Use of Physical Fitness Assessments in Tactical Populations

Robin M. Orr, PhD, TSAC-F*D,1,2 Robert Lockie, PhD, TSAC-F*D,1,3 Gemma Milligan, PhD,4 Cheryl Lim, BSc (Sport Science and Management), TSAC-F,5 and Jay Dawes, PhD, CSCS,*D, NSCA-CPT*D, TSAC-F, FNSCA1,6
1Tactical Research Unit, Bond University, Gold Coast, Australia; 2Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia; 3Department of Kinesiology, Center for Sport Performance, California State University, Fullerton, Fullerton, California; 4School of Sport, Health and Exercise Science, University of Portsmouth, Portsmouth, United Kingdom; 5Soldier Development Branch, Centre of Excellence for Soldier Performance, Singapore; and 6School of Kinesiology, Applied Health and Recreation, Oklahoma State University, Stillwater, Oklahoma

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ABSTRACT

Physical fitness assessments for tactical occupations (e.g., military, law enforcement, and emergency services) can include predictive tests of anaerobic power, cardiovascular fitness, muscular endurance, muscular power, strength, agility, and/or simulated occupational tasks. Not only can these tests be used to assess the ability of someone to undertake the job role but they can be used to determine injury risk, training failure, and/or general health. This review discusses different uses for physical fitness assessments and considerations for their use in tactical populations.

INTRODUCTION

Physical fitness assessments are widely used in public safety organizations where there is a high physical demand. These organizations include the military (2,30,95), law enforcement (66,67), firefighter (7,83), and other rescue services, such as beach lifeguards (77). These professions often perform tasks that are highly physical in nature, and as such, physical assessments are often used at all stages of career progression to ensure that applicants (11,51), trained personnel (2), and those seeking specialist selection (30,70) have the necessary physical fitness to meet their specific training or employment obligations. Occupational physical fitness assessments can include tests of sustained anaerobic power, for example, a 75-yard pursuit (11,51) and 300-yard pursuit (87); aerobic power, for example, shuttle run assessments (2,17,73) and 2.4 km (2,12,47) and 3.2 km (30,87) distance runs; muscular endurance, for example, push-ups (14,27,30,37,38,43,47,78,100), sit-ups (14,27,30,37,38,43,47,100), and grip endurance (59); muscle strength, for example, grip strength (68,78,82), leg/back dynamometers, that is, a midhigh pull (17,18), and one or three repetition maximum tests (72,79); muscular power, for example, vertical (27,43,46,66,72) and broad (27,72) jumps; agility, for example, a change in direction test and T test (4,13); and simulations of occupational tasks, for example, Work Sample Battery Test (WSBT) (47), the Physical Employment Standards—Army (19), and the Royal Air Force (RAF) COMBAT-T (96).

Assessments of physical fitness can be used as a measure of injury risk (73,84,94), to provide information on general health and well-being (16), or to ensure job-task capability and employability (19,47,96). As physical fitness assessments can be used for different purposes, it is important for employers to understand the purpose of the tests they are using. This understanding will mitigate against lawsuits (6) and ensure validity of the tests results (60). In understanding the use of an assessment, consideration also needs to be given to how the cut score (i.e., minimally acceptable standard) is derived. Much debate of whether cut scores for physical fitness assessments should, or should not, account for age and sex is presupposed by their application. If a test is meant for selection (i.e., a physical employment standard), it is argued it should be age and sex free (93) because the nature of the task does not change. However, if the test is being used to act as a health or general fitness screen, there is an argument for the assessments to consider the age and sex of the individual (17). Therefore, the intent

KEY WORDS: military; law enforcement; fire and rescue; army; police
of this article is to discuss some of the different uses for physical fitness assessments in tactical populations and what these assessments mean for the associated benchmarks in regard to age and sex equality or neutrality.

### ASSESSMENTS FOR PREDICTING INJURIES/ATTRITION

When tactical personnel undergo a period of training, the risks of injury are known to increase (65,73). In new trainees, this is due to recruitment taking place from the general population, who often display varying levels of physical fitness and training experience (11,50,51). As such, the sudden increase in activity requirements resulting from physical and occupational training may exceed a trainee’s previous training load and current capabilities (65). This change in physical load brought on by an increase in physical conditioning, complexity of new physical tasks, and a reduced opportunity for recovery increases the risk of overtraining and potential injury (9,34,41,75). This increase in injury risk is likewise found in trained personnel undergoing specialist selection, whereby the selection process and subsequent specialist training cycle are intensive and physically demanding (30,70).

