research team to be presented with representative scenarios. These skill set-specific scenarios were purpose-made to capture all the physical tasks considered important to job function and allowed researchers to appreciate the performance chronology and interconnectedness of tasks. With a dozen skill sets to develop physical assessments for, and with skill sets having a wide variety of response contexts and demonstrating a different array of criterion tasks, it would have been feasible to use these scenarios as the basis to design a unique and specific assessment battery for each skill set. This appears intuitive, as the criterion tasks that arise from each skill set are contextually unique and representative of only that skill set.

Therefore, it makes sense that each skill set would require its own unique set of assessments. However, while this may seem like an intuitive, sensible and appealing approach at first, it quickly runs into trouble when you consider implementing this at an organisational level. Take a moment to imagine the person in the organisation who holds qualifications in 10 skill sets. It would be time consuming and grossly inefficient to have them perform 10 completely unique and skill set-specific assessment batteries, and from the evidence collated, there would be a large amount of unnecessary duplication (due to the significant amount of fundamental movement and task performance overlap present amongst skill sets).

With bespoke skill set assessment batteries being untenable and also appearing unnecessary, there was a need to provide a solution to the problem of constructing a group of assessments that had the capability to be targeted enough to address the physical demands of individual skill sets but also flexible enough to allow for the variation in performance intensity observed between skill sets. A first step in finding a solution to this problem is understanding and appreciation of the similarities and differences observed across the range of skill sets.

When accounting for all the similarities and differences amongst skills sets, it was determined the most appropriate assessment development framework was one with a focus on (1) core job physicality (i.e. the fundamental physicality required to be demonstrated in order to meet job demands), (2) performance features of criterion tasks (e.g. is the task lifting and carrying, is it pulling or dragging, or is it prolonged walking) and (3) organisational desires and limitations. These three key focus areas were deemed the prototype assessment development 'inputs' and would help guide the initial design of the prototypes. Additionally, once the initial prototypes were designed and pilot tested, they would be evaluated based on two key performance indicators, referred to as the 'outputs'.

Pleasingly, at the completion of the conceptualisation phase the problem had been identified and a clear approach to guide the development of a solution had been developed. Additionally, the research team had detailed how they would hold the process accountable to ensure the outcome was going to satisfy scientific rigour and organisational desires. The next portion of this chapter is dedicated to these creative processes and to provide details of the prototype assessments that were developed.

3.2 Creation phase

Knowing the desirable outcomes of successfully combining all the INPUTS, the research team set about developing a prototype assessment programme. The role of the programme would be to provide the skeleton that would house the assessments in such a way that enabled job-related and specific fundamental movements to be assessed at an intensity representative of skill set task performance. The programme would be designed such that it removes the between-skill set duplication of like-tasks and movements that had been observed during earlier research phases. Each assessment would address a requisite physical demand as presented through criterion tasks. Once prototypes of the programme and assessments are developed, they are subject to pilot testing on a number of fronts including dissemination to members of the project team for socialising and comment, and formal field trials of assessments with SES volunteers to ascertain real-world capability and conduct.

This phase is presented in three parts: 1) programme design, 2) prototype assessment development and 3) pilot testing.

3.2.1 Programme design

The first step was to design a programme that was able to succinctly administer physical assessments to a wide range of volunteers, each with a varying portfolio of skill set qualifications. Through the previous phase the research team had established there was a degree of overlap in the physical nature of the work performed between skill sets. Whilst there were some clear differences and unique aspects observed for some skill sets there were also many common features. It was decided that for the programme to deliver the outcomes needed, it would be largely based on movement patterns and incorporate levels of performance to account for skill sets requiring higher intensity work. Building a programme structure such as this would allow for a 'plug and play' type of approach to enable each individual volunteer to be prescribed a battery of assessments as dictated by the skill sets they held qualifications in. A big advantage of a programme such as this is that even though a movement pattern (e.g., loading and moving equipment) could be common across multiple skill sets, it would only need to be assessed once.

3.2.2 Prototype assessment development

During the job analysis process, each skill set was subject to a series of research steps that ultimately resulted in criterion tasks being identified. These criterion tasks summarise the inherent physical nature of the skill set and are representative of the physicality required to safely perform the job role. The fundamental movement and task performance attributes of criterion tasks were identified and presented at Stage 5 (Categorisation) of the job analysis process.

As such, what follows in this section is a description of the creative process used by the research team to move from 67 individual criterion tasks, which were grouped into 7 task performance categories, to a set of prototype assessments capable of delivering intended organisational and project outcomes.

The 7 task performance feature categories:

1. LIFT & CARRY; 25 CRITERION TASKS; 1 ENVIRONMENT; 12 SKILL SETS
2. SINGLE-SIDED LIFT & CARRY; 7 CRITERION TASKS; 1 ENVIRONMENT; 7 SKILL SETS
3. PROLONGED WALK; 8 CRITERION TASKS; 1 ENVIRONMENT; 4 SKILL SETS
4. PULL / HAUL / GRIP / DRAG; 11 CRITERION TASKS; 2 ENVIRONMENTS; 9 SKILL SETS
5. CLIMB / CRAWL; 6 CRITERION TASKS; 2 ENVIRONMENTS; 6 SKILL SETS
6. LIFT & HOLD; 2 CRITERION TASKS; 1 ENVIRONMENT; 2 SKILL SETS
7. SWIM; 8 CRITERION TASKS; 1 ENVIRONMENT; 3 SKILL SETS

The desires and limitations expressed by the organisation were taken into account when designing the programme and also during the formulation of each prototype assessment. For SES these desires and limitations included the following:

- ➔ The assessments should 'look' and 'feel' like the job;
- ➔ Easy to 'sell' to volunteers;
- ➔ No major equipment purchases required;
- ➔ No major infrastructure projects necessary;
- ➔ Administered by anyone within SES (doesn't require specialist health and fitness qualification or training); and
- ➔ Testing should be time efficient and able to handle large groups.

As these desires and limitations were a common input element to all prototypes, they will not be addressed in replication in this section but will be referred to in the evaluation phase of this chapter, and specifically when ascertaining downstream confidence.

On the pages that follow, you will be introduced to the nine prototype assessments that have been developed from the seven task performance feature categories. For each category the 'inputs' (excluding organisational desires and limitations) will be detailed. After the 'inputs' the prototype assessments that align with the category will be presented in table format. These tables include information regarding the assessment name, the performance environment, equipment required, the intensity of effort required, and intra-assessment performance standards.

CATEGORY 1: LIFT & CARRY

Inputs: LIFT & CARRY is the largest category when it comes to sheer quantity of criterion tasks, with all tasks being characterised by picking an item up, carrying it for a distance, then placing it down. The between-task similarities and differences observed relate to the items, the method of carrying, and the location of placing. Through field observations it became clear that there were certain skill sets requiring heavier items to be moved around and also those which required more items to be moved. Carrying distances, terrain and method and placing location was found to be similar amongst the skill sets. A carry of 30m was recorded as being representative of the typical distance (rescue truck to worksite) with the most common lift height being the lower shelf of the rescue truck (1.2 m). To be successful in these tasks requires picking up a number of representative items, carrying them (over representative ground using a variety of common methods), then placing them down in a number of different locations including on the ground and also into the rescue truck.

Prototype Assessment: LIFT & SHIFT		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound.
Performance elements	Variety of common items and carry methods utilised as well as different pick up and put down locations (i.e., ground and platform). Ability to incorporate a "heavy" element into the assessment design and flow.	Representative of the common carry distances, the common items being carried, a mix of required carry methods including one and two handed, by the side, and in front of the body. "Heavy" element provided for by a number of repeated sandbag carries. Safety needs to include flat level ground to avoid trip hazard.
Equipment	Everyday equipment that would be found on the shelf, in the yard or on a truck. Items to have easy-to-identify handles or carry points, a variety of weights, dimensions, and carry methods required. Platform height at approx. 1.2m.	Items such as containers, chainsaw, sandbags, fuel, tool kit / bag (specifics not set in stone; item guidelines developed). Platform height represents that of a common rescue truck shelf, but anything can be used as long as it meets the guidelines and is stable and safe.
Intensity of effort	Intensity comes from item weights, carry distance, pick up and put down locations, performance speed, and carry methods. Self-paced; walking only.	Representative of on-the-job task performance. Platform to ground, ground to ground, and ground to platform combinations included. No urgency or life-critical element present in criterion tasks.
Assessment standards	30 m carry distances using a variety of carry methods and pick up and put down locations. Repeat sandbag carries incorporated into the assessment flow rather than simply a tack-on. Two performance levels developed with Level 2 applicable to those skill sets requiring handling of heavier items and equipment. No time limit imposed for this assessment.	All skill sets (with the exception of Air Search) have a LIFT & SHIFT assessment aligned with them. Arising from the analysis conducted with skill sets the following would have the "heavy" option aligned with them: Boat Operations, In-water Technician, Land-based Swiftwater, Land Search (Level 3), Road Crash Rescue, Storm Damage, and Vertical Rescue.

CATEGORY 2: SINGLE-SIDED LIFT & CARRY

Inputs: this was the category with the most like-tasks which related to being able to contribute to a stretcher carry (on one side or corner). Between-task differences mostly surrounded the distance or duration the stretcher would be required to be carried. One task from Category 4 (PULL / HAUL / GRIP / DRAG) was imported into this category for consideration due to it being closely related to the tasks found in this category (the USAR 'pass and receive a stretcher' task; classified as GRIP). Being able to successfully execute all tasks in this category requires a unilateral capability to lift a heavy object from the ground, maintain a grip on it, walk a distance, before safely placing it back down again. Additionally, the ability to repeat the effort gives rise to the confidence to contribute to the team task for a prolonged

period of time. A minimum expectation regarding single carry distance was conveyed to researchers to be around 50m, as less than this would result in the task being unacceptably stop / start in nature and thus unable to achieve its purpose.

Prototype Assessment: CARRY		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound.
Performance elements	Single-sided lift and carry of a surrogate designed to represent a corner of a laden stretcher (4-person carry). Both left and right sides of the body are assessed.	Left side pickup, carry, and put down assessed as well as right side pickup, carry, and put down.
Equipment	Stretcher surrogate that has similar attributes to a stretcher including weight, handle height, handle diameter and carry behaviour. A 24 kg kettlebell appears fit for purpose (to represent one quarter of total weight; 80 kg casualty + 15-20 kg stretcher).	The equipment for this assessment is especially important. How the surrogate feels and behaves are vital factors when designing an assessment to replicate what has been universally reported to be one of the most physically demanding tasks performed.
Intensity of effort	Intensity comes from item weight, carry distance and walking pace. Self-paced; walking only.	Picked up from, and returned to the ground each carry. Criterion tasks are team-based; assessment is individual (team will work to the pace of their slowest member).
Assessment standards	What is being lifted and carried doesn't change, but the distance or duration of the task does. Two performance levels have been developed: Level 1: 2 x 50 m efforts. Level 2: 4 x 50 m efforts. No time limits are set for this assessment.	Arising from the analysis conducted with skill sets the following would have the "longer or further" option aligned with them: In-water Technician, Land Search (Levels 2 and 3), USAR and Vertical Rescue. The lesser performance level is aligned with General Rescue, Land Search (Level 1), Boat Operations, Land-based Swiftwater and Road Crash Rescue.

CATEGORY 3: PROLONGED WALK

Inputs: the common theme presented by the tasks in this category was constant and consistent movement over ground. Between-task differences are observed when it comes to what is carried, the time or distance of movement, and the urgency of the mission. Through workshops and field observations it became clear which skill sets, and skill set subgroups, required carrying heavier external load, moving over a greater distance (and / or time), and this movement having mission-critical timing. Successful performance of these tasks requires consistent movement over ground and being able to negotiate representative obstacles at operational pace while carrying the requisite external load.

Prototype Assessment: HIKE		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Ready access to a suitable location with considerations including size of venue and ability to construct obstacles.
Performance elements	Constant movement over ground and negotiating representative obstacles while carrying the appropriate external load.	Walking at operational pace for a set distance. Requirements of the individual skill sets (and qualifications within when it comes to Land Search) dictate performance attributes.
Equipment	Repeated 400 m "lap" performed on firm level ground. Individual backpacks loaded to skill set-appropriate weight. Item capable of being climbed onto, over, and then safely off. 6' pickets and SES tape to construct step over (approx. 50 cm) and crawl under (approx. 60 cm). and cribbing to construct step ups (30 cm and 50 cm).	400m is considered ideal with guidance developed for shorter and longer. Importantly, independent of the assessment level, the weight of backpacks reflects what is carried on the job (and this will differ between skill sets). Necessary stores and equipment readily found at SES units.
Intensity of effort	It is a walking only assessment. Intensity comes from walking distance and pace, backpack weight, and completion of obstacles.	Evidence-based backpack weights are important as this impacts the overall effort required to perform the assessment. Obstacles are completed on all laps apart from the first and the last.
Assessment standards	Three performance levels developed. Performed on the same course with similar obstacles utilised. Level 1 (2 km) is self-paced and thus no time limit is imposed. Levels 2 (4 km) and 3 (6 km) have an overall time limit imposed (cut-offs to be determined).	Lower step-up height prescribed for performance Level 1; all other obstacles remain the same. Level 1 is simply about demonstrating continuous forward motion, whereas for Levels 2 and 3 factors such as predicting a greater duration or distance of travel or the performance of further work upon completion are important considerations.

CATEGORY 4: PULL / HAUL / GRIP / DRAG

Inputs: this category presented a diverse range of tasks performed in different environments (on-land and on-water); however, it was observed that they all required a degree of upper body strength and / power. The one GRIP task which was included (the USAR 'pass and receive a stretcher' task) has already been addressed in CATEGORY 2: SINGLE-SIDED LIFT & CARRY. It then became important to understand and define what was required in the remaining tasks, and due to performance similarities, this resulted in the formation of three groups. These groups were: 1) tasks which required the person to be stationary and move the 'something' from a distance away to close to them using a rope, 2) tasks which required the person to lift and support a portion of the 'something' while they move it to another location (with the remainder of the 'something' in contact with the ground), and 3) tasks which required the person (who was in the vessel) to pull another person from the water into the vessel (this last group was decided to be best covered during water-based testing and has been incorporated into the IN-WATER SAFETY assessment presented in CATEGORY 7: SWIM).

The between-task differences needing consideration included the 'somethings' (e.g., weight and quantity), the distance required to be moved, and how many people were contributing to performing the task. Successful completion of tasks

comes down to being able to impart the requisite amount of force or effort, in order to move the 'something' or person a representative distance, for a representative number of repetitions.

Prototype Assessment: HAUL		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound
Performance elements	Hauling movement (person is stationary and pulls the object toward them) over a representative distance, using a hand-over-hand technique with multiple haul efforts separated by a walk recovery.	Importantly, the haul distance and the effort required to complete the assessment must align with how it feels to execute the criterion tasks it represents.
Equipment	Something to grip on to and haul with, something to haul, a suitably robust surface, and the ability to add extra resistance.	25 m rope attached to a 13" spare wheel (rim + tyre). Level concrete (or non-marking) surface Additional objects weighing 10 kg to add weight where required.
Intensity of effort	Intensity in this assessment comes from the relationship between the surface and the object being hauled (i.e., the resistance being experienced and thus needing to be overcome) and the haul distance. A heavier object and repeated hauls increase the effort required.	It is important that the level of effort required during the performance of the assessments is representative of the tasks it is designed to replicate. In particular this extends to the resistance experienced and the number of repetitions performed. The walk recovery between repeated haul efforts replicates the natural on-the-job breaks experienced (e.g., to reset the equipment before hauling again).
Assessment standards	Two performance levels developed with both using the same rope, spare wheel, and haul distance. Level 1 is a single 20 m repetition performed with the spare wheel. Levels 2 is performed with the spare wheel + 10 kg and 2 x 20 m repetitions required to be completed.	A stable and stationary haul position needs to be demonstrated throughout the assessment. There is no time limit enforced for this assessment, however constant hauling movement during repetitions is required to be demonstrated. Level 1 is aligned with Boat Operations, Land-based Swiftwater and Storm Damage. Level 2 aligns with In-water Technician and Vertical Rescue.

