

Article

Inter-Rater and Intra-Rater Reliability of Return-to-Work Screening Tests for UK Firefighters Following Injury

Liam Noll *, Jason Moran and Adrian Mallows

School of Sport, Rehabilitation & Exercise Sciences, University of Essex, Colchester CO4 3SQ, UK

* Correspondence: lnoll@essex.ac.uk

Abstract: The aim of this study was to assess the inter-rater and intra-rater reliability of a return-to-work (RTW) screening test to be used on UK firefighters following injury. The inter rater and intra-rater reliability of eight tasks involved in a screening test was used to assess readiness to RTW for UK firefighters following injury. These tasks included the following: (1) putting on and removing a breathing apparatus set (BA), (2) a ladder lift simulation, (3) a ladder carry simulation, (4) a light portable pump (LPP) lift and carry simulation, (5) a hose run, (6) a ladder climb with leg lock, (7) a casualty evacuation and (8) a confined space crawl simulation. The inter-rater reliability between each individual screening task was interpreted as Excellent (ICC = 0.94–1.00) for eleven (68.75%) of the screening task videos and as Good (ICC = 0.75–0.88) for five (31.25%) of the screening task videos. Intra-rater reliability was interpreted as Excellent (ICC = 1) for twenty-six participants (74.3%), Good (ICC = 0.76–0.88) for eight participants (22.9%) and Moderate for one participant (2.8%). Due to the reliability of this screening test, it allows conclusions to be made from the results which can inform a RTW decision for a firefighter.

Keywords: return to work; firefighter; tactical athlete; functional capacity; physical assessment



Citation: Noll, L.; Moran, J.; Mallows, A. Inter-Rater and Intra-Rater Reliability of Return-to-Work Screening Tests for UK Firefighters Following Injury. *Healthcare* **2022**, *10*, 2381. <https://doi.org/10.3390/healthcare10122381>

Academic Editor: Alberto Modenese

Received: 7 November 2022

Accepted: 25 November 2022

Published: 27 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Musculoskeletal injuries can account for one-third of all workplace-related injuries [1,2]. Common causes include overexertion, contact with equipment, slips, trip and falls [2]. Many work tasks contain some risk of injury; however, the extent of these risks differs depending on the type of sector and job role [3]. The risk of a work-related injury increases for individuals with athletic occupations, including firefighters, military personnel, police officers and paramedics, whose job role requires higher physical demands; for example, heavy lifting, kneeling and crouching [4–7]. Of these injuries, more than 40% were musculoskeletal-related [4,8].

Following a musculoskeletal-related injury, assessing an individual's readiness to return to work (RTW) can be complex; many factors need to be considered, including physical performance in relation to the work task demands [9,10]. An individual may believe that they are ready to RTW, but if they are unable to meet the minimum work-related physical demands, an increase to reinjury has been shown [11,12].

To assess physical performance in relation to work task demands, during recruitment of athletic occupations, a physical screening test is used to determine if individuals possess the minimum required aerobic fitness and muscular strength standards [13–15]. However, no such test exists to determine if an individual can meet the minimum standards after injury. For example, the physical screening test used for recruitment of firefighters does not include all tasks involved during operational duties, including hose running and ladder carry [1,16]. Instead, UK firefighter selection tests were designed to help identify applicants physically suited to roles within UK fire and rescue services [17] and then once employed, individuals are trained in more specific tasks related to their firefighting role [18]. If operational tasks are unable to be completed effectively in emergency situations, a firefighter could put themselves at risk of danger, their operational colleagues and members of the public [19].

To date, limited research exists for the effectiveness of RTW screening tests to reduce reinjury rates for individuals returning to work in an athletic occupation, for example a professional athlete [20–24]. No research has included athletic populations who are not professional athletes, for example firefighters [20–24]. To start to address this, a recent study provided consensus for the inclusion of tasks to be adopted into a screening test that could be used to assess a firefighter's readiness to RTW following injury [1].

However, before any screening test can be used to assess readiness to RTW, its reliability must be determined [25]. The reliability of a screening test should be of important consideration especially in settings where decisions on an individual's ability to perform job related tasks at the required level as based on interpretation of the results [26]. A reliable screening test ensures the same or compatible results across different assessments, regardless of when the test took place, the environment in which the test is conducted in, or the professional administering the test [25,27]. Without sufficient inter-rater and intra-rater reliability, any screening test holds little value in determining if an individual is ready to return to the demands of their job role [28].

The aim of this study was to assess the inter-rater and intra-rater reliability of a RTW screening test to be used on UK firefighters following injury.

2. Materials and Methods

2.1. Study Design

An inter-rater and intra-rater reliability study of eight tasks involved in a screening test was used to assess readiness to RTW for UK firefighters following injury. The eight tasks in the screening test were gained by consensus during a recent Delphi study [1] and include the following; (1) putting on and removing a breathing apparatus set (BA), (2) a ladder lift simulation, (3) a ladder carry simulation, (4) a light portable pump (LPP) lift and carry simulation, (5) a hose run, (6) a ladder climb with leg lock, (7) a casualty evacuation and (8) a confined space crawl simulation.