Poor levels of fitness, both metabolic and musculoskeletal, have been associated with a higher risk of training-related injuries and attrition (8,22,30,32,33,40,42,57,66,70,73,74). Pope et al. (73) found that the risk of attrition through failing to complete military training was approximately 25 times greater in trainees who scored poorly (bottom percentiles) on the 20-m progressive shuttle run test (PSRT) when compared to trainees who scored highly (upper percentiles). These findings led to the establishment of a level 75 on the PSRT for entry into the Australian Army. Even after a sustained physical program during initial training, poor metabolic fitness has been associated with an increased risk of training injury (57). A study by Meigh et al. (57) found that Army cadets with lower levels of fitness, similarly measured by the PSRT, were more likely to be injured during a trainee field exercise than those with higher levels of fitness, even after 6 months of physical training. These findings of increased risks of training-related injuries and attrition in military trainees are supported by studies from both the United Kingdom (8,80) and the United States (40,42) and have likewise been found in law enforcement trainees (15,41,48,49,66). As such, a trainee’s level of fitness, regardless of their age and sex, serves as an indicator of injury risk during training.

Aspects of fitness, both metabolic and musculoskeletal, have been found to be associated with the ability of specialist military (30) and police personnel (70) to succeed in specialist selection. In a study by Hunt, et al. (30), the researchers found that those who performed poorly in a loaded pack march (20 km with 28 kg), push-ups (two-second cadence), and sit-ups (three-second cadence) were more likely to fail specialist selection. Orr et al. (70) found that levels of performance by specialist tactical response police officers in pull-ups and push-ups in 2 minutes, seven-stage sit-up, and a lift and carry task for time were significantly and positively correlated (rs = 0.362–0.508, p = 0.010–0.042) with a level of selection success. In both examples, the specialist selection courses were physically demanding and included relatively high volume of physical training, physical task performance (including loaded pack marching and victim recovery), and limited recovery opportunities. During these selection courses, a primary selection success saw applicants complete the course (i.e., did not suffer an injury). In addition, applicants were generally graded based on their performance on physical, tactical, and technical tasks. These findings suggest that even in well-trained personnel, any physical performance deficiencies relative to requirements can negatively affect selection.

Considering the use of physical fitness assessments as a predictor of injury, 2 points require consideration. First, although there may be concern that these fitness measures are not related to actual job requirements, it should be noted that undergoing and completing training is the job requirement for the trainee (i.e., they are employed to complete training) and often this training is more physically demanding than occupational service (65). Having these personnel injured or fail to complete training means that they in effect fail to perform their daily training duties.

Second, a noted concern lies in the use of physical fitness assessments to predict injury risk or training failure with research findings often conflicting. For example, in 9 studies investigating relationships between a push-up test and injuries, the results were almost evenly divided with 5 studies finding relationships (1,41,44,62,91) and 4 studies failing to document significant relationships (25,53,54,84). Two potential reasons for these conflicting results include the lack of contextual specificity and a ceiling effect. Context specificity refers to the similarity of the assessment regarding the mechanism of injury. For example, if one training institution completed high volumes of running as part of training, a run assessment may be a predictor of risk; the inverse may be true if the institution had a low run requirement, whereby a run assessment may not predict injury risk (94). As such, relationships between fitness measures and injury/failing attrition risk must be context specific (i.e., training environment). Similarly, research results may be conflicting because of a ceiling effect, whereby the fitness level of trainees is well above that required of a given task. For example, if the trainees are, in general, very aerobically fit, yet the training they are undertaking has a very low aerobic fitness requirement, they may be well above a potential injury threshold (94). Both Hunt et al. (30) and Orr et al. (70) discussed this limitation, whereby, in both of their studies, the aerobic fitness of the personnel involved in the studies was of a high performance standard. For example, in the study by Orr et al. (70), the specialist police undergoing selection training had an aerobic capacity of 52 mL·kg\(^{-1}\)·min\(^{-1}\), which is notably higher than that of general duties police, who average between 37.5 and 44.9 mL·kg\(^{-1}\)·min\(^{-1}\) (10,17). As such, this level of aerobic fitness in these specialist trainees may have been above that at which injuries are more likely occur.
Higher levels of physical fitness may greatly reduce the risk that personnel undergoing training will sustain an injury (36,39,81). Fitter personnel can perform activities at a lower percentage of their maximal capacity and are therefore able to perform tasks for longer, recover faster, and fatigue less rapidly (39). Thus, there is a rationale to use physical fitness assessments to identify those individuals who are at risk of injury or failure. Furthermore, as the occupational training undertaken by trainees is the same (i.e., all trainees of a given cohort complete the same activity), regardless of sex or age, the fitness requirements to meet the training physical demands without injury or failure are also the same. On this basis, physical fitness assessments designed to gauge injury and/or failure risk should be age and sex neutral.