Prototype Assessment: DRAG		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound.
Performance elements	Dragging movement (person moves and takes the object with them) over a representative distance, holding part of the object with the remainder of it trailing behind on the ground with multiple drag efforts separated by a walk recovery.	Importantly, the drag distance, the number of repetitions required, and the effort required to complete the lift and drag components of the assessment must align with how it feels to execute the criterion tasks it represents.
Equipment	Something to hold on to (weight through the hands), attached via a short connector to something to drag, a suitably robust surface, and the ability to add extra resistance where required.	4-5 kg object to hold (e.g., a dumbbell) attached to a 2 m rope attached to a 13" spare wheel (rim + tyre). Level concrete (or non-marking) surface. Additional objects weighing 10 kg to add weight where required.
Intensity of effort	Intensity in this assessment comes from the relationship between the surface and the object being dragged (i.e., the resistance being experienced and thus needing to be overcome), the weight being held in the hands and the haul distance. A heavier object as well as repeated drags increases the effort required.	It is important that the level of effort required during the performance of the assessments is representative of the tasks it is designed to replicate. In particular this extends to the resistance experienced and the number of repetitions performed. The walk recovery between each successive drag repetition forms the natural on-the-job break experienced to drop off the object and go back to collect the next.
Assessment standards	Two performance levels developed with both using the same weight-rope-spare wheel design. Drag distance is constant, but spare wheel resistance and repetitions differ between levels. Level 1 is 6 repetitions of a 20 m drag performed with the spare wheel (i.e., no additional weight). Levels 2 is performed with the spare wheel + 10 kg and 1 repetition of a 20 m drag is required to be completed.	Forward-facing or sideways drag technique permitted. Walking backwards or running is not allowed. There is no completion time set for this assessment, however constant movement during and between repetitions is required to be demonstrated. Level 1 is aligned with Chainsaw Operations, while Level 2 aligns with Firefighting Air Base Support and Road Crash Rescue.

CATEGORY 5: CLIMB / CRAWL

Inputs: a simple two-group split between on-land and in-water tasks was the starting point with this category. These groups were 1) the land tasks that all described being able to climb a ladder, and 2) the water tasks that all described being able to get yourself from in the water to a safe place out of the water. Successful performance of these tasks requires 1) being able to safely climb and descend a ladder up to the height of a single-storey roof, and 2) the ability to extricate yourself from the water (while wearing the appropriate Personal Protective Equipment & Clothing) to a point of safety out of the water (e.g., a vessel or the bank). The second grouping of tasks was decided to be best covered

during water-based testing and has been incorporated into the IN-WATER SAFETY assessment presented in CATEGORY 7: SWIM.

Prototype Assessment: LADDER CLIMB		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound.
Performance elements	Climbing and descending a ladder using the correct technique.	The height to be climbed (measured in number of rungs, for ease of communication) needs to be determined.
Equipment	Commonly used ladder type. A secure roof or wall with stable ladder footing. Safety harnesses.	The ladder type selected to be a commonly used variety and climb height should reflect the reasonable expectation for every qualified operator. Consideration also needs to be given to what is available at an average SES unit.
Intensity of effort	The intensity in this assessment comes from the number of rungs required to be climbed, the angle of the ladder, and any requirement to wear a harness.	Important that the prescribed climb height represents the minimum expectation and also that the angle the ladder is positioned remains consistent with on-the-job practices.
Assessment standards	A single performance level developed. Self-paced assessment with focus on correct technique. Only applied to those members who choose to climb ladders (noting that it is not a mandatory requirement).	One repetition (up and down) satisfies requirements. No time limit is enforced; however constant movement and correct technique, including 3 points of contact at all times, must be demonstrated. This assessment is aligned with General Rescue, Chainsaw Operations and Storm Damage.

CATEGORY 6: LIFT & HOLD

Inputs: the fewest number of tasks is found in this category. Similarities in task performances reside in the static hold positions required. The between-task difference was predominantly around the weight of the equipment used. To successfully execute these tasks the appropriate equipment (i.e., chainsaw or rescue tool) needs to be lifted and safely located in various static positions prior to it being activated.

Prototype Assessment: HOLD		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	On-land.	Important that the assessment can be conducted within the confines of an average SES unit compound.
Performance elements	Static holds in a number of different positions: shoulder height, waist height, and below knee.	The different positions chosen reflect anatomical locations rather than measured heights due to reasons of practicality and real life tool use.
Equipment	Commonly used variety of chainsaw. A combination rescue tool.	The equipment selected to be used should reflect the reasonable expectation for every qualified operator.
Intensity of effort	This is a static assessment (i.e., conducted in place). Intensity comes from the weight of the tool, the position it is being lifted to, and the duration of the hold and between-hold components.	Important that the selection of tools is correct and represents the minimum expectation. Requirement for 2 repetitions in each position (multiple efforts are common) with 5 second hold durations employed to replicate job demands (being able to accurately place the tool where it is intended prior to activating it).
Assessment standards	Two performance levels developed. Performed using the same positions, hold durations and number of repetitions. Level 1 is performed with the chainsaw. Level 2 is performed with the combination rescue tool.	Positions include shoulder, waist, and below knee. Each hold duration is 5 seconds with 2 repetitions at each position and 10 seconds between holds. Level 1 is aligned with Chainsaw Operations, while Level 2 aligns with Road Crash Rescue.

CATEGORY 7: SWIM

Inputs: this category exclusively contains tasks which are performed in-water. The between-task differences which required careful consideration were the nature and behaviour of the water, and the context and urgency of the task. The first grouping of tasks relates to inadvertently finding yourself in the water whilst wearing all appropriate PPE+C and then being able to safely get yourself to somewhere to exit the water. The second grouping being all about getting to a victim in the water and then towing them to an extrication point. The successful performance of these tasks will require a varying degree of effort which will depend on the nature and behaviour of the water. However, it can be seen that the first grouping presents to be more self-survival in nature and thus no real task urgency is present while the second grouping has a life-saving context and thus most certainly possesses a performance urgency component.

Prototype Assessment: IN-WATER SAFETY		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	In-water.	Conducted at a local pool or body of water or safe stretch of river (not flowing water).
Performance elements	Entering the water wearing appropriate Personal Protective Equipment + Clothing (PPE + C) and then safely negotiating a representative distance in the water before self-extricating to dry land. Removing a rescue mannequin from the water onto dry land.	This assessment is designed to replicate inadvertently finding yourself in the water, maintaining safety until a point of extrication presents, and then being able to self-extricate to dry land (or a vessel). The final component of the assessment became a relevant 'add-on' when considered alongside other criterion tasks and also the POWER SWIM & RESCUE assessment to avoid any unnecessary duplication.
Equipment	Minimum 25 m length of water with suitable and stable extrication point. All necessary PPE+C (including personal flotation device (PFD)). Aquatic rescue mannequin (approx. 40 kg).	A 50m pool is considered ideal but workarounds, where this is not achievable, are possible. Rescue mannequin weight is representative of the individual share contributed to the team task it replicates.
Intensity of effort	Intensity comes from the speed travelled in the water, the distance to be negotiated, the bulk of the worn PPE+C (including PFD), the height differential between the water and extrication point, and the weight of the rescue mannequin.	There is no time limit imposed on this assessment regardless of the performance level, and this is reflective of the criterion tasks that it has been designed to represent, none of which include a time-critical urgency.
Assessment standards	Three performance levels developed. Level 1 includes a 25 m distance to negotiate and then self-extricate to dry land. Level 2 is essentially the same as Level 1 but with the removal of the rescue mannequin performed after self-extrication. Level 3 is a 50m distance, self-extricate and then remove the rescue mannequin.	Level 1 is aligned with Land-based Swiftwater. Level 2 is aligned with Boat Operations when no on-the-job flowing water is experienced. Level 3 is aligned with In-water Technician and Boat Operations when there is flowing water encountered during on-the-job tasks.

Prototype Assessment: POWER SWIM & RESCUE		
COMPONENT	KEY DETAILS	COMMENTARY
Environment	In-water.	Conducted at a local pool or body of water or safe stretch of river (not flowing water).
Performance elements	From a stationary start and wearing the appropriate PPC+E (including PFD) perform a 25m sprint swim. Orientate a face down rescue mannequin into a safe position. Rescue tow the rescue mannequin back to the start point.	This assessment is designed to replicate a swimming rescue task where it is important to get to the casualty in as short a time as possible and then tow them to a point of safety where they can receive further assistance. It also provides confidence for the capability to swim across fast flowing water to be able to eddy out and then perform other tasks when on the other side.
Equipment	Minimum 25m length of water. All necessary and role specific PPE+C (including PFD). Aquatic rescue mannequin (approx. 80kg).	A 50m pool is considered ideal but workarounds where this is not achievable are possible. Rescue mannequin weight is representative of an average adult.
Intensity of effort	Intensity during this assessment comes from the speed travelled in the water to get to the rescue mannequin, the distance to be covered, the bulk of the worn PPE+C (including PFD) and the weight of the rescue mannequin.	There is a time limit imposed on the sprint swim component of this assessment, but not the rescue tow. This is reflective of the criterion tasks that it has been designed to represent, where there is a time-critical urgency associated with getting to the casualty. Once contact has been made no such urgency can be justified.
Assessment standards	A single performance level developed. It includes a 25m sprint swim to the rescue mannequin, orientating the mannequin into a safe position, and then rescue towing it back to the start point.	This assessment is aligned with In-water Technician. The cut-off time for the sprint swim component is yet to be established. Whilst the rescue tow does not have a completion time associated with it, it is expected that constant forward motion is displayed.

As has been detailed above, the approach taken by the research team has resulted in the development of several different assessment types and varying within-assessment performance levels. It is at this stage of development that the prototypes are required to undergo testing to understand whether they possess sufficient real-world capability and alignment with skill sets. This kind of testing will also allow for the research team to establish what works and what doesn't regarding assessment conduct, and to refine performance parameters and equipment where appropriate.

3.2.3 Pilot testing

In general terms, pilot testing involves a selected group of end users (volunteers) trying the system under test conditions and providing feedback before the full deployment of the system. In other words, a dress rehearsal for the usability test that follows. Pilot testing helps in early detection of bugs in the system. Without pilot testing, the programme and assessments cannot be viewed as being validation ready. Essentially, pilot testing allows the research team to arrive at

the best possible versions of the programme and prototype assessments before they are taken forward into formal validation trials.

As applied to this project, pilot testing has included presenting to and gaining feedback from the Project Team about such things as the programme design and framework, and prototype assessments as well as a dedicated field trial (pilot trial) that was focussed on the seven land-based assessments.

PROJECT TEAM MEETINGS AND WORKSHOPS (PRE-PILOT TRIAL)

Over the course of the project and leading up to the pilot trial for the land-based assessments, periodic meetings and workshops were scheduled with the Project Team. One purpose of the meetings was to pilot test concepts and prototypes prior to formalising and progressing future works. Two important aspects of the project were tested with the Project Team, these being 1) the Fit for Task Programme and 2) the prototype Fit for Task assessments. After discussions regarding these broader topics, endorsement from Project Team members was sought.

Workshops and meetings were held across February and March 2021. During discussions with Project Team members, there was overwhelming support for the programme framework and prototype assessments. Discussions were held regarding some of the finer details surrounding implementation and monitoring, including equipment use and access to facilities.

PILOT TRIAL FOR LAND-BASED ASSESSMENTS: WANGARATTA, VIC, MARCH 2021

A pilot trial can be defined as a small study to test research protocols, data collection instruments, sample recruitment strategies, and other research techniques in preparation for a larger study. A pilot trial is one of the important stages in a research project and is conducted to identify potential problem areas or opportunities for improvement or refinement.

The prototype assessments that have been developed are designed to replicate the most important, physically demanding and frequently performed tasks (i.e., the criterion tasks). Although every effort is taken to ensure the prototype assessments closely replicate the criterion tasks, as they are a step removed from actual job performance there is no guarantee that they are a valid representation. Therefore, at this stage of the project, it is necessary to ask a group of qualified SES volunteers (spanning the widest possible range of skill sets) to perform the assessments and provide feedback on how well they think the assessments replicate the criterion tasks.

The information and evidence collected during the trial helps the research team evaluate whether the prototype assessments are well-aligned with the criterion tasks as they would be experienced in the 'real world'. It is also the first opportunity for the research team to evaluate whether the assessments that have been created meet the organisational desires and limitations, such as requiring no major equipment purchases or infrastructure projects.

Objectives

The key outcomes of the pilot trial were to modify the prototype land-based assessments based on objective and subjective feedback received from the volunteers. This feedback process is designed to maximise the assessments' representation of the criterion tasks and to ensure the prototypes meet the organisational desires regarding use of common SES equipment. Researchers took this opportunity to refine assessment layouts and performance flow.

Approach

The pilot trial was conducted at the Wangaratta (Victoria) SES facility in March 2021. Volunteers (n=9; Male=5; Female=4) from Victoria and New South Wales participated. All seven land-based assessments, including the various performance levels, were evaluated across the weekend. This included: LIFT & SHIFT (Level 1 and 2); CARRY (Level 1 and 2); LADDER CLIMB; DRAG (Level 1 and 2); HOLD (Level 1 and 2); HAUL (Level 1 and 2), and HIKE (Level 1, 2 and 3). Importantly, volunteers were only asked to perform assessments that were relevant to, and aligned with, the skill sets they were qualified for.

Before each assessment was conducted, volunteers were provided a verbal explanation (including a brief background of the development and the criterion tasks being represented), a physical demonstration of the assessment as well as an opportunity to familiarise and practice. Each assessment was conducted one at a time and volunteers were provided ample time to experience the assessments and converse with the research staff and each other to consider what they

thought was good, bad, or indifferent about the assessment. After each volunteer completed an assessment, they were asked to provide written feedback on a data collection sheet, which included fidelity ratings and general comments.

Fidelity ratings were provided on a visual scale from 0 “no likeness” to 10 “exact likeness”, on how well the assessment replicated the job demands across five rating categories, including: types of movements, distances, height and weight of objects, repetitions performed, and clothing and equipment used. When providing these ratings, volunteers were asked to consider what they would experience in a typical callout for each skill set that was aligned with that assessment.

If an assessment was aligned with more than one skill set, volunteers were asked to complete a separate feedback form for each skill set whilst considering how closely the assessment aligned with the specific demands of that skill set. Importantly, it is these measures of fidelity that help the research team quantify how well the assessments align with the criterion tasks and ultimately, the job demands. The general comments section of the feedback form was used as a way for volunteers to provide more in-depth or specific feedback, which was used to promote detailed discussions during the pilot trial weekend that led to valuable assessment modifications. Some examples of these written comments are presented later in this section.

Data treatment

Due to the qualitative nature of collecting written feedback and likeness ratings, scores of fidelity are most appropriately reported through the use of median scores. Although numerical values are assigned to the ratings, the responses indicate an ordered categorical structure, rather than a numerical value that can be treated arithmetically (e.g., addition, subtraction). Therefore, traditional statistical analysis (i.e., parametric analysis), such as calculating mean and standard deviation, must be avoided (Svensson, 2001). In this project, the prototype assessments have been purposely designed to represent the physical nature of what is experienced during operational call outs and not to be pure replications of job tasks or scenarios, and therefore they do not possess all possible performance elements.