2.2. Participant Criteria

A purposive sample, of occupational health, fitness professionals or operational firefighters working within fire services in the United Kingdom (UK) was recruited to be participants. Purposive sampling aimed to capture experts within the fire service. All participants were currently involved in health and fitness assessments of operational firefighters. There was no requisite on the number of years a participant had worked within their role.

2.3. Recruitment

Participants were recruited from the National Fire Chiefs Council Fitness Advisers and Occupational Health online groups. The researcher (LN) emailed fitness advisors, occupational health managers, occupational health nurses, occupational health advisors and operational firefighter trainers who currently work for UK fire and rescue services, inviting them to participate in the study. The email included a hyperlink to the study website page and a participant information sheet (PIS). All participants were required to give their consent by answering the pre-study questions before progressing further in the study.

2.4. Sample Size

A priori power analysis was conducted to estimate the sample size required using G* Power software (version 3.1.9.4), Franz Faul, Germany [29]. The results estimated that a sample size of thirty-five would be required to establish inter-rater and intra-rater reliability ($H_0 = 0.00$, $H_1 = 0.70$, $\alpha = 0.05$, single tail, power = 0.95) [30]. To allow for attrition, we increased this estimated sample size by 10% and rounded up to the nearest whole number [31,32], leaving a sample size of thirty-nine.

2.5. Data Collection/Testing Procedure

Participants were provided access to a website, created using the E-learning tool Moodle [33]. The website hosted videos of the screening tests were recorded in 1080p HD video at 60 frames per second using an iPhone 12 and were edited in iMovie [34]. The iPhone 12 was set up on a tripod at approximately two meters [35] from the individual being recorded, from a front view. Each screening test was recorded two times with predetermined outcomes, (1. Pass, 2. Fail). All participants were unaware of the predetermined outcome for each video. The scoring criteria were based on the current national firefighter guidance for correct technique required for the tests [17].

All participants were required to watch an online training video detailing the online screening criteria form (SCF) before completing any rating as part of this study. The online training video was created by one of the researchers (LN) by screen recording of a mock screening test rating using Microsoft Teams [36]. The mock screening test was different from the included screening tests to avoid any influence on participants rating. After viewing the online training video, all participants were required to complete a multiple-choice questionnaire based on the training video with 100% pass mark required to pass the training. If any participants had difficulties with the online training, they were able to contact one of the researchers (LN) via email for assistance. To ensure audio and video quality, a pilot test was undertaken by one of the researchers (LN).

Participants visually assessed the technique used in the video for each screening test using a score criteria ("Pass" or "Fail"). Scores were based on a participant's judgment regarding technique throughout the task using the scoring criteria provided for each task as a reference (Appendix A Table A1).

For each participant, two rating sessions were performed with two weeks separating each session as used in previous reliability studies [25,26]. The measures obtained from both rating sessions were used to estimate inter-rater reliability. The initial and follow up testing measures from participants were used to estimate intra-rater reliability. All participants were blinded to other participants' scores by viewing the videos of the screening test online individually. All participants were advised to prevent any communication about the screening videos and/or ratings between each other. All videos were required to be rated in one sitting.

2.6. Statistical Analysis

Descriptive data were used to characterise the participants using means with standard deviations (SD) where applicable using a Microsoft Excel spreadsheet. Scores from the participants were initially stored in a Microsoft Excel spreadsheet.

Inter-rater and intra-rater reliability was assessed using Intra-class Correlation Coefficients (ICC) [25]. For inter-rater reliability, a two-way random-effects model, mean of k raters, and absolute agreement (ICC(2, k)) was used. For intra-rater reliability, a two-way mixed-effects model, mean of k measurements, and absolute agreement (ICC(3,1)) was used. Interpretation of reliability results was based on the following criterion: Excellent reliability (>0.90), Good reliability (0.75–0.90), Moderate reliability (0.50–0.75) and Poor reliability (<0.50) [37]. All statistical analysis were conducted using Statistical Package for the Social Services (SPSS) version 27 for Windows [38].

3. Results

3.1. Participants

Forty-two participants volunteered to participate in this study. Participants' job roles within their service included fitness advisors ($n = 14$) (40%), occupational health doctor ($n = 1$) (2.8%), occupational health manager ($n = 1$) (2.8%), occupational health nurse ($n = 1$) (2.8%), occupational health advisor ($n = 7$) (20%) and operational firefighter trainer ($n = 11$) (31.4%) (Figure 1). From these, a total of thirty-five participants completed both online rating screening sessions (83.3% retention rate). There was representation from different fire and rescue services across the UK ($n = 8$) (Figure 2). Overall, the demographic of the

participants was proportionally representative of the original invitation list. The mean age of the participants in this study was 40.34 + 9.02 years and the mean duration they had worked for their fire service was 12.40 + 8.11 years (Table 1).



Figure 1. Bar Chart showing the job role of the participants.

