Applied Ergonomics 60 (2017) 255–259

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

The sensitivity of a military-based occupational fitness test of muscular strength

Kane J. Middleton [a, b, *], Greg L. Carstairs [c], Joanne N. Caldwell [a], Daniel C. Billing [c], Ben Beck [c, d]

[a] Centre for Human and Applied Physiology, School of Medicine, University of Wollongong, Wollongong, Australia
[b] La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, La Trobe University, Melbourne, Australia
[c] Land Division, Defence Science and Technology Group, Melbourne, Australia
[d] Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 25 October 2015
Received in revised form
6 December 2016
Accepted 7 December 2016

Keywords:
Physical employment standards
Task-related predictive test
Specificity
Ergonomics
Lifting

ABSTRACT

The risk of low back pain and injury during manual materials handling is increased if personnel are not physically capable of safely performing such tasks. To establish predictive relationships and develop a test cut-score, 69 participants performed a critical military lifting task to a 1.5-m platform (pack lift) and two task-related predictive tests (box lift to 1.5 m and 1.3 m). The pack lift was strongly correlated with both the 1.5-m ($R^2 = 0.85$) and 1.3-m box lifts ($R^2 = 0.82$). Both tests had similar sensitivity (range 0.85–0.94) with the 1.3-m test having higher specificity when compared with the 1.5-m lift. Increasing the test cut-score with the application of a safety factor increased the number of false positives and true negatives for both tests. Organisations must carefully assess their risk acceptance when applying safety factors to test cut-scores as the classification (pass/fail) of personnel may be affected.

1. Introduction

Manual materials handling is a high risk work activity for the occurrence of low back pain and injury (Lötters et al., 2003). This injury risk is increased if personnel are not physically capable of safely performing set tasks (Harbin and Olsen, 2005; Rosenblum and Shankar, 2006). To reduce the likelihood of these injuries occurring, commensurate levels of muscular strength and endurance relative to task demand are necessary. Specifically, lifting is a common component of manual materials handling within many occupations (Rayson, 1998) and can account for up to 78% of all manual handling tasks within certain roles (Beck et al., 2012).

Utilising physical employment standards can ensure employees possess physical capabilities commensurate with job demands (Taylor and Groeller, 2003). Task simulations, task-related predictive tests and generic predictive tests are the three most common types of physical employment tests used to assess a person's ability to complete job task requirements (Gumieniak et al., 2011; Rayson, 1998; Rayson et al., 2000; Taylor and Groeller, 2003). Task simulations replicate job tasks (Rayson, 1998), whereas task-related predictive tests maintain some job task characteristics (Payne and Harvey, 2010). In contrast, generic predictive tests assess general physical capacity but generally lack face validity relative to job tasks (Rayson et al., 2000).

Generic predictive tests have previously been used to assess muscular strength (Harman et al., 2008; Vanderburgh, 2008) in a military setting. However, it has recently been shown that tests such as push-ups and pull-ups are weaker predictors of lifting task performance when compared with task-related predictive tests (Carstairs et al., 2016). Several occupational lifting-based tests have previously been developed (e.g. LIFTEST by Kroemer (1985); incremental lift test by Dempsey et al. (1998)) and successfully implemented in the military environment (Ayoub et al., 1987; Rayson et al., 2000; Williams and Wilkinson, 2007). Carstairs et al. (2016) assessed the utility of an Australian Army in-service task-related predictive test, the box-lift-and-place, and found that it was strongly correlated with performance in a number of common simulated army job tasks, including a pack-lift-and-place. This test has been developed in order to overcome limitations in the implementation of a task simulation using a pack in a real world

* Corresponding author. La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, La Trobe University, Melbourne, Australia.
E-mail address: k.middleton@latrobe.edu.au (K.J. Middleton).

http://dx.doi.org/10.1016/j.apergo.2016.12.004
0003-6870/

environment, including standardisation (of object and weighting) and safety (handles). The box-lift-and-place test has also been shown to be able to predict lift performance across a range of lift heights (Savage et al., 2015). Although this test strongly predicts lift performance across heights (Savage et al., 2015), the sensitivity and specificity of this test has not previously been assessed.

Sensitivity and specificity are often critical for the scientific defensibility of a physical employment standard assessment. Minimising false negatives will increase test sensitivity and reduce the number of personnel that fail the assessment even though they are capable of performing the job task. Minimising false positives will increase specificity and reduce the number of personnel that may be deemed as suitable but in fact may not be physically capable of performing the job task, thereby exposing them to an increased risk of injury (Tipton et al., 2013). The first aim of this study was to assess the criterion validity of two potential task-related predictive tests of muscular strength and develop test standards. The second aim was to assess each tests' sensitivity and specificity and investigate the trade-off between sensitivity and specificity when adjusting test cut-off points.