With these factors in mind, the research team was satisfied that there was acceptable likeness and connection between the assessments and ‘real-world’ job performance with fidelity ratings of 7 or above. This score was selected by researchers based on multiple conversations with participants selecting different fidelity scores. A trend has been observed where participants who select 7 or above don’t tend to suggest that the task could be run any differently. A score of 7, 8 or 9 usually implies that the environmental conditions (e.g., temperature), may be different, or there are other factors (such as adrenaline, psychological stress), associated with the task, that cannot be replicated in a simulation. Scores below 7 however, are usually associated with feedback that the task was missing something that would normally occur during operations.

Assessment focus: To understand the overall performance of the seven assessments, all fidelity scores for each individual assessment, regardless of the skill set, were aggregated and the median score was determined as the assessment rating. For example, for the CARRY assessment, all fidelity ratings associated with CARRY from General Rescue, Land Search, Vertical Rescue and Urban Search & Rescue were pooled together and the middle (median) number was taken as the assessment rating.

Skill set focus: The fidelity ratings for each assessment were aggregated for each skill set and presented as the median of the group score. For example, there were six volunteers qualified for Storm Damage, so the ‘movement’ score from each volunteer was ordered and the middle number (median) was taken as the whole group ‘movement’ score. This process was then repeated for each fidelity category in each assessment. These fidelity ratings were calculated for all assessments for all skill sets represented at the pilot trial.

It should be noted that due to there being no assessments aligned with the Air Search skill set, there are no ratings of fidelity presented. Additionally, as there were no Firefighting Air Base Support qualified volunteers present at the pilot trial, no fidelity ratings for the relevant assessments could be collected (please see section 3.2.4 for further discussion of assessments for these skill sets).

For a video representation of the pilot trial assessments, please see the video in the following link:

https://www.youtube.com/watch?v=BO5dzIxbH0g&t=2s

Findings

Based on the overall assessment ratings (presented in Figure 18), the prototypes can be considered a good representation of the criterion tasks as they are experienced by SES volunteers in the real world. Whilst the overall

assessment ratings are all well above 7, based on skill set specific analysis and feedback from volunteers, there was opportunity for improvement.

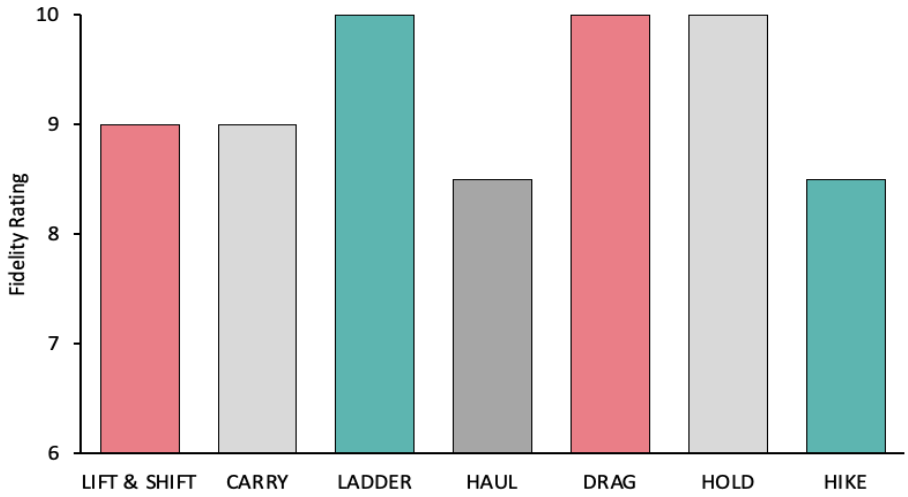


Figure 18: Combined median rating of fidelity from all rating categories for each individual land-based assessment conducted during the pilot trial (Lift & Shift, n=31; Carry, n=15; Ladder, n=17; Haul, n=7; Drag, n=8; Hold, n=8; Hike, n=6).

Ratings of fidelity across all individual rating categories were 7 or above for General Rescue, Storm Damage, Chainsaw, Off-road Driving, Land Search, Urban Search & Rescue and Road Crash Rescue. However, for Vertical Rescue, the 'distance' category in CARRY and the 'clothing/equipment' category in HAUL received a rating of 6. It is important to highlight that only one qualified volunteer was present for this skill set and it is likely that a different result would have been attained if a larger cohort was available. Based on follow-up conversations with the volunteer, the assessments were well-received, and the decision was made to progress the prototypes without need for modification. Individual skill set fidelity results are presented in Table 3.

Table 3 Median fidelity ratings for all pilot assessments provided by qualified personnel in General Rescue (n=6), Chainsaw (n=5), Land Search (n=5), Off-road Driving (n=2), Road Crash Rescue (n=3), Storm Damage (n=6), USAR - Urban Search & Rescue (n=3), Vertical Rescue (n=1).

		General Rescue	Chainsaw	Land Search	Off-road Driving	Road Crash Rescue*	Storm Damage	USAR*	Vertical Rescue
LIFT & SHIFT	Movement	8.5	8	8	8.5	8	9.5	8	7
	Distance	9	10	9	8	9	9	9	8
	Height/weight	8	10	9	9.5	7	9.5	10	7
	Repetitions	10	10	9	10	8	10	10	8
	Clothing	10	10	10	10	9	10	10	7
CARRY	Movement	9	-	8	-	-	-	9	8
	Distance	8	-	9	-	-	-	10	6
	Height/weight	9.5	-	8	-	-	-	10	9
	Repetitions	9	-	9	-	-	-	10	7
	Clothing	10	-	10	-	-	-	10	7
LADDER CLIMB	Movement	9.5	8	-	-	-	9.5	-	-
	Distance	9.5	9	-	-	-	10	-	-
	Height/weight	10	10	-	-	-	10	-	-
	Repetitions	10	10	-	-	-	10	-	-
	Clothing	10	10	-	-	-	10	-	-
HAUL	Movement	-	-	-	-	-	9	-	8
	Distance	-	-	-	-	-	10	-	7
	Height/weight	-	-	-	-	-	10	-	7
	Repetitions	-	-	-	-	-	10	-	7
	Clothing	-	-	-	-	-	10	-	6
DRAG	Movement	-	9	-	-	10	-	-	-
	Distance	-	8	-	-	10	-	-	-
	Height/weight	-	10	-	-	9	-	-	-
	Repetitions	-	10	-	-	8	-	-	-
	Clothing	-	10	-	-	10	-	-	-
HOLD	Movement	-	10	-	-	10	-	-	-
	Distance	-	10	-	-	10	-	-	-
	Height/weight	-	10	-	-	10	-	-	-
	Repetitions	-	10	-	-	10	-	-	-
	Clothing	-	10	-	-	10	-	-	-
HIKE	Movement	-	-	8	-	-	-	-	9
	Distance	-	-	8	-	-	-	-	8
	Height/weight	-	-	9	-	-	-	-	9
	Repetitions	-	-	8	-	-	-	-	9
	Clothing	-	-	10	-	-	-	-	7

*Please see section 3.2.4 *Post Pilot Trial* for further task and assessment clarification.

The conversations had with volunteers and written feedback received across the period were an incredibly valuable aspect of the pilot trial. This form of evidence was rich in detail and afforded the researchers the opportunity to understand the specific roles more deeply, but perhaps more importantly, it helped the researchers understand exactly how to make the assessments even more relevant to the criterion tasks.

In the very early stages of the pilot trial, feedback from SMEs and experienced volunteers indicated that the 50-m distance within the CARRY assessment would not be achievable for the majority of SES locations (and in fact for the compound at Wangaratta). A distance of 30 m was settled on as being achievable in the set up and conduct (organisational desires) and did not compromise the most critical physical components of the assessments.

For the HAUL and HOLD assessments, feedback from the volunteers was very supportive, with no major concerns or suggestions for modifications.

Robust conversations were had about the LIFT & SHIFT assessment. Largely, there was consensus that when carrying items to and from the truck or around the work site, no matter the callout or skill set, there is often the requirement to navigate small obstacles, such as gutters and debris, whilst carrying the items. Examples of written feedback for this include “Look at adding walk on or over or around objects (gutters, puddles)” and “Needs obstacles to negotiate”.

When conducting the LADDER CLIMB assessment, the volunteers agreed that whilst there is an important physical component to climbing the ladder itself, there is almost always a requirement to restow the ladder in the overhead compartment of the truck, or on the roof of the ute. Examples of written feedback from volunteers include “Need to put ladder away” and “Good representation but need overhead movement to put ladder away”.

During the DRAG Level 1 assessment, the chainsaw-qualified volunteers agreed that when clearing felled logs and debris, there are often numerous shorter distance drags, rather than few longer distance drags. Examples of written feedback for this include “Very well replicated, shorter distance but more repetitions” and “Shorter course with more repetitions”.

During the DRAG Level 2 assessment, task clarification was sought from the Road Crash Rescue qualified and experienced volunteers that were present. The consensus was that there were two (vehicle mounted) hydraulic hose reels needing to be pulled out, each approximately 20 m in length. Often this task was performed by a single person (one hose at a time) over uneven terrain and often negotiating obstacles. Integral to the task as described by volunteers is a requirement to stop and then start again during the drag task.

After completing the HIKE assessment, feedback from the volunteers was very positive, but they felt there could be more frequent performance of the obstacles. The volunteers described that when conducting land searches, there are often numerous branches, bushes, or structures to navigate, rather than a few. Examples of written feedback for this include “Greater frequency of obstacle performance required” and “Needs more obstacle repetitions”.

Learnings and Modifications

Researchers were satisfied that the prototype assessments were developed to a level that closely represents the demands of the criterion tasks as they would be commonly experienced in the real world. Whilst overall ratings were all above 7, the feedback received indicated there was room for improvement on some prototypes. As described earlier, a key outcome from the pilot trial was to modify the prototype assessments as deemed appropriate to maximise the representation of the criterion tasks. The main modifications arising from the pilot trial were:

1. Reduction of the distance in the CARRY assessment from 50 m to 30 m.
2. Addition of three different obstacles during the 30-metre ‘ground-to-ground’ carry section of LIFT & SHIFT;
3. Addition of an overhead lift component to LADDER CLIMB to replicate stowing it away on a vehicle. This resulted in a renaming of this assessment to LADDER CLIMB & LIFT;
4. Reduction of drag distance and increase in number of repetitions in DRAG Level 1;
5. DRAG Level 2 increased to two repetitions with a brief pause to be performed at the halfway point; and
6. Increased performance frequency of obstacles during HIKE, alteration to step up heights, alteration to the height of the crawl under obstacle, and the removal of the climb over (or scale) obstacle.

3.2.4 Post pilot trial

UPDATES TO ASSESSMENTS

Firefighting Air Base Support: as there was no Firefighting Air Base Support qualified or experienced volunteer present at the pilot trial, a follow up discussion was scheduled with an experienced SME for comment on the fidelity of the

assessments. LIFT & SHIFT Level 1 and DRAG Level 2 were believed to be valid and appropriate representations of the physicality required not only of the skill set criterion tasks but also of performing the role in general.

Road Crash Rescue: after clarification was provided by SMEs regarding the task "Remove the casualty from vehicle" and the inclusion of a requirement to transport (i.e., carry) the casualty up to 100 m to get further medical assistance it was decided that this new information elevate the task to criterion and resulted in the addition of CARRY Level 1 to the suite of assessments aligned with the skill set.

Urban Search & Rescue: further conversations and task analysis resulted in the addition of HIKE Level 1 to the assessments aligned with this skill set in order to provide a better level of coverage to the criterion task "Perform repetitive scrambling over rubble and unstable debris".

PROJECT TEAM MEETING: MAY 2021

Representatives from all states and territories attended this meeting and were privy to being updated on the project including the findings from the Wangaratta pilot trial and any subsequent modifications and updates made to the assessments. Additionally, researchers presented the Project Team with the prototype water-based assessments.

Overwhelming support was expressed by the Project Team for all the prototypes. The commentary and debate that was had during this meeting was primarily focussed on the implementation and integration of assessments into what may currently be in existence in certain jurisdictions around the country, rather than on the specifics of assessments or overall design of the programme. Agreement was reached that the project was in a good position and the Project Team was satisfied to progress the prototype assessments to validation trials.

3.3 Evalution phase

The evaluation phase is aligned with the OUTPUTS (or KPIs) of the development process and forms the basis for accountability. The two key outputs during this phase of the research are 1) the assessments being validation ready, and 2) the assessments inspiring 'upstream' and 'downstream' confidence. As a reminder, validation ready = the prototype is (at least) the 80% solution (in reference to design and performance characteristics) and capable of progressing to the next research stage (i.e., entering validation trials). And inspiring confidence = the prototype must ensure that it captures the requisite physicality of what it has been designed to represent (i.e. the 'upstream' criterion tasks), and it must be able to meet the expressed desires of end-users and be able to be implemented and supported ('downstream' usability).

OUTPUT 1

Confidence that the programme and prototype assessments are ready to move to the next research stage is evidenced by quantified and documented findings of the pilot trial (both objective and subjective) and the subsequent approval and sign-off given by the Project Team. The ability of the research team to make slight but meaningful modifications to the prototypes as a result of the pilot has added to the quality of the product. Each assessment now has a clear and detailed set of instructions that enable it to be conducted in a repeatable and organisationally appropriate manner.

Fit For Task Assessment	Types and Associated	Performance Levels:
1. Lift & Shift	2 performance levels	movement of equipment/stores
2. Carry	2 performance levels	stretcher carry
3. Ladder Climb & Lift	1 performance level	on-ladder then stowing it
4. Haul	2 performance levels	pulling rope/object toward you
5. Hold	2 performance levels	static hold (chainsaw/rescue tool)
6. Drag	2 performance levels	dragging something behind you
7. Hike	3 performance levels	walking wearing a backpack
8. In-water Safety	3 performance levels	water safety and (self) rescue
9. Power swim & Rescue	1 performance level	sprint swim and tow rescue

The design specifics and conduct of each assessment and performance level is described in detail in a separate part of the report (please see section 5). The combination of assessment type and performance level allows for a tailored

approach when it comes to aligning assessments (and performance levels) with the physical demands of individual skill sets and implementing the Fit for Task programme in the real world. The alignment and allocation of assessments (and their associated performance levels) to skill sets is illustrated in the ASSESSMENT MATRIX (please see section 5).

OUTPUT 2

The research team acknowledges that for a project such as this to be successful it not only requires a scientifically robust process, but it must also pass the judgement of those it has been developed to assess, the volunteers. Additionally, the programme and the individual assessments need to be capable of being implemented by each state and territory. The keys to achieving success in the aforementioned are to design assessments that are functional in nature and look and feel like the job tasks they have been developed to replicate, and to develop a programme that is simple, efficient and uses (for the most) equipment and assets already in-service.

The overwhelming response from the volunteers involved in the pilot trial was that the assessments were highly functional and the movements and actions they were required to perform were very similar to the ones experienced regularly on the job. This came through in the reported fidelity ratings and in conversations and discussions. The fact that the research team was also able to make several modifications utilising the feedback received has added to the job-like feel and connectedness of the assessments.

Having the ability to keep the Project Team periodically updated on the progress of the Fit for Task project and being able to harness their thoughts as to what would be appropriate in regards to programme design assisted researchers in developing a capable and implementable programme. Notably, at each stage the Project Team provided their approval and sign off.

What's next?

Now the programme and prototype assessments have been pilot tested and modified, where appropriate, they are considered to be in their most mature state. It is necessary for these final assessments to be performed by a wider range of volunteers over multiple occasions to ensure the assessments are replicable in other settings and with other participants. The process provides a level of confidence that the assessments are capable of producing results for the wider organisation that are truthful and meaningful.

4 Validating assessments

4.1 Introduction

The seven land-based and two water-based assessments (including their various performance levels) are now developed to a level that will very closely represent their 'final format'. To date, the assessments have been born out of the research team's collective and creative ideas based on the most important, frequent and physical tasks (i.e., the criterion tasks). Further, the initial assessment designs have been performed by a subset of volunteers who have provided detailed feedback, coupled with valuable conversations with SMEs who provided important feedback from an organisational perspective (i.e., pilot testing). Lastly, the assessment designs have been reviewed and modified based on this feedback and are now considered more robust, suitable and ready for amalgamation into the SES Fit for Task Programme.