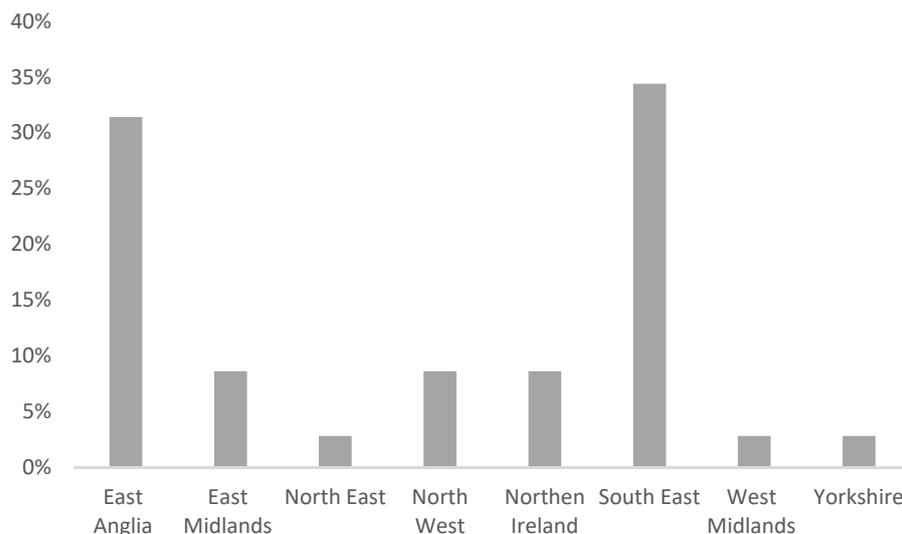


Figure 2. Bar chart showing the region representation in the United Kingdom of the participants.

Table 1. Overall inter-rater reliability of all screening tests for both rating sessions. ICC = Intraclass correlation coefficients, CI = Confidence interval.

Rating Session	Inter-Rater Reliability		
	ICC _{2,35}	95% CI	Interpretation
1	0.77	0.67–0.85	Good
2	0.79	0.71–0.87	Good

3.2. Inter-Rater Reliability between All Screening Tasks

The inter-rater reliability between all screening tasks during both rating sessions was interpreted as Good (ICC = 0.77–0.79) (Table 1). For participants with 0–9 years of service, the inter-rater reliability between all screening tasks during both rating sessions was interpreted as Good (ICC = 0.76–0.81) and Good (ICC = 0.77–0.82) for participants with more than nine years of service (Table 2).

Table 2. Inter-rater reliability of all screening tests for both rating sessions based on years worked with the fire service. ICC = Intraclass correlation coefficients, CI = Confidence interval.

	Rating Session	Inter-Rater Reliability		
		ICC _{2,35}	95% CI	Interpretation
0–9 years of service	1	0.76	0.66–0.85	Good
	2	0.81	0.72–0.89	Good
9+ years of service	1	0.77	0.68–0.86	Good
	2	0.82	0.75–0.89	Good

3.3. Inter-Rater Reliability between Each Individual Screening Task

The inter-rater reliability between each individual screening task was interpreted as Excellent (ICC = 0.94–1.00) for eleven (68.75%) of the screening task videos across both rating sessions. These tasks included, Ladder lift (Pass Video), Putting on a BA set (Fail Video), Ladder carry (Pass and Fail video), LPP lift and carry (Pass and Fail video), Hose run (Pass and Fail video), Casualty evacuation (Pass and Fail video) and Confined Space (Fail video) (Table 3).

Table 3. Inter-rater reliability of each individual screening test video over two rating sessions. ICC = Intraclass correlation coefficients, CI = Confidence interval.

	Rating Session	Inter-Rater Reliability		
		ICC _{2,35}	95% CI	Interpretation
Ladder lift (pass)	1	1.00	1.00–1.00	Excellent
	2	1.00	1.00–1.00	Excellent
Ladder lift (fail)	1	0.88	0.55–1.00	Good
	2	0.88	0.55–1.00	Good
Putting on a BA set (pass)	1	0.76	0.32–1.00	Good
	2	0.76	0.33–1.00	Good
Putting on a BA set (fail)	1	0.94	0.74–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
Ladder carry (pass)	1	0.94	0.74–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
Ladder carry (fail)	1	0.94	0.74–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
LPP lift and carry (pass)	1	1.00	1.00–1.00	Excellent
	2	1.00	1.00–1.00	Excellent
LPP lift and carry (fail)	1	1.00	1.00–1.00	Excellent
	2	1.00	1.00–1.00	Excellent
Hose run (pass)	1	0.94	0.74–1.00	Excellent
	2	1.00	1.00–1.00	Excellent
Hose run (fail)	1	0.91	0.70–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
Ladder climb and leg lock (pass)	1	0.82	0.43–1.00	Good
	2	0.82	0.43–1.00	Good
Ladder climb and leg lock (fail)	1	0.81	0.42–1.00	Good
	2	0.88	0.55–1.00	Good
Casualty evacuation (pass)	1	0.94	0.74–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
Casualty evacuation (fail)	1	0.94	0.74–1.00	Excellent
	2	0.94	0.74–1.00	Excellent
Confined space (pass)	1	0.81	0.42–1.00	Good
	2	0.81	0.42–1.00	Good
Confined space (fail)	1	1.00	1.00–1.00	Excellent
	2	1.00	1.00–1.00	Excellent

Inter-rater reliability was interpreted as Good (ICC = 0.75–0.88) for five (31.25%) of the screening task videos across both rating sessions. These tasks included, Ladder lift (Fail video), Putting on a BA Set (Pass video), Ladder climb and leg lock (Pass and Fail video), Confined space (Pass video) (Table 3).