2. Material and methods

2.1. Participants

Sixty-nine soldiers with (mean ± SD) 18.4 ± 10.8 years of military service participated in testing and included 45 males and 24 females. Participants were 41.2 ± 10.4 years of age, 1.72 ± 0.09 m tall with a shoulder height of 1.42 ± 0.09 m and a body mass of 81.9 ± 15.9 kg. According to the responses to a single-item question on leisure-time vigorous physical activity, participants were undertaking vigorous physical activity at least three times per week. All participants gave written informed consent to procedures approved by the Australian Defence Human Research Ethics Committee (Protocol number: 491–07, ADF Physical Employment Standards).

2.2. Experimental approach to the problem

Three lifting activities [pack lift to 1.5 m (criterion task simulation); box lift to 1.5 m and box lift to 1.3 m (task-related predictive tests)] were conducted over three testing sessions. A minimum of 48-hours rest was provided between testing sessions. The order in which participants conducted the lifting sessions was balanced to reduce the effects of systematic bias and any learning effect. To ensure safety and minimise day-to-day variation in performance as a result of environmental conditions, all experimentation was conducted indoors.

2.3. Lifting assessments

Prior to testing, participants performed a general warm-up followed by task-specific practice and familiarisation. Participants were provided with a task demonstration that included verbal instruction and visual demonstration of each task emphasising correct lifting technique. Participants wore standard issue disruptive pattern combat uniform and boots with 22-kg weighted vests to simulate the weight of webbing, weapon, and body armour.

Within the context of the Australian Army, a number of lifting tasks were previously identified as critical for soldiers to successfully complete; including stores and pack lifting (Carstairs et al., 2016; Savage et al., 2012). Specifically, the 20-kg individual pack lift to a 1.5-m platform has been identified as the criterion, or critical, muscular strength task of the Australian Army. Given that the hand-to-coupling location of the pack and box were different (pack: inferior-to-lateral side, box: 0.2 m from base on lateral side), it was decided to also test a lift to 1.3 m.

2.3.1. Individual pack lift

The individual pack lift (PL) required participants to squat down and lift a standard military pack (Large Field Pack, 1994, 0.8 × 0.65 × 0.3 m) from the ground to between knuckle and chest height, pausing briefly, taking one step forward and subsequently lifting from between knuckle and chest height onto a 1.5-m platform.

2.3.2. Box lift and place to 1.5-m

The 1.5-m box lift and place ($BLP_{1.5}$) required participants to squat down and lift a plastic box (Trimcast Rotomoulders Pty Ltd, 0.35 × 0.35 × 0.35 m; metal handles at 0.2 m from base of box) from the ground to between knuckle and chest height, pausing briefly, taking one step forward and subsequently lifting from between knuckle and chest height onto a 1.5-m platform.

2.3.3. Box lift and place to 1.3-m

The 1.3-m box lift and place ($BLP_{1.3}$) required participants to squat down and lift a plastic box (Trimcast Rotomoulders Pty Ltd, 0.35 × 0.35 × 0.35 m; metal handles at 0.2 m from base of box) from the ground to between knuckle and chest height, pausing briefly, taking one step forward and subsequently lifting from between knuckle and chest height onto a 1.3-m platform.

2.3.4. Lifting procedures

Participants conducted all three lifting assessments with increments in mass (5 kg) until lift failure. At lift failure, mass was reduced by 2.5 kg and the participant repeated the assessment. The final successful mass lifted was used as the performance measure. For participant safety, performance for each of the lifting tasks was capped at a maximum lifted mass of 60 kg. Technical exclusion criteria were also applied to ensure that safe lifting techniques were employed at all times. This included enforcing distinct stages of the lift to reduce the use of momentum throughout the movement, subjectively assessing levels of back hyperextension to ensure that the spine was in a neutral position throughout the lift and controlled placement of the item prior to and at lift completion.

2.4. Statistical analysis

A total of five participants were excluded from analysis due to reaching the *a priori* exclusion criteria of a 60 kg lift in more than one activity. All statistical analyses were conducted in IBM SPSS Statistics (v21.0.0.0, IBM Corporation, USA). Lift data violated assumptions of normality. Therefore, a Friedman test ($\alpha = 0.05$) was used to compare performance scores in the muscular strength tasks. Where a significant difference between mean ranks was found, a Wilcoxon signed-rank test with a Bonferroni correction ($\alpha = 0.017$) was performed to determine differences between each pair of lifting tasks. Bivariate correlations were used to determine the relationships between pack lift mass and the box lift and place to 1.5 m and 1.3 m. A calculation method developed by Lee and Preacher (2013) was used to test the difference between the two dependent correlations (box lifts) with one variable in common (pack lift). This calculation converts each correlation coefficient into a z-score using Fisher's r-to-z transformation then computes the asymptotic covariance of the estimates using equations from Steiger (1980).

Simple linear regression was used to determine the box-lift-and-place (dependent variable) test cut-score from maximal lifting capacity in the pack-lift-and-place task simulation (independent variable) using an inverse prediction approach as per Beck