The final step in the scientific process is to ensure the results and feedback that have been provided on the assessments to date are congruent with results and feedback from a wider range of volunteers and SMEs across the country. This process is known as 'Validation' and can be thought of as ensuring that the results we have observed to date aren't unique to the specific volunteers and SMEs that have been involved in the aforementioned steps. Further, validation allows the research team to determine the impact, if any, of having different assessors conduct the same assessment on the same individual.

The aim of this stage of the project was to validate all nine assessments (and the various performance levels) by conducting multiple trials with a broad range of SES volunteers. It was expected this process would provide a level of confidence that the assessments which progress from this stage will be representative of the job demands for the entire population of SES volunteers and will produce consistent outcomes regardless of the assessor.

4.2 Methods

Given the ongoing delays of intra- and inter-state travel with the COVID-19 pandemic, researchers were unable to travel to conduct on-site validation procedures. After several adjustments to the research design, it was agreed between HPS and the Project Team that the only way to complete the validation was remotely. From these discussions, the idea of 'Validation Socialisation' was developed.

Validation Socialisation required states and territories to run their own validation procedures following instructions provided by HPS. This initiative was driven by state and territory representatives who coordinated validation trials amongst targeted regions and units. All states and territories were provided a 6-week window to complete validation procedures.

To accommodate for the different levels of commitment that states and territories could afford to provide to this stage of the project, three different categories of engagement were offered: ALPHA, BRAVO and CHARLIE. The categories differed in the data that was collected and the type of feedback that was required. Regardless of the level of engagement and subsequent data collection requirements, it was anticipated the 6-week window would provide ample time for numerous units across the country to complete some level of validation. Each category is outlined in Table 4.

Table 4 Brief description of the three Validation Socialisation engagement options.

ALPHA	BRAVO	CHARLIE
● Performance of each volunteer recorded ● Time to completion recorded ● RPE measures recorded ● Test-retest (volunteers perform assessments on two occasions) ● Volunteers complete online surveys for feedback ● Assessors complete online survey (in group) for feedback ● Priority to complete HIKE assessment	● Performance of each volunteer recorded ● Time to completion recorded ● RPE measures recorded ● Single test only (volunteers perform assessments once) ● Volunteers complete online surveys for feedback ● Assessors complete online survey (in group) for feedback ● Priority to complete HIKE assessment	● Performance of each volunteer recorded ● Time to completion (only for HIKE and POWER SWIM) ● Single test only (volunteers perform assessments once) ● Volunteers complete online surveys for feedback ● Assessors complete online survey (in group) for feedback ● Priority to complete HIKE assessment

All participating units were free to choose the engagement option that best suited them. However, an emphasis was put on completing assessments with a timed component (i.e., HIKE L2 and L3 and POWER SWIM & RESCUE). Information packages containing all necessary information to complete all three engagement options were sent to state and territory representatives to be disseminated to interested units. This information package included:

1. A detailed "Start Here" document thoroughly explaining the background of the project, descriptions and explanations of the resources provided, explanations of the different engagement levels, FAQs, detailed descriptions of how to record the information correctly, detailed descriptions and explanations of the feedback surveys, an example timeline of how to conduct the assessments and contact details of the research team.
2. A short video illustrating examples of some of the land-based assessments.
3. An assessment matrix detailing the assessments that each SES skill set is required to perform.
4. Individual instructions for each of the nine assessments, including diagrams, written instructions and word-for-word scripts that should be read aloud to all participants prior to each assessment.
5. Data recording sheets appropriate for each of the three engagement options.
6. RPE scale.
7. QR codes for convenient access to the feedback surveys.

PERFORMANCE DATA

For all three engagement options, units were required to record whether the volunteers passed or failed. This pass or fail data required assessors to note the performance level the participant was attempting, a tick or a cross denoting a pass or fail and how many attempts were undertaken. For the ALPHA and BRAVO options, assessors were also required to record how long it took each volunteer to complete each assessment in minutes:seconds format.

SURVEYS

At the completion of each testing session, two surveys were available to participating units. Each participant was asked to provide feedback on the 'individual participant survey' and the assessors at each unit were asked to provide feedback as a group on the 'assessor group survey'.

First, the individual participant survey was primarily used to help the research team understand how well the assessments represented the type of physical work the volunteers experience during call outs and also how well the types of assessments covered the physical movements and demands of the skill sets. Questions of this nature included: "How well do you feel that the Fit for Task assessments represented the type of work you do on a turn out or during training?" and "In regards to the individual skill sets you hold current qualifications for, and are operational in, how well do you think the physical requirements are addressed by the Fit for Task assessments that are aligned with them?".

The questions required participants to provide their response on a sliding scale from 0 to 100, where 0 represented "no representation" and "did not address at all", respectively, and 100 represented "perfect representation" and "perfectly

addressed", respectively. This survey also asked participants to rate on a 7-point Likert scale how well they believe the assessments were received by themselves and their peers. This rating scale ranged from "negatively" to "positively". The survey also included questions regarding some demographic details and provided an opportunity for participants to add any further comments they thought might be relevant, whether negative or positive with respect to the new assessments.

Second, the assessor group survey was primarily used to help the research team understand how the assessors experienced the set-up process, conduct, and recording of assessments. This survey included questions regarding assessment feasibility and sourcing equipment, communicating expectations to volunteers, setting up the assessments, the comprehension of assessment instructions, and how representative the sample of volunteers they assessed was with respect to their entire unit. These questions required assessors to provide their response on a sliding scale from 0 to 100, where 0 represented "not at all helpful" or "very challenging", and 100 represented "very helpful" and "very easy", depending on the question. In total there were 11 questions with a rating scale and each of these was accompanied by a short answer option for assessors to provide additional comments to their rating. This survey also asked assessors to rate on a 7-point Likert scale how well they believe the assessments were received by the participants. This rating scale ranged from "negatively" to "positively". The survey also included questions regarding some demographic details and provided an opportunity for assessors to add any further comments they thought might be relevant, whether negative or positive with respect to the new assessments.

DATA TREATMENT AND STATISTICS

Performance data validation

Prior to the commencement of the Validation Socialisation, the research team predicted expected completion times for the majority of the assessments. These times were based on a combination of average movement speeds coupled with other associated assessment-specific factors (e.g., changes of direction, manual handling requirements, negotiating obstacles). The predicted average movement speed ranged between 65 m/min and 100 m/min (i.e., roughly 4-6 km/h) and was based on an average to brisk walking pace (Waters et al., 1988). For example, the LIFT & SHIFT L1 assessment is 435 metres in length from start to finish. Walking at 65 m/min and 100 m/min, this assessment should take approximately 6:40 and 4:20 to complete, respectively. Therefore, it would be expected that any volunteer walking at operational tempo, should complete this assessment within, or close to, this time frame. Any performance times that were significantly faster or slower than the predicted times were identified as requiring further investigation. Further investigation considered the time it would ordinarily take to complete the associated assessment-specific factors, such as handling equipment or navigating obstacles, combined with the expected locomotion duration. If, after this investigation, the data were considered invalid, they were subsequently omitted from further analysis.

Test-retest reliability

To assess test-retest reliability for any units that chose the ALPHA option, both coefficient of variation (CV%) (Shechtman, 2013) and intraclass correlation coefficients with 95% confidence intervals (ICC, Model 3,1: Hopkins et al., 2009; Weir, 2005) will be calculated between Trial 1 and Trial 2 for all performance time data. Although there is no standard for classifying the magnitude of ICC values (Nunan et al., 2009), retest correlations of >0.81 have been described as suitably 'trustworthy' (Hopkins, 2000). The CV% is a unitless percentage between 0 and 100, where a number closer to 0 represents better reliability. More accurately, the CV% is a representation of the standard deviation as a percentage of the mean, where a lower CV% represents smaller variation from trial to trial (Schechtman, 2013). The ICC is a unitless value between 0 and 1, where a number closer to 1 represents better reliability. There are different models of ICC that are suitable for either test retest or intrarater reliability and model 3,1 will be employed for calculating test-retest reliability in the current phase of the project (Weir, 2005).

4.3 Results

ASSESSMENT PERFORMANCE DATA

Demographic information

The Validation Socialisation was completed by a total of 139 participants across 7 SES units representing Queensland (2 locations), NSW (2 locations), SA, WA and VIC. For a full list of participant demographics, see Appendix A. For most of this results section, the data will be de-identified and presented as Unit 1 through Unit 7.

Assessment coverage

Of the 18 Fit for Task assessments (when accounting for assessment type and performance levels), a combined total of 13 were completed during the Validation Socialisation. Assessments that were not attempted during this phase include HIKE L1, IN-WATER SAFETY L1, L2 and L3 and POWER SWIM & RESCUE L1.

Performance data

All 7 participating SES units provided performance data in the form of performance times and pass/fail results. Performance times represent the total time it took each participant to complete the assessment. For the HOLD L1 and L2 assessments, only pass/fail data was reported, as each HOLD assessment involves a standardised completion time. The pass/fail results simply represent how many participants were deemed to have satisfactorily passed the assessment or not. These results have been summarised in section 'Pass & Fail'. It should be noted that no reliability analysis was conducted as insufficient test-retest data were provided by participating units.

Performance times

Performance time data from all SES units are summarised in Table 5. Reported performance times from all participating units were compared to the predicted expected completion times (see methods for explanation). These predicted times are included in Table 5. Based on the predicted times, the majority of the data were thought to be valid, however, there were some examples of invalid data. These invalid cases are shown in red text in Table 5.

Table 5 Reported performance data (time to completion) for all units that participated in the Validation Socialisation. Performance data is reported as a range from lowest to highest of all participants who completed each assessment at each unit. The total distance and predicted completion time are presented where appropriate. Red text indicates invalid data. m, meters; km, kilometres; L1, level 1; L2, level 2; n, number of participants.

			REPORTED PERFORMANCE TIME (min:sec)						
ASSESSMENT	**DISTANCE**	**PREDICTED TIME**	**Unit 1**	**Unit 2**	**Unit 3**	**Unit 4**	**Unit 5**	**Unit 6**	**Unit 7**
LIFT & SHIFT L1	435 m (240 / 195)	4:20 - 6:40	-	-	1:55 - 3:47 n=27	-	4:00 - 6:40 n=21	-	-
LIFT & SHIFT L2	645 m (360 / 285)	6:37 - 10:14	9:00 n=1	5:10 - 5:56 n=5	-	-	-	-	-
CARRY L1	60 m	0:35 - 0:55	-	0:43 - 0:51 n=3	No time reported n=21	-	0:40 - 1:10 n=21	-	-
CARRY L2	120 m	1:11 - 1:50	1:21 - 2:06 n=9	1:23 - 1:32 n=3	-	-	-	-	-
LADDER CLIMB & LIFT	NA	NA	0:45 - 1:07 n=10	0:20 - 0:26 n=5 (of 11)	0:22 - 1:12 n=24	-	0:24 - 4:00 n=21	-	-
HAUL L1	20 m	Approx. 0:30	0:24 - 0:53 n=7	0:14 - 0:24 n=7	0:14 - 0:48 n=27	0:10 - 0:32 n=24	0:10 - 0:25 n=21	-	-
HAUL L2	60 m (40 / 20)	Approx. 1:20	1:37 - 2:23 n=4	0:41 - 1:08 n=5	-	-	-	-	-
DRAG L1	190 m (100 / 90)	1:52 - 2:53	2:16 - 4:49 n=11	2:53 - 3:18 n=2 (of 8)	0:09 - 0:27 n=27	0:45 - 1:36 n=26	1:45 - 3:25 n=21	-	-
DRAG L2	60 m (40 / 20)	0:35 - 0:55	-	0:45 - 0:53 n=5	-	-	-	-	-
HIKE L2	4 km (+ 8 x obstacles)	39:13 - 60:36	-	48:00 - 52:40 n=7	-	-	-	50:00 - 59:00 n=3	42:32 - 54:35 n=11
HIKE L3	6 km (+ 13 x obstacles)	58:49 - 90:55	-	-	-	-	-	63:00 - 83:00 n=13	-

Any data that was determined to be invalid was removed from any further analysis. Please see below for descriptions and explanations of the invalid data.

LIFT & SHIFT: L1 from Unit 3 and L2 from Unit 2. Completion times reported by Unit 3 for L1 ranged from approximately 2 - 4 minutes. Additionally, Unit 2 reported completion times for L2 ranging from 5 - 6 minutes. Importantly, the total distance required to be covered in these assessments is 435 m and 645 m respectively. Included in this total distance for both performance levels are multiple changes in direction, lifting and lowering of equipment, and obstacles to be negotiated. At the lower completion time provided by Unit 3, participants had an average speed of around 220 m/min (13 km/h), while at the upper completion time this average speed is approximately 115 m/min (7 km/h). In the instance of Unit 2 (and L2), reported average movement speeds ranged between 125 m/min (8 km/h) and 110 m/min (7 km/h). Considering the stop/start nature of the assessment, coupled with the requirement to carry items and negotiate obstacles, it appears unlikely that these assessments could have been conducted in accordance with the instructions (that stipulate moving at what would be considered operational tempo).

DRAG: DRAG L1 from Unit 4 and Unit 3. Completion times ranging from 45 seconds to approximately 90 seconds were provided. The total distance to be covered in this assessment is 190 m (10 x 10-m drags and 9 x 10-m recovery walks). Included in the total distance are multiple changes in direction, dragging external load, and picking up and placing down the handle attachment. According to the completion times (fastest to slowest) provided by Unit 4, this would equate to average movement speeds of between 250 m/min and 125 m/min (or 15 km/h and 7.5 km/h). Similar to LIFT & SHIFT, it appears unlikely that the assessment was conducted in accordance with the instructions. Unit 3 provided performance times that were much quicker than Unit 4 and are therefore also considered invalid.

HAUL: HAUL L2 from Unit 1. Completion times provided by Unit 1 for this assessment ranged from around 1.5 to almost 2.5 minutes. This assessment includes 2 x 20-m hand-over-hand hauls separated by a 20-m walk, for a total assessment distance of 60 m. The performance time range reported equates to an average movement speed of between 40 m/min (2.5 km/h) and 25 m/min (1.5 km/h). Instructions stipulate that during the haul movement there is to be constant motion maintained throughout (unless a brief pause is required to adjust grip). The inflated times provided could tend to indicate lengthy rest periods were taken during either or both of the two haul movements or a large time gap was observed between them (where only a 20-m walk is required to be performed). Considering these factors, it seems unlikely that the assessment was conducted in accordance with the instructions.

PASS & FAIL

When all individual assessments are combined, there were a total of 393 valid attempts completed by all participants. Of the 393 attempts undertaken, 392 attempts were successfully passed on the first attempt. This is inclusive of the HOLD assessments for L1 and L2, where there were 63 and 4 successful attempts, respectively. Only one failed attempt was reported, which was for the CARRY assessment (Unit 3).

SURVEY FEEDBACK

Participating units provided feedback on the Validation Socialisation in either the individual participant survey or the assessor group survey. The feedback provided by participating units is summarised in the following two sections.

Individual participant survey responses

All participants who completed assessments were given an opportunity to provide written feedback on their experience with the assessments. There was a total of 43 survey responses from all participating volunteers (approximately 31% of total participants) across the SES units that took part (see Appendix B for demographics).

As a result of SES units performing a different mix of Fit for Task assessments and appearing to adhere to assessment instructions and techniques to varying degrees of accuracy, survey responses have been categorised into 3 groups: INVALID, HIKE ONLY and MIXED. It was necessary to isolate INVALID responses for this aspect of the analysis to ensure the volunteers' feedback was valid and relevant to the assessments as they were intended to be performed. For example, written feedback provided by volunteers who were deemed to have completed the new assessments incorrectly, would be based on an experience that is not comparable to those who completed the assessments appropriately. As such, their experience may be unfairly tainted by an assessment that was not setup, performed or administered correctly.