3.4. Intra-Rater Reliability

Intra-rater reliability was interpreted as Excellent (ICC = 1) for twenty-six participants (74.3%), Good (ICC = 0.76–0.88) for eight participants (22.9%) and Moderate for one participant (2.8%) (Table 4).

Table 4. Intra-rater reliability between each rating session. ICC = Intraclass correlation coefficients, CI = Confidence interval.

Intra-Rater Reliability			
	ICC _(3,1)	95% CI	Interpretation
Participant 1	1.00	1.00–1.00	Excellent
Participant 2	1.00	1.00–1.00	Excellent
Participant 3	0.63	0.37–0.80	Moderate
Participant 4	1.00	1.00–1.00	Excellent
Participant 5	1.00	1.00–1.00	Excellent
Participant 6	0.80	0.77–0.94	Good
Participant 7	0.81	0.66–0.91	Good
Participant 8	1.00	1.00–1.00	Excellent
Participant 9	0.76	0.56–0.87	Good
Participant 10	1.00	1.00–1.00	Excellent
Participant 11	1.00	1.00–1.00	Excellent
Participant 12	1.00	1.00–1.00	Excellent
Participant 13	0.76	0.56–0.87	Good
Participant 14	1.00	1.00–1.00	Excellent
Participant 15	0.88	0.77–0.94	Good
Participant 16	1.00	1.00–1.00	Excellent
Participant 17	1.00	1.00–1.00	Excellent
Participant 18	1.00	1.00–1.00	Excellent
Participant 19	0.88	0.77–0.94	Good
Participant 20	1.00	1.00–1.00	Excellent
Participant 21	1.00	1.00–1.00	Excellent
Participant 22	1.00	1.00–1.00	Excellent
Participant 23	0.88	0.77–0.94	Good
Participant 24	1.00	1.00–1.00	Excellent
Participant 25	1.00	1.00–1.00	Excellent
Participant 26	1.00	1.00–1.00	Excellent
Participant 27	1.00	1.00–1.00	Excellent
Participant 28	0.80	0.77–0.94	Good
Participant 29	1.00	1.00–1.00	Excellent
Participant 30	1.00	1.00–1.00	Excellent
Participant 31	1.00	1.00–1.00	Excellent
Participant 32	1.00	1.00–1.00	Excellent
Participant 33	1.00	1.00–1.00	Excellent
Participant 34	1.00	1.00–1.00	Excellent
Participant 35	1.00	1.00–1.00	Excellent

4. Discussion

Currently, no nationally agreed RTW screening test exists within UK fire services. To develop a nationally agreed test, previous research identified the tasks to be included [1]; however, the reliability was yet to be determined. This study aimed to assess the inter-rater and intra-rater reliability of a RTW screening test to be used on UK firefighters following injury. Results showed that the overall inter-rater reliability between all screening tasks was interpreted as Good (ICC = 0.77–0.79) for both rating sessions and the intra-rater reliability was interpreted between Moderate-Excellent (ICC = 0.63–1.00), with 97% of participants reliability being interpreted between Good-Excellent (ICC = 0.76–1.00).

Employers often reply upon screening tests assessing functional capacity to assist in determining an individual's work capacity relevant to their specific job role [39]. The results from these screening tests can aid with the decision to allow an individual to return to their

job role or help provide further rehabilitation interventions [39]. In addition, screening tests help provide a consistent method of assessment used within a workforce [39,40].

Similar studies assessing functional capacity set an ICC criterion of >0.75 for screening tests to be classed as “reliable” [41,42]. The inter-rater results from this study (ICC = 0.77–0.79) suggest that this screening test can be used to identify if a firefighter undertaking the RTW tasks passes or fails on a reliable basis. These data are important, as it is essential to have reliable screening methods when assessing a firefighter’s ability to complete operational tasks with the correct technique to determine their physical readiness to return to operational duties [43]. By identifying reliable RTW screening tests for the physically demanding role of a firefighter is key to help highlight those firefighters who are able to undertake their role effectively, therefore improving the safety of themselves, their colleagues and the public on their RTW [44]. Previous research concluded that reliability studies should focus on multiple raters of varying background and experiences [45,46]. This was achieved as thirty-five participants from eight fire and rescue service regions across the UK completed both of the required screening sessions. The results obtained were provided from professionals working across a range of occupational health, fitness and operational training departments, with an average of 12.40 + 8.11 years’ experience.

Intra-rater reliability is important in such measures because it determines the accuracy of an assessment where a single rater may make multiple assessments over time [47,48]. Our study showed that intra-rater reliability ICC ranged from 0.63–1.00 with 97% of participants achieving a reliability interpretation above the ICC criterion of >0.75 as shown in previous studies [41,42]. This suggests that the RTW screening test for firefighters following musculoskeletal injury used in this current study is suitable for repeated measures in assessing a firefighter’s readiness to RTW.