INVALID: Unit 3 (n=9); these responses were isolated due to survey feedback indicating that inaccuracies in and violations to assessment conduct, performance technique and equipment were likely experienced.

Actual feedback received included:

- "The 25L carry of a plastic container (in lieu of a kettlebell) did not have an appropriate handle to allow correct posture and grip. This inhibited participants' ability to complete the task which is not necessarily a reflection on the members' capabilities but also goes against most manual handling procedures and may cause injury."
- "The body mechanics and manual handling of some of the tasks were inappropriate. When carrying a 25L water container it would have been more appropriate to balance the load and carry 2 half or 3/4 filled drums."
- "Chainsaw positioning should not ever raise the chainsaw above shoulder height."
- "Tire drag is a poor one-armed shoulder drag putting excessive strain on the shoulder and back. Would have been better to allow them to drag with two arms behind with squared shoulders and hips. Would also have been better to have either a shoulder hook, or a two-hand grip against the hip on one side."

HIKE ONLY: Units 6 and 7 (n=13 and n=7, respectively); these responses were grouped together as they only have application to a narrow portion of the Fit for Task assessment programme. Actual feedback received included:

- "The Fit for Task assessment was much more realistic in the context of land search than the Arduous & Moderate Pack tests, which are way too fast with a much too heavy pack. I think the Fit for Task assessment felt much more like going on a land search, rather than the MPT/APT which is almost at a running speed. The Fit for Task assessment is the minimum fitness required though for basic searches - a lot of searches require a higher fitness level."
- "Significantly better than the APT. Less stress on the body and more aligned with what we do in the field."
- "The level 3 test for rugged search is comfortably completed at operational pace in under 80 minutes with a pack load of 12 kg as the recommended weight for day searches in rugged terrain."
- "If I could change just one thing, it would be to put a minimum and maximum time on the assessment. I can do the assessment in 65 mins, but speed is not the objective of searching. A time window would take the focus away from attempting to do it in the shortest time ever and put it back to endurance and technique. I also did the course in 123 mins, which I believe is a better representation of what pace a search should be done at. I believe the obstacle activities are a great idea. Thank you."

MIXED: Units 1 and 5 (n=3 and n=11, respectively); these responses came from participants that performed a mix of land-based assessments (excluding HIKE).

Actual feedback received included:

- "Fun but time consuming for night training."
- "The drag and haul assessments were impacted by the surface they were performed on and will likely return inconsistent results unless the surface is specified."
- "Also the vertical rescue role I usually perform (edge attendant or escort) does not require the haul abilities required by the task. So the test doesn't reflect the requirements for the role."
- "The testing needs scaling for age and gender!!"
- "Working with a chainsaw above shoulder height is not realistic."
- "The ladder test is confusing. Putting up the ladder should be separated from climbing the ladder."
- "Need a baseline endurance activity. Ideally incorporate level 1 land search as part of general rescue."
- "Tasks take too long and will be a pain for units with lots of members."

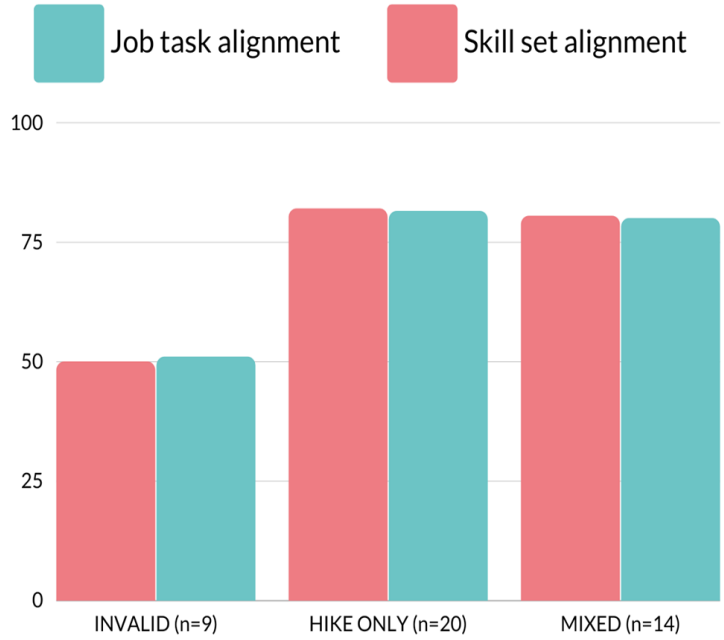


Figure 19: Individual survey responses for questions pertaining to how well the assessments represented the type of work performed during call outs and training (Job task alignment) and how well the physical requirements of individual skill sets are addressed by the Fit for Task assessments (Skill set alignment). Median responses presented. 0 = lowest possible rating and 100 = highest possible rating.

The results for survey questions regarding how well participants thought the assessments captured the physical nature of the tasks they regularly perform and the degree to which assessments align with the skill sets they are associated with, are presented in Figure 19.

The results for survey questions regarding participants' overall feelings towards the Fit for Task assessments and programme and if they considered it is a realistic expectation of each and every volunteer, are presented in Figure 20.

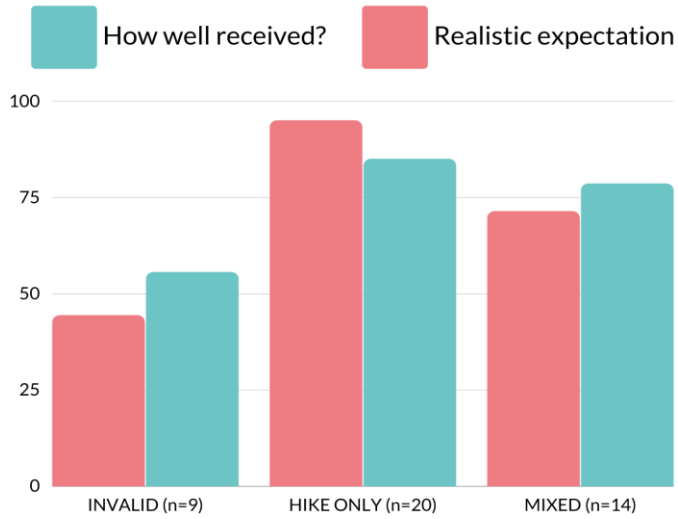


FIGURE 5: INDIVIDUAL SURVEY RESPONSES FOR QUESTIONS PERTAINING TO A PARTICIPANT'S OVERALL FEELINGS TOWARDS THE FIT FOR TASK ASSESSMENTS (HOW WELL RECEIVED?; BLUE) AND PROGRAMME AND IF THEY CONSIDERED IT TO BE A REALISTIC EXPECTATION OF EACH AND EVERY VOLUNTEER (REALISTIC EXPECTATION; RED). BLUE = COMBINED 'POSITIVELY' AND 'MOSTLY POSITIVE' RESPONSES EXPRESSED AS A PERCENTAGE OF TOTAL RESPONSES. RED = NUMBER OF 'YES' RESPONSES EXPRESSED AS A PERCENTAGE OF TOTAL RESPONSES.

Assessor group survey responses

A total of 3 units responded to the Assessor Group Survey. Whilst the majority of questions within this survey align with implementation, as opposed to the scientific development and validation aspect, two questions present as being of interest here.

The first question of interest (Q1) requests information regarding how well the volunteer sample they recruited for validation represented their wider volunteer population. This is important as it speaks to the degree of confidence we can have when extending and extrapolating the findings. The second question (Q2) is focussed on the connection between what is required of volunteers during the assessments and the physical work they commonly perform during call outs and training. Both questions utilised a sliding scale response where 0 = lowest possible rating and 100 = highest possible rating. The numerical responses are summarised below.

Q1. Representative sample:

- Unit 1 - 100,
- Unit 6 - 80, and
- Unit 7 - 84.

Q2. Job task alignment:

- Unit 1 - 66,
- Unit 6 - 100, and
- Unit 7 - 84 (accompanied with the comment "Definitely an improvement over the MPT").

4.4 Discussion of findings

The following sections are designed to answer the salient questions the research team were looking to answer from the Validation Socialisation stage. It is noted that herein, the commentary and inferences around the data relate to all the 'valid' or usable data. The research team has not included the INVALID data as part of this commentary. The data which has been assigned INVALID includes performance times that appear outside the realm of possibility for some of the assessments, with respect to moving speed and completion time. It appears that the units who set up these

assessments have done so in a way that differs from the intended design. Consequently, there is a lack of confidence in using any of the data from these units as it can be reasonably assumed that many, if not all of the assessments, were not run according to the intended design.

Though the identification of the INVALID group decreases the overall data set, it does provide a valuable lesson in the context of implementation. Namely, the importance of clear communication of the assessment instructions and the checks and balances needed to ensure consistency in the conduct of assessments. It has identified to the research team that instructions may need to be more explicit and prescriptive, so that there is no ambiguity in setting up or running assessments. It has also highlighted how SES agencies need to ensure cross checks for each assessment, ensuring multiple personnel are available to set up the assessments adequately. These lessons are elaborated on within the SES *Fit for Task Implementation Strategy* document.

WERE THE ASSESSMENTS VALID AND RELIABLE PREDICTORS OF SES TASKS?

The Validation Socialisation comprised 18 assessments, accounting for all performance levels. Currently, 13 of those assessments have been appropriately performed by qualified SES volunteers (n=132). Of those 13 assessments, 5 have been performed by 13 or less people on one occasion. Additionally, only small numbers of surveys were completed for the assessments that were performed. With these numbers in mind, we consider the validity and reliability of the assessments.

Validity is the capacity of an assessment to measure what it should be measuring, in this case the physical capacity to carry out criterion tasks. Validity was measured using fidelity scores, where participants provided subjective ratings for the assessments they completed. In large, the fidelity scores supported the suitability of 13 out of 18 assessments that were completed. However, as previously commented, survey numbers were very low, with no more than 13 responses for any individual assessment. Evidently, despite there being a trend to suggest there is good support for the assessments, further data would be encouraged to boost the confidence in this assessment.

Reliability is the measure of consistency of an assessment. A participant should receive the same result regardless of when they are assessed, or who assesses them. No single assessment was measured for reliability during the Validation Socialisation stage. However, reliability is only applicable to performance measured on an ordinal, or continuous, scale, such as time to completion. For all the SES Fit for Task assessments, this is only applicable to HIKE L2, HIKE L3 and POWER SWIM & RESCUE L1. Due to the untimed nature of most assessments, it is unlikely conducting any formal reliability investigations on these assessments will provide novel or unexpected support or refutation for their reliability.

In summary, the Validation Socialisation stage was only able to provide partial validity of the Fit for Task assessments. Since participant numbers were low, further data collection, during soft implementation efforts, could provide more confidence in the validity of the assessments. Furthermore, no reliability data was collected on the relevant assessments: HIKE L2, HIKE L3 and POWER SWIM & RESCUE L1. The implications of this are discussed in the *Summary and Next Steps* section.

WAS THERE SUFFICIENT DATA TO DETERMINE PERFORMANCE STANDARDS FOR THE ASSESSMENTS? WHETHER ‘MEETS REQUIREMENTS/ DOES NOT MEET REQUIREMENTS’ OR TIME-TO-COMPLETION.

Of all the Fit for Task assessments, the HIKE L2, HIKE L3 and POWER SWIM & RESCUE L1 assessments require a time-to-completion standard, ensuring volunteers demonstrate an intensity within the assessment reflective of safely executing a criterion task. All remaining assessments have a binary ‘complete’ or ‘did not complete’ outcome (i.e., there is no requirement to complete the assessment within a given time).

When analysing time-to-completion data, researchers can determine a suitable cut score (performance standard) based on the distribution of data, assuming it is normally distributed (Zumbo, 2016). The distribution of the data refers to the mean scores and the standard deviation of the scores. For the data collected amongst the two assessments requiring a performance standard (HIKE L2 and HIKE L3) the data were treated to account for 95% of the population who completed the assessments, equivalent to two standard deviations. This threshold was determined since it can be assumed all participants who completed the assessment are competent, and all their completion times represent their version of operational tempo. However, the data indicated that the difference between the slowest participant and the second slowest participant was greater than all other inter-participants differences; subsequently removing the slowest time was deemed appropriate.

Using the mean and standard deviation, the HIKE performance standards can be calculated by accounting for the mean, the standard deviation and a 5% safety zone. The 5% safety zone (Beck et al. 2016), ensures that the standard is elevated slightly, to ensure that volunteers who meet the standard are not doing so at maximum exertion. If someone needs to exert maximum effort to achieve the minimal performance standard, they pre-dispose themselves to a high risk of injury, compared to someone who is achieving the standard at a sub-maximal exertion level. Once this 5% safety zone had been factored into the calculation, the final time was rounded to the nearest 5-minute interval, for practical reasons when timing the assessment.

The HIKE L2 assessment saw times (in minutes:seconds) range between 42:32 and 59:00. A total of 21 participants completed this assessment, with an average time of 50:10 and a standard deviation of 03:39. When calculating the performance standard for HIKE L2, the time volunteers should be able to complete the assessment is 55 minutes. With 8 minutes of time taken up with obstacles, this leaves an average walking speed of 5.1 km/h for the hiking components.

The HIKE L3 assessment saw times (in hours:minutes:seconds) range between 1:03:00 and 1:23:00. A total of 13 participants completed this assessment, with an average time of 1:12:55 and a standard deviation of 04:35. When calculating the performance standard for HIKE L3, the time volunteers should be able to complete the assessment is 80 minutes. With 17.5 minutes of time taken up with obstacles, this leaves an average walking speed of 5.8 km/h for the hiking components.

For both HIKE L2 and L3 assessments, there is also the consideration of a *minimum* time volunteers should complete the assessment. Assessments have been designed to provide confidence that a volunteer who meets the performance requirements can safely perform the job task. Therefore, the outcome of the assessment becomes a matter of being competent or not competent, without there being graded or ranked levels of assessment performance. Assigning a minimum time accounts for safe performance of the assessment. If volunteers are moving too quickly, it is unlikely representative of a team-based hiking tempo. Further, moving too quickly can be indicative of shuffling or running and result in abnormal gait kinematics, which is not safe when carrying external loads.

For HIKE L2 the fastest completion time from the Validation Socialisation population was 45 minutes, equivalent to hiking at 6.5 km/h (excluding obstacles). Since the boundary between walking and running (commonly referred to as preferred transition speed) typically occurs at approximately 2 m/s (or 7.2 km/h) for normal populations, without external load (Hreljac, 1993), 6.5 km/h is considered the maximum safe walking speed for the HIKE assessment. Therefore 45 minutes is the recommended minimum completion time for HIKE L2 assessment. For the HIKE L3 assessment, this calculates to being a minimum time of 70 minutes.

For the POWER SWIM & RESCUE L1 a performance standard of 40 seconds has been assigned for the 25 m power swim component. Though no data has been collected during the current Validation Socialisation period, it is important SES agencies have an indicative standard for volunteers. The 40 second limit has been determined from previous research amongst police specialist units undertaking a similar assessment, inclusive of a 25 m power swim (Savage & Silk, 2019). As per the HIKE L2 and HIKE L3, it is strongly recommended that agencies capture time-to-completion measures for the POWER SWIM & RESCUE L1 power swim component.

HIKE L2: Completion time = 55 minutes. Recommended minimum time = 45 minutes.
HIKE L3: Completion time = 80 minutes. Recommended minimum time = 70 minutes.
POWER SWIM & RESCUE L1: Completion time of power swim = 40 seconds.

WAS THERE SUFFICIENT DATA TO GUIDE ASSESSORS FOR EXPECTED COMPLETION TIMES FOR EACH ASSESSMENT?

The Validation Socialisation stage saw assessors record times across the following assessments: LIFT & SHIFT L1, LIFT & SHIFT L2, CARRY L1, CARRY L2, LADDER CLIMB & LIFT, HAUL L1, HAUL L2, DRAG L1, DRAG L2. Participant numbers ranged between n=5 and n=66 across these assessments, meaning there are various levels of confidence in the expected times for the assessments. Nevertheless, there is preliminary data that can help guide assessors regarding expected times for these assessments. This data can be used in the planning of assessments; primarily how much time may be needed to run through different sized groups of volunteers (note: the expected completion times do not represent performance standards.) A full list of these completion times is provided in the instruction manual, separate to this report (please see section *1.4 Supporting Documentation*).

4.5 Summary and next steps

The Validation Socialisation research stage has partially achieved what it was designed to achieve. In lieu of researchers being able to run validation trials, the Validation Socialisation trials provided useful data in the context of the project. Primarily, the utility in this phase was that volunteers across five different states were able to experience the majority of the newly developed Fit for Task assessments. Their feedback, in addition to the assessor's feedback, serves as a useful tool for implementation efforts. Their high ratings of fidelity also suggest the assessments that were completed were functional, applicable, and well received.

Despite a range of useful data from the Validation Socialisation stage, there is insufficient data to have complete confidence in the SES Fit for Task assessments which have time-to-completion performance standards (HIKE L2, HIKE L3, POWER SWIM & RESCUE L1). This has implications for the legal defensibility of these assessments; the capacity to withstand legal challenge if a member were to formally challenge the performance standards.

It is important to recognise that for assessments with time-to-completion performance standards, legal defensibility applies to all components of the assessment: the movement patterns, equipment required, pacing, practicality AND performance standards. Though the insufficient data during Validation Socialisation cannot ensure legal defensibility of the performance standards, there is no compromise to the legal defensibility of the remaining components of the assessments. In other words, from all the research preceding the final Validation Socialisation phase, the nature and composition of the assessments remain legally defensible. This is because validation steps have been taken through all stages up until Validation Socialisation.

All nine assessments that comprise the Fit for Task Programme, are still strongly recommended as bona fide minimal physical standards for SES personnel. The endorsement of these assessments has not changed following incomplete validation testing. However, it is strongly encouraged that SES agencies run further validation trials, effectively ‘practice trials’ during a soft implementation phase.

The purpose of running additional trials during a soft implementation would be twofold: (1) agencies to iron out any logistics issues (e.g., what equipment may need to be substituted - such as cribbing in the LIFT AND SHIFT) and (2) collecting more time-to-completion data for HIKE L2, HIKE L3 and especially POWER SWIM & RESCUE L1 (since no data has been collected). Additional data in these assessments can provide more confidence in any final performance standards that are set. As it currently stands, the performance standards recommended in this report may remain unchanged. More data may simply consolidate the confidence in these times.

It is the firm position of the research team that the assessments, in their current format, represent the most appropriate screening tools for the physical competency of SES volunteers. Since a high level of scrutiny has been applied to the assessments throughout the entire research process, agencies should be confident in a product that is representative of both scientific and SES input.

5 Final assessments

As a result of the preceding research stages, nine new SES Fit for Task assessments have been developed. Each assessment represents an evidence-based screening measure modelled on the functional demands of SES roles. Sections 5.2 to 5.10 provide a brief description of each of the nine assessments, and the associated performance levels. For a full description of each assessment, including who should complete the assessment, what equipment is required, how to set up the assessment, and safety considerations, please see the individual instructional documents (see Table 2).

All SES Fit for Task assessments are sex and age neutral. That is, the same standard applies regardless of sex or age. As described in section 1.3.1 *Assessment Standards, Sex and Age*, assessment standards should reflect the *minimum* requirements of performing a job safely and efficiently (Adams, 2016). Scaling standards to accommodate for sex or age is considered discriminatory unless it is demonstrated there are clear differences in the demands of job tasks. The SES Fit for Task research process identified no differences in the requirements between personnel of different sex or age; therefore, the same assessment standard applies to all members.

5.1 Implementation

Upon the successful development of nine new SES Fit for Task assessments, the delivery and implementation of the assessments is a necessary final step to ensure successful uptake of the assessments. Considerations around implementation have been integrated throughout each stage of the research process. Conversations and workshops between the research team and SES SMEs and volunteers, in addition to the Project team, have helped provide a framework for the practical considerations in implementing the programme. A full summary of these considerations is provided in the 'Implementation Strategy' (see Table 2). Included in this report is the Assessment Matrix, which forms a significant part of the implementation process. The Assessment Matrix illustrates the skill sets that are linked to each of the nine assessments (please see following page).

The SES Fit for Task Assessment Matrix illustrates the range of physical assessments matched to SES skill sets. When qualifications are held in two or more skill sets, volunteers need only complete an individual assessment type once (at the highest applicable performance level); there is no requirement to repeat the same assessment type multiple times.

LEGEND FOR TABLE

- Numbers after the Land Search skill set indicates the increasing level of competency or qualification.
- The increasing performance levels within an assessment type are indicated using L1, L2, and L3.

Note: If you use different terminology for skill sets (e.g., "Flood Operations" instead of "Boat Operations") assume they represent the same thing.

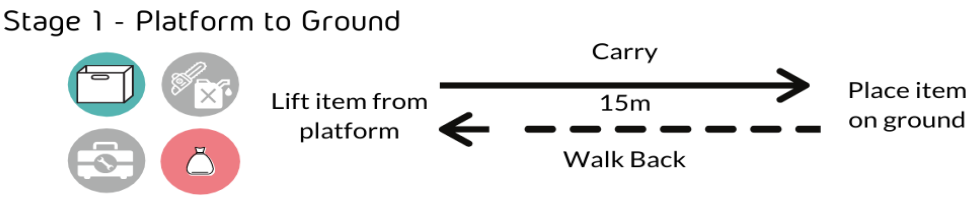
SES SKILL SET	ASSESSMENT TYPE and PERFORMANCE LEVEL								
	LIFT & SHIFT	CARRY	HIKE	DRAG	HAUL	HOLD	LADDER CLIMB & LIFT	IN-WATER SAFETY	POWER SWIM & RESCUE
General Rescue	L1	L1					L1		
Air Search									
Boat Operations	L2	L1			L1			L2 / L3	
Chainsaw Operations	L1			L1		L1	L1		
Firefighting Air Base Support	L1			L2					
In-water Technician	L2	L2	L2		L2			L3	L1
Land-based Swiftwater Rescue	L2	L1			L1			L1	
Land Search 1 (urban or light rural)	L1	L1	L1						
Land Search 2 (rural or rugged)	L1	L2	L2						
Land Search 3 (hard rural - including alpine)_	L2	L2	L3						
Off Road Driving	L1								
Road Crash Rescue	L2	L1		L2		L2			
Storm Damage	L2				L1		L1		
USAR	L1	L2	L1						
Vertical Rescue	L2	L2	L2		L2				

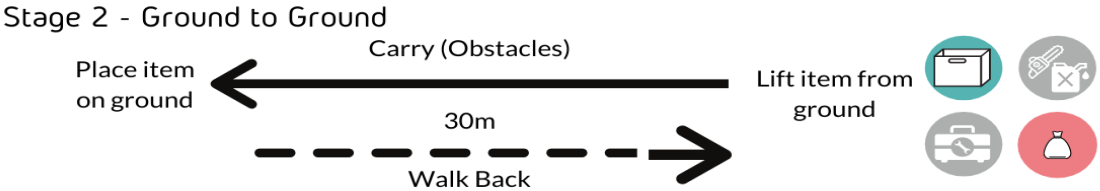
5.2 Lift and shift

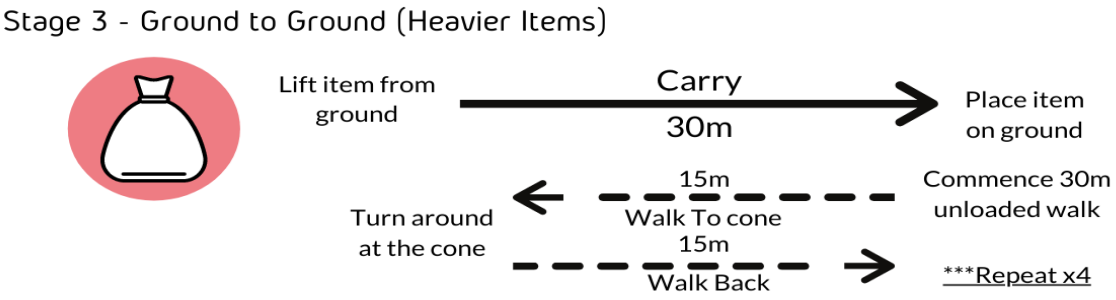
Volunteers are to lift, carry and place a variety of items, while demonstrating safe technique and movement. Items are required to be picked up from the ground or an elevated platform and carried for a distance before being placed back down on the ground or onto an elevated platform. Individuals will then walk back to collect the next item. A portion of the assessment course (Stage 2) will include low obstacles to be negotiated.

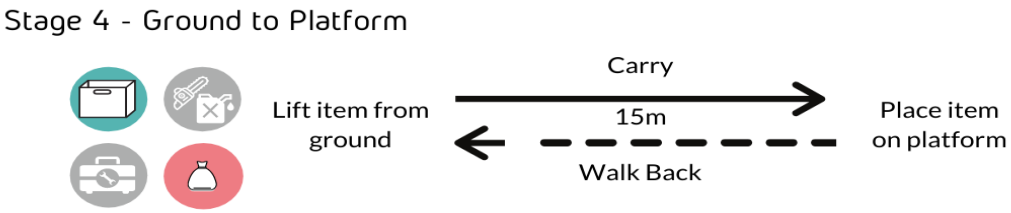
The carry techniques will include:

- ITEM 1: Two-handed carry in front of body
- ITEM 2: Two-handed carry, one item in each hand at side of body
- ITEM 3: Two-handed carry in front of body
- ITEM 4: One-handed carry by side of body
- ITEM 5: However the volunteer chooses









PERFORMANCE LEVELS

LEVEL 1: Complete Stages 1, 2 and 4 (Stage 2 includes three obstacles to negotiate during the carry).

LEVEL 2: Complete all 4 stages in numerical order.

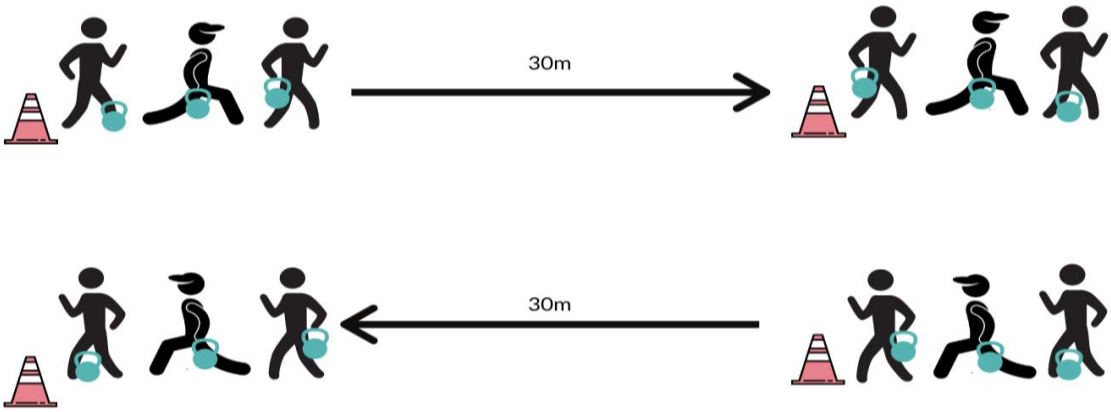
SAFETY & PERFORMANCE

It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job. No running.
- When lifting and lowering items, volunteers should demonstrate a safe lifting technique. If a volunteer cannot lift or lower an item safely, they should be advised by the instructor to ensure their next lift or lower is performed safely. If unsafe lifting or lowering technique continues, the assessment should stop. If items are dropped, the assessment should stop.

5.3 Carry

Members are to lift and carry a 24-kg kettlebell for the prescribed distance, in 30-metre segments. First, members will safely lift the kettlebell from the ground in one hand and carry it for 30 metres, then safely lower it to the ground. Members will then turn 180 degrees and pick up the kettlebell with the other hand, before carrying it back to where they started from. During each 30-metre carry segment, members are permitted one opportunity to pause and readjust their grip.



PERFORMANCE LEVELS

LEVEL 1: a total of 60 m is to be completed.

LEVEL 2: a total of 120 m is to be completed.

SAFETY & PERFORMANCE

It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

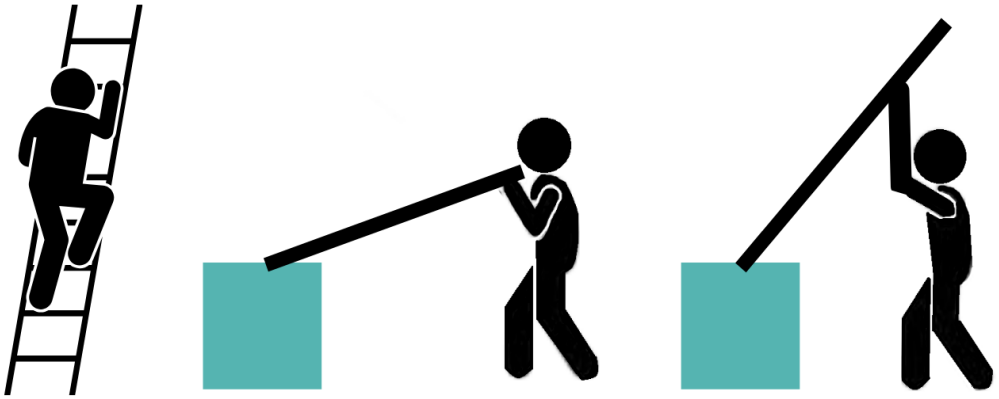
- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job. No running.
- When lifting and lowering, volunteers should demonstrate safe lifting technique. If a volunteer cannot lift or lower the kettlebell safely, they should be advised by the instructor to ensure their next lift or lower is performed safely. If unsafe lifting or lowering technique continues, the assessment must stop. If the item is dropped, the assessment must stop.

5.4 Ladder climb and lift

NOTE: Only members required to climb ladders need to complete this test

PART A: ascend (and then subsequently descend) a ladder, one rung at a time, for a total of 6-8 rungs (representing the height required to gain visual on a single-storey roof / gutter system) whilst maintaining three points of contact. The ladder is to be securely positioned against a wall or building.

PART B: with the far end of the ladder resting on a stable platform (i.e., ready to stow), lift the near end from the ground and press it to an overhead position with elbows locked out.