Reliability for repeated measures is especially important in assessing the consistency of the RTW screening test. A lack of consistency for RTW assessments following injury was perceived as a barrier amongst firefighters experienced during their RTW process [49]. Therefore, if this RTW screening test was used as good practice within UK fire and rescue services, it could potentially remove this barrier by adding trustworthiness to the RTW process and help to increase the consistency of the RTW assessment.

The online design of the RTW screening test used in this study increased the ease of access for participants, as they were able to complete the rating sessions for the RTW screening test on desktop or portable devices, including laptops, smartphones, and tablets. As a result, future practice could allow for this RTW screening test to be used in various locations across different fire and rescue services provided they have the required equipment for the screening test. This could increase the availability in RTW screening test appointments within fire and rescue services and as a result, help decrease potential waiting times for firefighters looking to return to their job role. Further research is needed to assess the validity of the use of this RTW screening test to help reduce firefighter reinjury rates in UK fire and rescue services.

Strengths and Limitations

This study included experts from fire service fitness and occupational health departments as well as operational firefighters in the UK. Experts from fire and rescue services across the UK were invited to participate but this study participation did not include representation from every fire and rescue service in the UK. Nevertheless, those who did take part provided representation from a large range of UK fire and rescue services. The online approach helped reduce the impact on the participants. This study was focused on participants working for UK fire and rescue services. The online approach allows for representation from fire and rescue services internationally in future studies.

A training video and clear SCF provided the participants with the information required of what was required from them. The videos filmed, provided clear visual information for participants to decide if the video should be marked as pass or fail. The design of the website allowed the SCF and the assessment videos to be on the webpage. This allowed

participants to use one screen/device and it could be completed on a computer desktop, tablet or mobile device.

5. Conclusions

The return-to-work screening test used in this study provided evidence that it has good inter-rater reliability (ICC = 0.77–0.79) and good-excellent intra-rater reliability (ICC = 0.76–1.00) for 97% of participants. Due to the reliability of this screening test, it allows conclusions to be made from the results which can inform a return-to-work decision for a firefighter. This return-to-work screening test provides a method for fitness and occupational health experts as well as operational trainers working for UK fire and rescue services to refer to when assessing the readiness of a firefighter to return to operational duties. If used, this screening test could increase the consistency of return-to-work process within UK fire and rescue services and add trustworthiness to the decisions made. Further research is needed on the validity of this return-to-work screening test in reducing reinjury rates within firefighters.

Author Contributions: Conceptualization, L.N.; methodology, L.N.; software, L.N.; validation, L.N.; formal analysis, L.N.; investigation, L.N.; resources, L.N.; data curation, L.N.; writing—original draft preparation, L.N.; writing—review and editing, L.N.; visualization, L.N.; supervision, J.M. and A.M.; project administration, L.N.; funding acquisition, Not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The University of Essex research ethics committee granted approval for the study. (Ethics reference; ETH2122-1516).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Screening Video Criteria Form (SCF).

Screening Test	Pass Criteria
Putting on and Removal of Breathing Apparatus Set	Firefighter squats behind the BA set with the top of the cylinder between their feet.
	Firefighter stands the set onto the cylinder bump stop so that it is in a vertical position
	Firefighter draws the set close to their body, bending the knees and keeping the spine in a neutral position whilst standing up.
	Firefighter places the right-hand shoulder strap over their right shoulder and then places left arm into the left shoulder strap.
Ladder lift simulator	Firefighter fastens shoulder straps and then fastens waist belt buckle ensuring that the belt is not twisted.
	Firefighter fastens chest and waist clips then stands up straight.
	Firefighter starts with an underhand grip on the bar with palms facing upwards.
	Firefighter bicep curls the bar, keeping back straight.
	Firefighter rotates their wrists one at a time so that the bar is now gripped with the bottom of their palms facing outwards.
Firefighter shoulder presses the bar, without any assistance from the lower body, ensuring that the bar is above the designated yellow marker.	
Firefighter lowers the bar in a controlled manner back to chest height, changing wrists back over so that the bottoms of their palms are facing towards them.	
Firefighter lowers bar to the start position by extending their arm and places the bar into the rest position, bending their knees if required.	

Table A1. Cont.

Screening Test	Pass Criteria
Ladder carry simulator	<p>Firefighter starts with their feet flat on the ground and positioned between hip and shoulder width apart. Firefighter squats down and grasps the dumbbell in one hand. Firefighter lifts the dumbbell off the floor, by extending their knees and hips, until standing in an upright position. Firefighters back should maintain a rigid spine with a constant torso angle to the floor. Firefighter holds the dumbbell down by their side with a straight arm and proceeds to walk forwards, keeping an upright position. Once the firefighter has reached the required distance, they lower the dumbbell to the floor whilst maintaining a neutral spine, flexing the hips and squatting. Firefighter turns around and repeats the process, lifting the dumbbell with the opposite hand. Once the firefighter has reached the required distance, they lower the dumbbell to the floor whilst maintaining a neutral spine, flexing the hips and squatting.</p>
Light portable pump lift and carry simulator	<p>Firefighter starts with their feet flat on the ground and positioned between hip and shoulder width apart. Firefighter squats down and grasps the barbell with both hands. Firefighter lifts the barbell off the floor by extending knees and hips until they are in an upright position. Their back should maintain a rigid spine with a constant torso angle to the floor. Firefighter holds the barbell down in front of them with straight arms and proceeds to walk forwards keeping in an upright position. Once the firefighter has reached the required distance, the barbell is lowered to the floor whilst maintaining a neutral spine, flexing the hips and squatting.</p>
Casualty evacuation	<p>Firefighter grasps the casualty, with both hands, by the carrying handle located at the back of the dummy's head. Firefighter positions themselves body upright, back neutral and legs slightly bent. Firefighter drags casualty by walking backwards. Once the firefighter reaches the required distance, grasp on the carrying handle is released in a controlled manner.</p>
Hose run	<p>Firefighter places their foot on the hose and grasps the lugs with their hands. Firefighter lifts hose to shoulder height and holds it to the side of their body. Firefighter runs hose out until the end is reached and the female coupling is placed carefully on the ground. Firefighter runs back and underruns the hose.</p>
Ladder climb and leg lock	<p>Firefighter climbs the ladder and takes a leg lock. Firefighter releases their hands from the ladder, outstretches both arms to the side and looks over each shoulder. Firefighter regains hand hold on the ladder, removes their leg lock and descend the ladder to the ground.</p>
Confined space crawl	<p>Firefighter crawls on their hands and knees through the confined space. Once the firefighter reaches the end of the confined space, they turn around and make their way back to the start. The crawl should be completed in a calm and controlled manner by the firefighter.</p>

References

- Noll, L.; Mallows, A.; Moran, J. Consensus on tasks to be included in a return to work assessment for a UK firefighter following an injury: An online Delphi study. *Int. Arch. Occup. Environ. Health* **2021**, *94*, 1085–1095. [[CrossRef](#)] [[PubMed](#)]
- Yedulla, N.R.; Battista, E.B.; Koolmees, D.S.; Montgomery, Z.A.; Day, C.S. Workplace-related musculoskeletal injury trends in the United States from 1992 to 2018. *Injury* **2022**, *53*, 1920–1926. [[CrossRef](#)] [[PubMed](#)]
- Mock, C.N.; Kobusingye, O.; Nugent, R.; Smith, K.R.; Mock, C.N.; Kobusingye, O.; Nugent, R.; Smith, K. *Injury Prevention and Environmental Health: Key Messages from Disease Control Priorities*; World Bank: Washington, DC, USA, 2018.
- Wise, S.R.; Trigg, S.D. Optimizing health, wellness, and performance of the tactical athlete. *Curr. Sports Med. Rep.* **2020**, *19*, 70–75. [[CrossRef](#)] [[PubMed](#)]
- Orr, R.; Simas, V.; Canetti, E.; Schram, B. A profile of injuries sustained by firefighters: A critical review. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3931. [[CrossRef](#)] [[PubMed](#)]
- Scofield, D.E.; Kardouni, J.R. The tactical athlete: A product of 21st century strength and conditioning. *Strength Cond. J.* **2015**, *37*, 2–7. [[CrossRef](#)]
- Baidwan, N.K.; Gerberich, S.G.; Kim, H.; Ryan, A.D.; Church, T.R.; Capistrant, B. A longitudinal study of work-related injuries: Comparisons of health and work-related consequences between injured and uninjured aging United States adults. *Inj. Epidemiol.* **2018**, *5*, 1–9. [[CrossRef](#)] [[PubMed](#)]
- Cornell, D.J.; Gnacinski, S.L.; Ebersole, K.T. Functional movement quality of firefighter recruits: Longitudinal changes from the academy to active-duty status. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3656. [[CrossRef](#)] [[PubMed](#)]
- Cancelliere, C.; Donovan, J.; Stockkendahl, M.J.; Biscardi, M.; Ammendolia, C.; Myburgh, C.; Cassidy, J.D. Factors affecting return to work after injury or illness: Best evidence synthesis of systematic reviews. *Chiropr. Man. Ther.* **2016**, *24*, 1–23. [[CrossRef](#)]