PART A PART B

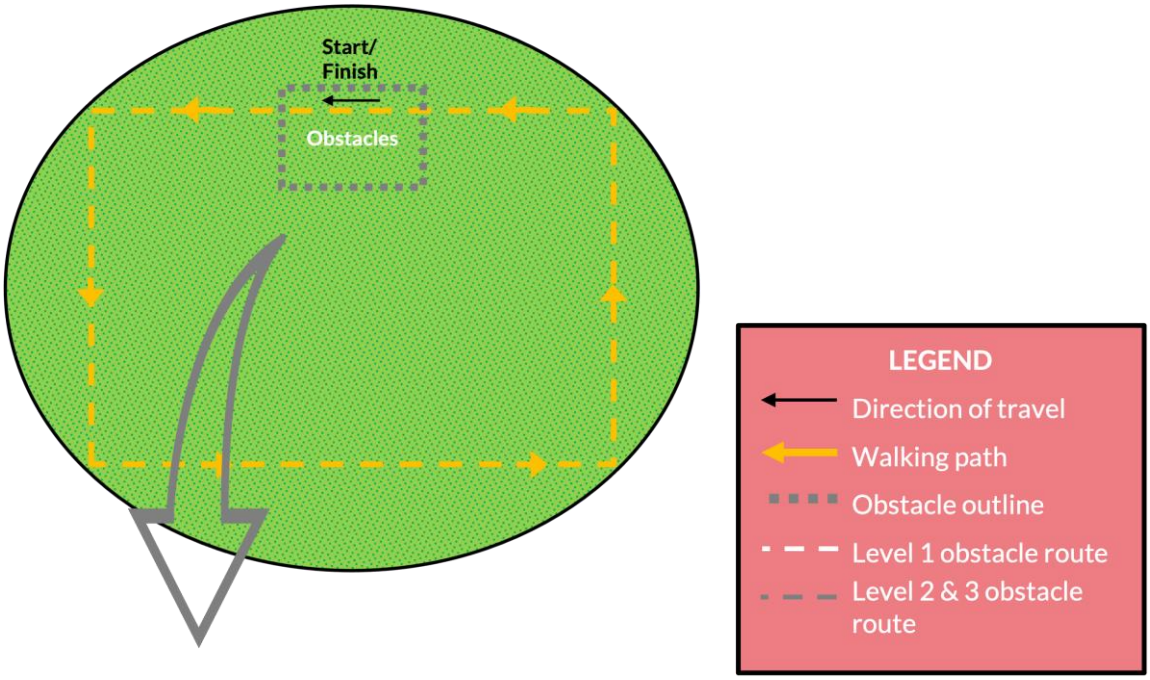
SAFETY & PERFORMANCE

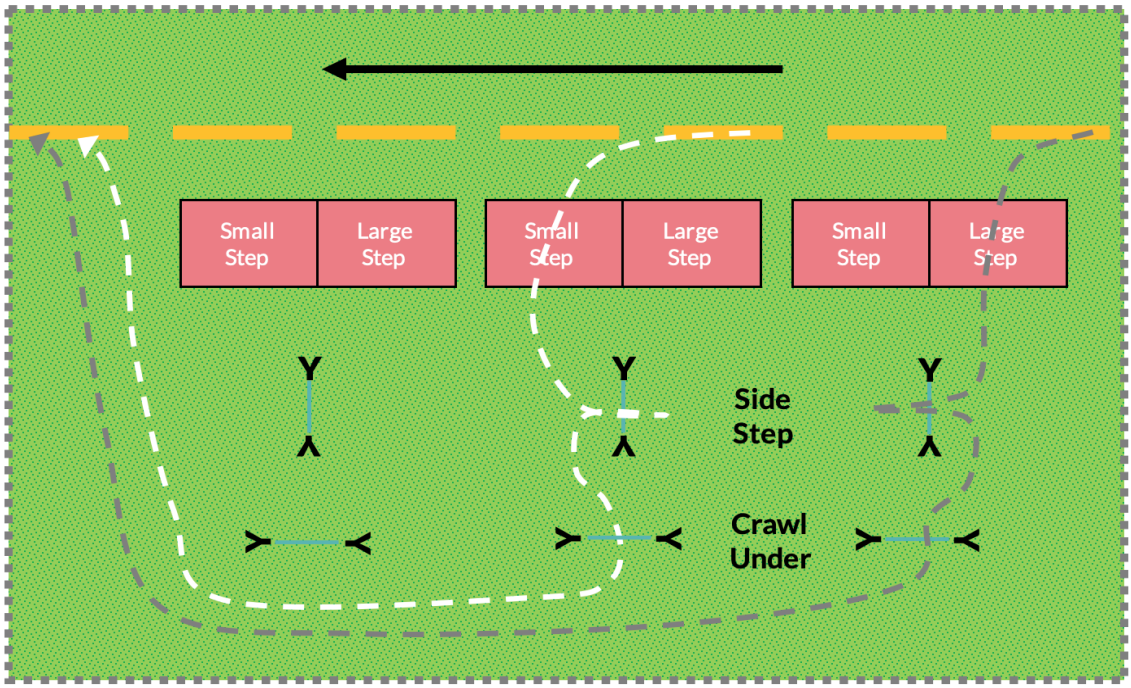
It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- All volunteers undertaking this assessment will be required to have undertaken instruction on the SES requirements for climbing ladders. Anyone climbing a ladder will be required to comply with relevant SES SOPs and Work Instructions (this may include wearing a harness). Safety instructors are to monitor the member at all times whilst they are on the ladder.
- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job.

5.5 Hike

Members are to walk, wearing a backpack of appropriate weight, for a specified distance. The assessment will include some additional physical activities (obstacles) to perform along the way.





PERFORMANCE LEVELS

LEVEL 1: 2 km / 5 LAPS / no completion time enforced; Backpack = skill set specific weight; Obstacles x 3 repetitions: small step ups x 10; sidestep x 2 (i.e., one in each direction) + crawl under.

LEVEL 2: 4 km / 10 LAPS / completion window 45 - 55 minutes; Backpack = skill set specific weight; Obstacles x 8 repetitions: large step ups x 10; sidestep x 2 (i.e., one in each direction) + crawl under.

LEVEL 3: 6 km / 15 LAPS / completion window 70 - 80 minutes; Backpack = skill set specific weight; Obstacles x 13 repetitions: large step ups x 20; sidestep x 2 (i.e., one in each direction) + crawl under.

SAFETY & PERFORMANCE

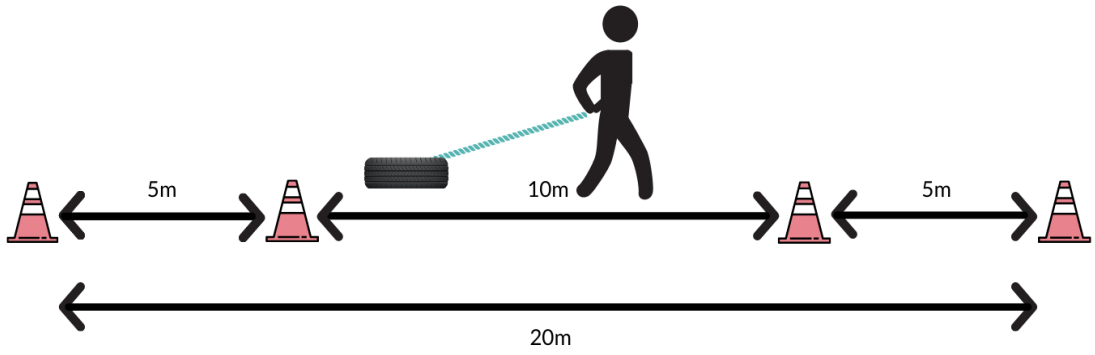
It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- The assessment should be performed at a pace that results in completion within the specified time window for Levels 2 and 3, and at what is considered 'operational tempo' for those performing Level 1. No running.
- Volunteers should not voluntarily pause and remain stationary for more than 1 minute at any stage of the hike.
- If volunteers cannot maintain balance during any of the obstacles, instructors should advise them to slow down if possible, and reattempt the obstacle. If balance still cannot be maintained the instructor should advise the volunteer to stop the assessment.
- If any of the obstacles cannot be completed, or the volunteer cannot finish the total distance in the required time (Level 2 and Level 3 only), the assessment will be considered not to meet the required performance standard.

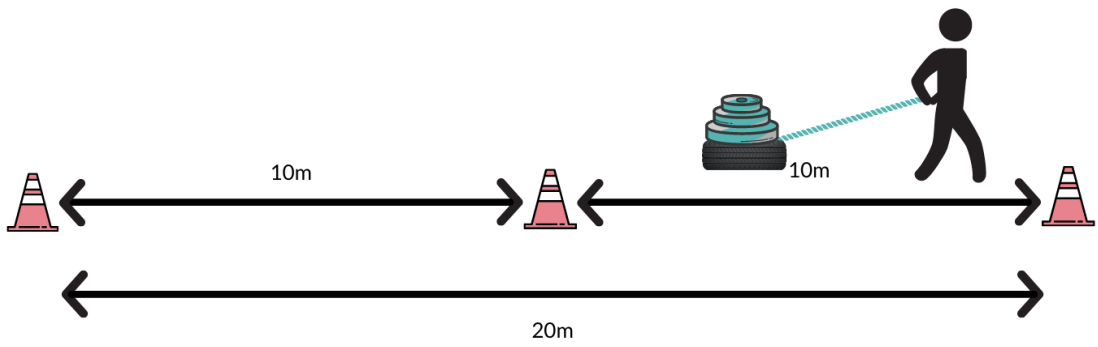
5.6 Drag

Using a forward walking motion (i.e., NOT backwards), using both hands, drag a tyre+rim by the handle and rope attachment for the prescribed distance for the required number of repetitions. (Note: in the below diagram the 'drop zone' cones have been collapsed for simplicity).

LEVEL 1



LEVEL 2



PERFORMANCE LEVELS

LEVEL 1: 10 x 10 m drag repetitions of a tyre+rim with a 10-m "out and back" unloaded recovery walk between drags.

LEVEL 2: 2 x 20 m drag repetitions of a tyre+rim loaded with 10 kg; brief pause performed at the halfway point of each drag before continuing to the end; 20-m "halfway and back" unloaded recovery walk between drags.

SAFETY & PERFORMANCE

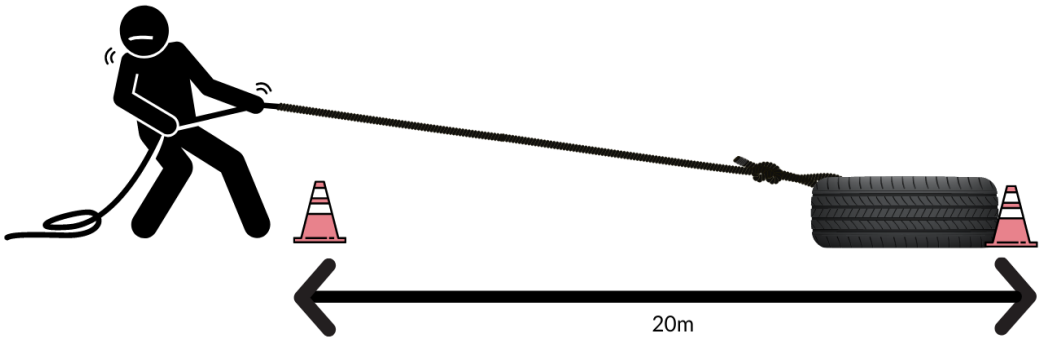
It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job. No running.
- The volunteer should always walk in a forwards motion. No backwards walking.
- There should be no prolonged pausing (greater than 10 seconds) during each drag.

5.7 Haul

Employing a hand-over-hand technique, members are required to haul a rope attached to a tyre + rim. The haul is to be performed in a continuous motion over the prescribed distance. The participants will stand with their feet between the pair of cones and haul the tyre from 'drop zone' to 'drop zone'.

(Note: in the below diagram not all cones have been represented when compared to the bird's-eye view diagram).



PERFORMANCE LEVELS

LEVEL 1: 1 x 20-m haul repetition of a tyre + rim.

LEVEL 2: 2 x 20-m haul repetitions of a tyre + rim loaded with 10 kg; taking the rope, walk to the other end to complete the second haul.

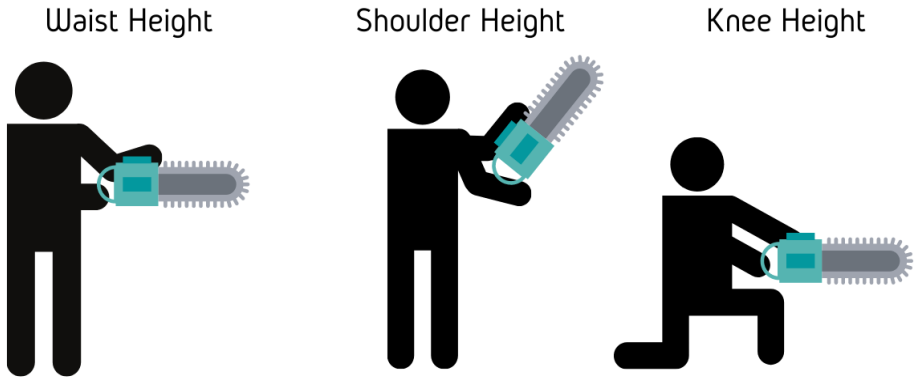
SAFETY & PERFORMANCE

It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job.
- The volunteer should not move their feet or sway their back.
- The tyre should maintain constant forward motion (no pausing once the haul has begun).

5.8 Hold

This assessment requires the execution of an equipment hold in three different positions (waist, then shoulder and then knee), demonstrating strength, balance and control. Members will complete 2 hold repetitions in each position for 5 seconds, with a 10-second rest period in between. During each hold repetition, the position must be maintained and there is to be clear separation between the equipment and the body or clothing. All holds should be performed using two hands.



PERFORMANCE LEVELS

LEVEL 1: performed with a commonly used chainsaw.

LEVEL 2: performed with a commonly used combination rescue tool.

SAFETY & PERFORMANCE

It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- The volunteer should remain still during each hold.
- The volunteer should not arch their back.
- The volunteer should not rest the item against their body during the hold.

5.9 In-water safety

Whilst wearing role-appropriate Personal Protective Equipment & Clothing, including a personal floatation device, volunteers are to enter the water and move a specified distance using a swimming stroke of their choice, then demonstrate the ability to maintain safety in the water. It is desirable to have the volunteer start at the end of the pool (A), then swim to the midpoint of the pool (if a 50-m pool) or to the other end of the pool (if a 25-m pool). However, if the pool has a natural height differential between the water and the edge (30 - 50 cm) at the end of the pool, then the volunteer may start in the middle of the pool (if it is a 50-m pool) and swim to the end. No matter how the assessment is run, volunteers are not permitted to kick off the wall at any stage.

Once the volunteer has completed the swim, they are required to pull themself out of the water and onto the pool deck, without the use of steps or aids (B). Depending on their performance level, they may then be required to safely lift a 40-kg aquatic rescue dummy from the water onto the pool deck (C).

PERFORMANCE LEVELS

LEVEL 1: Complete a 25-metre swim and self-extrication.

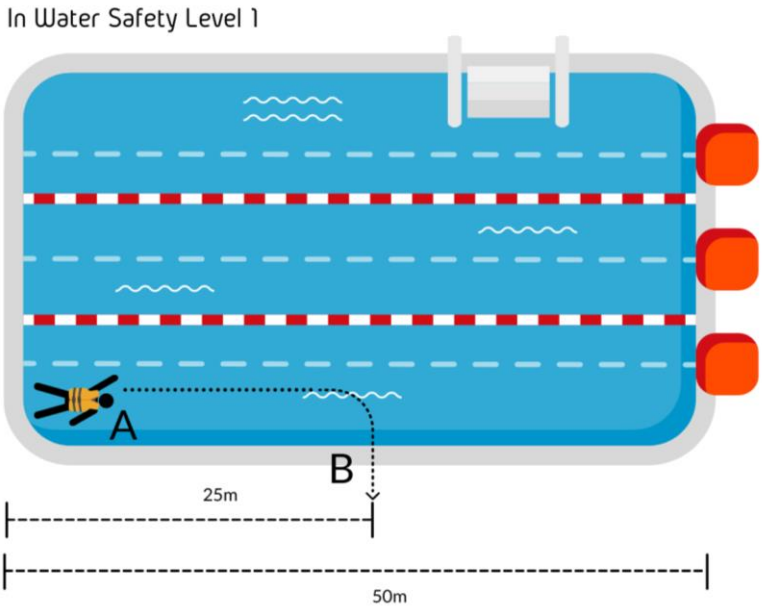
LEVEL 2: Complete a 25-metre swim and self-extrication, plus the safe rescue of a rescue dummy.