10. Tuckwell, N.L.; Straker, L.; Barrett, T.E. Test-retest reliability on nine tasks of the Physical Work Performance Evaluation. *Work* **2002**, *19*, 243–253.
11. La Reau, A.C.; Urso, M.L.; Long, B. Specified Training to Improve Functional Fitness and Reduce Injury and Lost Workdays in Active Duty Firefighters. *J. Exerc. Physiol. Online* **2018**, *21*, 5.
12. Anderson, C.; Briggs, J. A study of the effectiveness of ergonomically-based functional screening tests and their relationship to reducing worker compensation injuries. *Work* **2008**, *31*, 27–37. [[PubMed](#)]
13. Siddall, A.G.; Stevenson, R.D.; Turner, P.J.; Bilzon, J.L. Physical and physiological performance determinants of a firefighting simulation test. *J. Occup. Environ. Med.* **2018**, *60*, 637–643. [[CrossRef](#)] [[PubMed](#)]
14. Treweek, A.J.; Tipton, M.J.; Milligan, G.S. Development of a physical employment standard for a branch of the UK military. *Ergonomics* **2019**, *62*, 1572–1584. [[CrossRef](#)] [[PubMed](#)]
15. Bissett, D.; Bissett, J.; Snell, C. Physical agility tests and fitness standards: Perceptions of law enforcement officers. *Police Pract. Res.* **2012**, *13*, 208–223. [[CrossRef](#)]
16. Stevenson, R.D.; Siddall, A.G.; Turner, P.F.; Bilzon, J.L. A task analysis methodology for the development of minimum physical employment standards. *J. Occup. Environ. Med.* **2016**, *58*, 846–851. [[CrossRef](#)] [[PubMed](#)]
17. Blacker, S.D.; Rayson, M.P.; Wilkinson, D.M.; Carter, J.M.; Nevill, A.M.; Richmond, V.L. Physical employment standards for UK fire and rescue service personnel. *Occup. Med.* **2016**, *66*, 38–45. [[CrossRef](#)] [[PubMed](#)]
18. Young, P.M.; Gibson, A.S.C.; Partington, E.; Partington, S.; Wetherell, M.A. Psychophysiological responses in experienced firefighters undertaking repeated self-contained breathing apparatus tasks. *Ergonomics* **2014**, *57*, 1898–1906. [[CrossRef](#)]
19. Stevenson, R.D.; Siddall, A.G.; Turner, P.F.; Bilzon, J.L. Implementation of physical employment standards for physically demanding occupations. *J. Occup. Environ. Med.* **2020**, *62*, 647–653. [[CrossRef](#)]
20. Van Melick, N.; Pronk, Y.; Nijhuis-van der Sanden, M.; Rutten, S.; van Tienen, T.; Hoogeboom, T. Meeting movement quantity or quality return to sport criteria is associated with reduced second ACL injury rate. *J. Orthop. Res.* **2021**, *40*, 117–128. [[CrossRef](#)]
21. Zore, M.R.; Kregar Velikonja, N.; Hussein, M. Pre-and Post-Operative Limb Symmetry Indexes and Estimated Preinjury Capacity Index of Muscle Strength as Predictive Factors for the Risk of ACL Reinjury: A Retrospective Cohort Study of Athletes after ACLR. *Appl. Sci.* **2021**, *11*, 3498. [[CrossRef](#)]
22. De Vos, R.-J.; Reurink, G.; Goudswaard, G.-J.; Moen, M.H.; Weir, A.; Tol, J.L. Clinical findings just after return to play predict hamstring re-injury, but baseline MRI findings do not. *Br. J. Sports Med.* **2014**, *48*, 1377–1384. [[CrossRef](#)] [[PubMed](#)]
23. King, E.; Richter, C.; Daniels, K.A.; Franklyn-Miller, A.; Falvey, E.; Myer, G.D.; Jackson, M.; Moran, R.; Strike, S. Biomechanical but Not Strength or Performance Measures Differentiate Male Athletes Who Experience ACL Reinjury on Return to Level 1 Sports. *Am. J. Sports Med.* **2021**, *49*, 918–927. [[CrossRef](#)] [[PubMed](#)]
24. Fältström, A.; Kvist, J.; Bittencourt, N.F.N.; Mendonça, L.D.; Hägglund, M. Clinical Risk Profile for a Second Anterior Cruciate Ligament Injury in Female Soccer Players After Anterior Cruciate Ligament Reconstruction. *Am. J. Sports Med.* **2021**, *49*, 1421–1430. [[CrossRef](#)]
25. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [[CrossRef](#)] [[PubMed](#)]
26. Evans, A.M.; Rome, K.; Peet, L. The foot posture index, ankle lunge test, Beighton scale and the lower limb assessment score in healthy children: A reliability study. *J. Foot Ankle Res.* **2012**, *5*, 1–5. [[CrossRef](#)]
27. Lee, K.M.; Lee, J.; Chung, C.Y.; Ahn, S.; Sung, K.H.; Kim, T.W.; Lee, H.J.; Park, M.S. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clin. Orthop. Surg.* **2012**, *4*, 149–155. [[CrossRef](#)]
28. Apeldoorn, A.T.; Den Arend, M.C.; Schuitemaker, R.; Egmond, D.; Hekman, K.; Van Der Ploeg, T.; Kamper, S.J.; Van Tulder, M.W.; Ostelo, R.W. Interrater agreement and reliability of clinical tests for assessment of patients with shoulder pain in primary care. *Physiother. Theory Pract.* **2021**, *37*, 177–196. [[CrossRef](#)] [[PubMed](#)]
29. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.-G. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 1149–1160. [[CrossRef](#)]
30. Bronner, S.; Lassey, I.; Lesar, J.R.; Shaver, Z.G.; Turner, C. Intra-and inter-rater reliability of a ballet-based dance technique screening instrument. *Med. Problic Perform. Artist.* **2020**, *35*, 28–34. [[CrossRef](#)]
31. Yurcheshen, M.E.; Pigeon, W.; Marcus, C.Z.; Marcus, J.A.; Messing, S.; Nguyen, K.; Marsella, J. Interrater reliability between in-person and telemedicine evaluations in obstructive sleep apnea. *J. Clin. Sleep Med.* **2021**, *17*, 1435–1440. [[CrossRef](#)]
32. Dürregger, C.; Adamer, K.A.; Pirchl, M.; Fischer, M.J. Inter-rater reliability of a newly developed gait analysis and motion score. *J. Orthop. Trauma Rehabil.* **2020**, 2210491720967366. [[CrossRef](#)]
33. Muhsen, Z.F.; Maaita, A.; Odah, A.; Nsour, A. Moodle and e-learning Tools. *Int. J. Modern Educ. Comput. Sci.* **2013**, *5*, 1. [[CrossRef](#)]
34. Mayers, S.; Lee, M. *iMovie. Learn OS X Lion*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 267–274.
35. Oliver, B.; Cheema, S.; Dunbar, A.; Richards, J. The reliability of a new functional balance protocol for use in sports requiring jump landing tasks. *Physiother. Pract. Res.* **2017**, *38*, 79–85. [[CrossRef](#)]
36. Ilag, B.N. Microsoft Teams Overview. In *Understanding Microsoft Teams Administration*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 1–36.
37. Wei, C.-H.; Shen, S.-C.; Duh, Y.-C.; Tsai, K.Y.; Chen, H.-A.; Huang, S.-W. Inter-rater and intra-rater reliability of the current assessment model and tools for laparoscopic suturing. *Surg. Endosc.* **2022**, *36*, 6586–6591. [[CrossRef](#)] [[PubMed](#)]