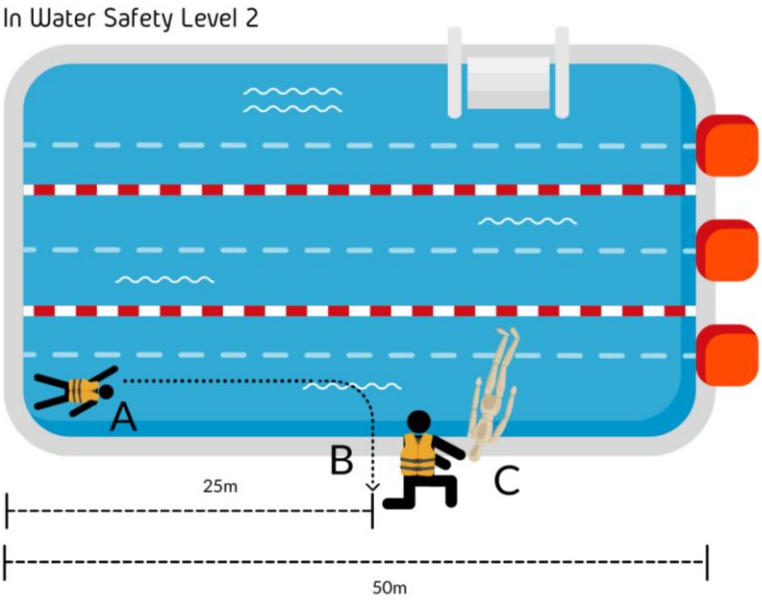
LEVEL 3: Complete a 50-metre swim and self-extrication, plus the safe rescue of a rescue dummy.

SAFETY & PERFORMANCE

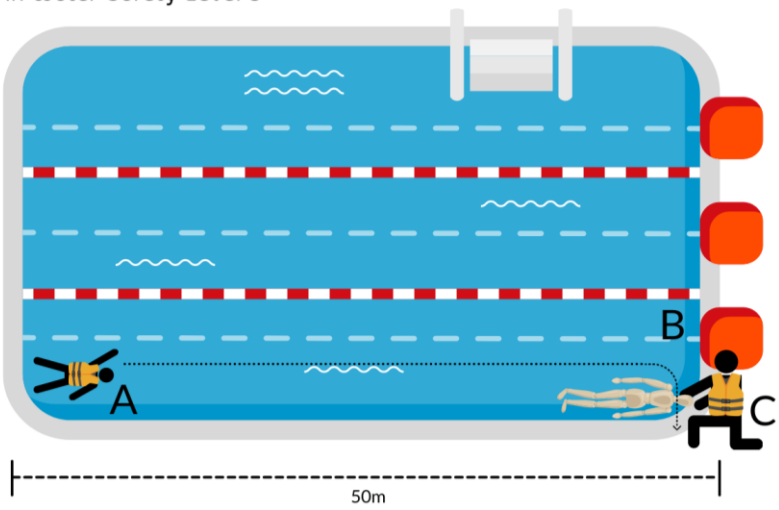
It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- In addition to the instructor, there must be a safety person watching the volunteer in the water at all times.
- The assessment should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job.
- There is no time limit imposed on the assessment, however, volunteers must demonstrate constant progress throughout. Stopping for more than a few seconds is not permitted.
- If a volunteer becomes fatigued or unable to continue, they are permitted to hold onto the lane rope or side of the pool. If this occurs, they must stop the assessment and exit the pool. There is no consequence or penalty for stopping, and volunteers may have another attempt at the assessment.



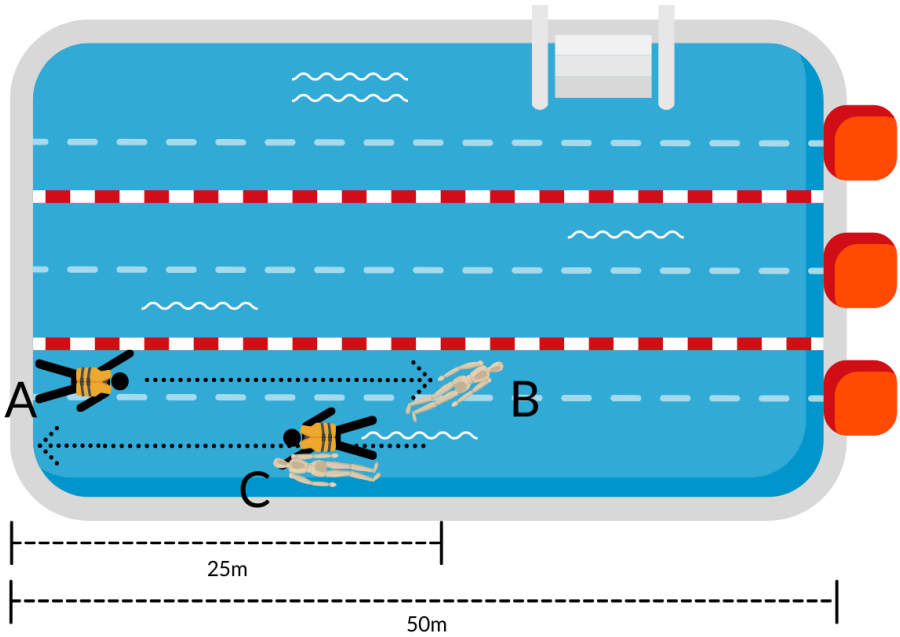


In Water Safety Level 3



5.10 Power swim and rescue

Whilst wearing role-appropriate Personal Protective Equipment & Clothing, including personal floatation device, whilst starting in the water (A), volunteers are to swim 25 metres in no more than 40 seconds, using a swimming stroke of their choice, to reach a rescue dummy (B). Upon making contact with the rescue dummy, volunteers must tread water for 30 seconds whilst they orientate it to a safe and stable posture, ready for it to be towed. After 30 seconds the volunteer must tow it back to where they started (C). There is no time limit imposed on the rescue, however, the volunteer must demonstrate constant forward motion.



SAFETY & PERFORMANCE

It is important that volunteers complete the assessment in a safe manner. Here are some rules around safety and performance:

- In addition to the assessor, there must be a safety person watching the volunteer in the water at all times.

- The power swim is required to be completed in no more than 40 seconds, representing the scenario of quickly closing the distance to someone who is in trouble in the water or rescuing a casualty in fast flowing water.
- The assessor should time the 30 seconds for treading water and instruct the volunteer when they can commence the rescue.
- The rescue should be performed at 'operational tempo'. This means it is completed at the same intensity as if doing a job.
- If a volunteer becomes fatigued or unable to continue, they are permitted to hold onto the lane rope or side of the pool. If this occurs, they should stop the assessment and exit the pool. There is no consequence or penalty for stopping, and volunteers may have another attempt at the assessment.

6 References

Adams, E.M., 2016. Human rights at work: Physical standards for employment and human rights law. *Applied Physiology, Nutrition, and Metabolism*, 41(6):S63-S73.

Australian Human Rights Commission, 2015. *Australian Human Rights Commission: Annual Report 2015-16,* Sydney.

Beck, B., Billing, D.C., Carr, A.J., 2016. Developing physical and physiological employment standards: Translation of job analysis findings to assessments and performance standards – A systematic review. *International Journal of Industrial Ergonomics.* 56:9-16.

Hopkins, W.G., 2000. Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1):1-15.

Hopkins, W., Marshall, S., Batterham, A. and Hanin, J., 2009. Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise*, 41(1):3-12.

Hreljac A., 1993. Preferred and energetically optimal gait transition speeds in human locomotion. *Medicine & Science in Sports & Exercise*, 25:1158–1162.

Jamnik, V., Gumienak, R., Gledhill, N., 2012. Developing legally defensible physiological employment standards for prominent physically demanding public safety occupations: A Canadian perspective. *European Journal of Applied Physiology*, 113(10):2447-2457.

Kenny, G.P., Groeller, H., McGinn, R. and Flouris, A.D., 2016. Age, human performance, and physical employment standards. *Applied Physiology, Nutrition, and Metabolism*. 41(6):S92-S107.

Larsen, B., Aisbett B., 2012. Subjective job task analyses for physically demanding occupations: What is best practice? *Ergonomics,* 55(10):1266-77.

Nunan, D., Donovan, G.A.Y., Jakovljevic, D.G., Hodges, L.D., Sandercock, G.R. and Brodie, D.A., 2009. Validity and reliability of short-term heart-rate variability from the Polar S810. *Medicine & Science in Sports & Exercise*, 41(1):243-250.

Payne, W., Harvey, J., 2010. A framework for the design and development of physical employment tests and standards. *Ergonomics*. 53(7): 858-71.

Petersen, S.R., Anderson, G.S., Tipton, M.J., Docherty, D., Graham, T.E., Sharkey, B.J. and Taylor, N.A., 2016. Towards best practice in physical and physiological employment standards. *Applied Physiology, Nutrition, and Metabolism*, 41(6):S47-S62.

Rayson M. 1998. The development of physical selection procedures. Phase 1: Job analysis. *Contemporary Ergonomics,* 393-397.

Roberts, D., Gebhardt, D.L., Gaskill, S.E., Roy, T.C. and Sharp, M.A., 2016. Current considerations related to physiological differences between the sexes and physical employment standards. *Applied Physiology, Nutrition, and Metabolism*, 41(6):S108-S120.

Savage, R., and Silk, A., 2019. Specialist Response Division: Fit for Duty 2019. *Human Performance Science.*
Shechtman, O., 2013. The coefficient of variation as an index of measurement reliability, in S.A.R. Doi & G.M. Williams (eds), *Methods of Clinical Epidemiology*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 39-49.
Shookster, D., Lindsey, B., Cortes, N. and Martin, J.R., 2020. Accuracy of commonly used age-predicted maximal heart rate equations. *International Journal of Exercise Science*, 13(7):1242-1250.

Svensson, E., 2001. Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine*, 33(1):47-48.

Taylor, N.A.S., Groeller, H., 2003. Work-based physiological assessment of physically-demanding trades: A methodological overview. *Journal of Physiological Anthropology and Applied Human Science,* 22(2):73-81.
Waters, R.L., Lunsford, B.R., Perry, J. and Byrd, R., 1988. Energy-speed relationship of walking: Standard tables. *Journal of Orthopaedic Research*, 6(2):215-222.

Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, 19(1):231-240.

Zumbo, B.D., 2016. Standard-setting methodology: Establishing performance standards and setting cut-scores to assist score interpretation. *Applied Physiology, Nutrition, and Metabolism*, 41(6):S74-S82.

Appendices

Appendix A – Participant information

The following pages illustrate participant demographics throughout each stage of the research process.

A.1 Verifiction workshops

Table 6 Details of subject matter experts (SMEs) who attended workshops. STDEV = standard deviation.

SME DETAILS		NSW & ACT	QLD	VIC	TAS	SA	WA	NT	TOTAL	AVERAGE	STDEV
ATTENDEES	TOTAL	10	15	16	7	7	10	8	73	10.4	3.7
	MALE	7	11	14	6	6	8	8	60	8.6	2.9
	FEMALE	3	4	2	1	1	2	0	13	1.9	1.3
AGE	AVERAGE	49.6	49.4	44.1	60.1	47.0	47.9	41.1		48.5	6
	STDEV	6.7	11.5	13.5	10.1	13.9	8.5	14.2			
	YOUNGEST	38	27	27	46	24	35	27		32	8
	OLDEST	56	69	66	72	62	61	61		63.9	5.5
YEAR'S SES EXPERIENCE	AVERAGE	19.4	12.2	14.2	19.8	19.4	11.2	6.7		14.7	5
	STDEV	12.3	7.7	10.4	11.7	11.5	8.5	3.6		9.4	3.1
	LEAST	5	1	3	5.5	4	2.0	2		3.2	1.7
	MOST	38	33	37	30	35	19.0	12		29.1	9.9
UNITS	UNITS*	7	10	9	1	3	5	1	36	5.1	3.7
	COMMAND^	3	2	4	5	4	2	1	21	3	1.4

* Number of different Units represented by SMEs

^ Number of SMEs in a command role. E.g., headquarters, regional office, district office

A.2 Surveys

Table 7 Number of survey responses for each state and territory.

STATE OR TERRITORY	RESPONSES
TOTAL	872
ACT	69
NSW	194
NT	9
QLD	130
SA	75
TAS	43
VIC	216
WA	136

Table 8 Number of survey responses for each gender.

GENDER	RESPONSES
TOTAL	868
Male	619
Female	246
Transgender	2
Prefer not to say	1

Table 9 Details of personnel who completed the survey. Total entries = 872. STDEV = standard deviation.

	AGE	SES YEARS	QUALIFICATIONS HELD
AVERAGE	49.3	11.3	4.1
STDEV	14.5	10.0	3.5
MIN	16	<1	1
MAX	81	53	12

A.3 Field trials

Table 10 Demographic data for all participants taking place in field trials for current SES Fit for Task Project. Total participants = 184. STDEV = standard deviation.

SEX	UNIT		SES YEARS	AGE	HEIGHT (cm)	WEIGHT (kg)
Male = 132	Number of Units participating = 54	AVERAGE	7.9	43.9	175.5	81.9
Female = 52		STDEV	8.5	14.9	9.5	15.8
		MIN	<1	18	148	46
		MAX	47	85	196	130

Table 11 Demographic data for all participants taking place in field trials for previous SES Fit for Task Project. Total participants = 78. STDEV = standard deviation.

SEX	UNIT		SES YEARS	AGE	HEIGHT (cm)	WEIGHT (kg)
Male = 58	(Data not available)	AVERAGE	7.6	44.8	175.6	84
Female = 28		STDEV	4.17	6.2	3.3	3.2
		MIN	<1	18	151	50
		MAX	43	73	196	127

A.4 Pilot validation

Table 12 Demographic data for all participants taking place during Pilot Validation. Total participants = 9. STDEV = standard deviation.

SEX	UNIT		AGE	HEIGHT (cm)	WEIGHT (kg)
Male = 5	Number of Units participating = 7	AVERAGE	46.1	178.7	89
Female = 4		STDEV	12.6	9.0	24.3
		MIN	26	163	53
		MAX	57	1187	118

A.5 Validation socialisation

Table 13 Demographic information relating to respondents to the individual survey during the Validation Socialisation phase.

AGE (years)		GENDER		LOCATION	
25 and under	4	Male	29	WA	3
26 - 35	13	Female	13	QLD	9
36 - 45	3	Non-binary or gender diverse	1	VIC	11
46 - 55	12			NSW	20
56 - 65	9			SA	0
Over 65	2				
YEARS IN SES		RECENT EXPERIENCE			
Less than 1 year	10	Within 1 month	33		
1 - 2 years	7	Within 3 months	4		
3 - 5 years	15	Within 6 months	1		
6 - 10 years	6	Within 12 months	4		
More than 10 years	5	Within 2 years	1		

Appendix B – Survey responses

The following table outlines the number of responses for each skill set as part of the Surveys stage of the research.

Table 14 Number of survey responses for each Skill Set. Note: the sum of individual Skill Set responses does not equal the total responses, since personnel often held qualifications for two or more Skill Sets.

SKILL SET	RESPONSES
TOTAL	872
GENERAL RESCUE	756
STORM DAMAGE	777
CHAINSAW OPERATIONS	660
LAND SEARCH & RESCUE	750
AIR SEARCH	122
VERTICAL RESCUE	194
ROAD CRASH RESCUE	277
USAR	373
BOAT OPERATIONS*	291
OFF-ROAD DRIVING	456
FIREFIGHTING AIR-BASE SUPPORT	77

*Boat Operations later split to include 'IN-WATER TECHNICIAN' and 'LAND-BASED SWIFTWATER'.

Appendix C – Data processing and detailed results

The following link will direct you to a spreadsheet with the detailed results from each of the stages during the job task analysis. Please read the 'Start Here' tab before viewing the subsequent tabs.

https://www.naturalhazards.com.au/sites/default/files/2023-06/SES%20task%20lists%20to%20criterion%20tasks%20-%20Roadmap.xlsx

If you cannot access the link, or you are reading this in printed form, the spreadsheet is provided as a separate document to this report (please see Table 2).