38. Wagner, W.E., III. *Using IBM® SPSS® Statistics for Research Methods and Social Science Statistics*; Sage Publications: New York, NY, USA, 2019.
39. Pransky, G.S.; Dempsey, P.G. Practical aspects of functional capacity evaluations. *J. Occup. Rehabil.* **2004**, *14*, 217–229. [[CrossRef](#)]
40. Hoyle, D.; Mecham, J. Current Concepts in Functional Capacity Evaluation: A Best Practices Guideline. 2018. Available online: https://www.orthopt.org/uploads/content_files/files/2018_Current_Concepts_in_OH_PT_FCE_06_20_18_FINAL%281%29.pdf (accessed on 7 November 2022).
41. Reneman, M.; Brouwer, S.; Meinema, A.; Dijkstra, P.; Geertzen, J.; Groothoff, J. Test–retest reliability of the Isernhagen work systems functional capacity evaluation in healthy adults. *J. Occup. Rehabil.* **2004**, *14*, 295–305. [[CrossRef](#)]
42. Reneman, M.; Dijkstra, P.; Westmaas, M.; Göeken, L. Test-retest reliability of lifting and carrying in a 2-day functional capacity evaluation. *J. Occup. Rehabil.* **2002**, *12*, 269–275. [[CrossRef](#)]
43. Matheson, L.; Duffy, S.; Maroof, A.; Gibbons, R.; Duffy, C.; Roth, J. Intra-and inter-rater reliability of jumping mechanography muscle function assessments. *J. Musculoskelet. Neuronal Interact.* **2013**, *13*, 480–486.
44. Stevenson, R.D.; Siddall, A.G.; Turner, P.J.; Bilzon, J.L. Validity and reliability of firefighting simulation test performance. *J. Occup. Environ. Med.* **2019**, *61*, 479–483. [[CrossRef](#)]
45. Onate, J.A.; Dewey, T.; Kollock, R.O.; Thomas, K.S.; Van Lunen, B.L.; DeMaio, M.; Ringleb, S.I. Real-time intersession and interrater reliability of the functional movement screen. *J. Strength Cond. Res.* **2012**, *26*, 408–415. [[CrossRef](#)]
46. Valdez, R. Health Care Professionals Confidence and Experience With Functional Movement Screen Testing. *Mast. Educ. Hum. Mov. Sport Leisure Stud. Grad. Proj.* **2017**, *41*. Available online: https://scholarworks.bgsu.edu/hmsls_mastersprojects/412017 (accessed on 7 November 2022).
47. Stanton, R.; Wintour, S.-A.; Kean, C.O. Validity and intra-rater reliability of MyJump app on iPhone 6s in jump performance. *J. Sci. Med. Sport* **2017**, *20*, 518–523. [[CrossRef](#)] [[PubMed](#)]
48. Bartolo, C.; Miller, K.; Seals, R.; Stotesbery, C. Examination of Tester Reliability Utilizing the Limits of Stability Test on the Neurocom Balance Master for Assessing Balance in Healthy Individuals. *Phys. Ther. Sch. Proj.* **2002**, *34*. Available online: <https://commons.und.edu/pt-grad/342002> (accessed on 7 November 2022).
49. Noll, L.; Mallows, A.; Moran, J. Psychosocial barriers and facilitators for a successful return to work following injury within firefighters. *Int. Arch. Occup. Environ. Health* **2022**, *95*, 331–339. [[CrossRef](#)] [[PubMed](#)]