When establishing cut scores for injury risk fitness assessments, the organization in question must decide on the level of risk they are willing to accept. If standards are raised, there may be less risk of trainee injury; however, the recruitment pool of trainees will be smaller. Conversely, lowering the fitness requirements could increase the recruitment pool of trainees with more people passing the assessment but likewise, increase the risk and incidence of injuries and increase the risk of agency separation (48). On this basis, the level of risk that an organization is willing to accept will affect recruitment levels, workforce size, and injury rates.

ASSESSMENTS FOR DETERMINING HEALTH AND WELL-BEING

The nature of tactical occupations can leave personnel exposed to a myriad of associated health-related concerns. Shift work, poor sleep, poor nutritional habits, smoking, and alcohol consumption have the potential to negatively affect the health and well-being of tactical personnel (29,97). Poor health can lead to long-term risks of comorbid diseases (3), which together with poor health, leads to increases in absenteeism (leading to increased organizational demands) (45,56,88).

Health-related conditions are of concern, for example, police officers and firefighters are at a greater risk of cardiovascular disease when compared to the general population (6,76,103). As such, physical fitness assessments can augment any medical health-related screening (blood pressure, waist-to-hip ratios, etc.). Orr et al. (63) and Sorenson et al. (85) both identified a loss in general fitness of law enforcement personnel over time. In the study by Orr et al. (63), differences in fitness were found to exist between new police trainees and serving officers. As age was not found to be a significant predictor of fitness test results (push-ups, \( p = 0.419 \); sit-ups, \( p = 0.111 \); 1.5-mile run, \( p = 0.81 \)), the authors suggested it was the nature of the work environment that led to these losses in fitness. The results are not surprising with research showing the negative impact of police work environment and shift work on desire to participate in physical exercise (26), nutritional choices (52), and sleep (24)—all of which affect physical fitness. Conversely, strong evidence supports the protective effects of high levels of fitness on major chronic diseases, such as coronary heart disease, hypertension, stroke, diabetes mellitus, osteoporosis, depression, and anxiety (35,58,99,101).

Regardless of the level of general health-related fitness selected, there are known differences in general fitness components between men and women (98,104) as well as those associated with aging (102). For example, in general, women and older persons tend to have lower levels of fitness than men or younger persons (17). Having comparative standards to those expected of the general population suggests that fitness standards designed to ensure basic health and well-being should consider an individual’s sex and age. Examples of sex-referenced and age-referenced assessments are commonly found in the military (2,20) and law enforcement (21,92). However, some research does suggest that there are differences in the fitness levels of new trainees from different subpopulations joining U.S. law enforcement agencies (61). Alternatively, in a study by Dawes et al. (16), the push-up scores of 518 police officers were at a “very good” or higher standard (20–29 years = 88%; 30–39 years = 94%; 40–49 years = 98%; 50–59 years = 100%) when compared to normative population standards. These findings in law enforcement populations bear consideration when developing standards through which to compare the general health and fitness of law enforcement personnel and raise the consideration as to whether standards should be set against the general population or against the specific population of that organization.

ASSESSMENTS FOR MEASURING OCCUPATIONAL PERFORMANCE

Typically, occupational assessments are based on the ability of personnel to complete required tasks deemed to be critical to the completion of the job (93). Given that these tasks remain extant regardless of the sex or age of the individual, they should be sex and age neutral (93). For example, if the job requirement is to lift and carry a 13 kg artillery shell, the weight of the shell will not change with the lifter’s sex or age. The importance of sex-neutral and age-neutral assessments of task capability is typified by the Australian Army Physical Employments Standards—Army (PESA) assessment (19), the U.S. Army Occupational Physical Assessment Test (23), the Army Combat Fitness Test (28), the U.S. Firefighter Candidate Physical Ability Test (31), the UK RAF COMBAT-T (96), and the South Australian Police “Fit for Duty” assessment (86). These assessments require personnel to complete given fitness tasks that are meant to replicate key occupational tasks in a given time, over a given distance, or with a given load, regardless of individual characteristics.

Defining occupational tasks and establishing commensurate assessments are challenging for some occupations where the frequency, duration, and work intensity of tasks demonstrate large variations and are difficult to
quantify and provide clear task descriptions. Police officers may have a predominantly sedentary occupation (e.g., completing desk work or driving a patrol car), although their duties can be physically demanding, ranging from patrolling large areas on foot and attending to a domestic incident to effecting an arrest of an uncooperative offender, all while carrying up to 10 kg in additional load (5, 55, 69, 90). Likewise, some tasks can be highly physical but performed infrequently, whereas less physically demanding tasks performed more frequently. Establishing which occupational tasks are physically demanding should indicate that occupational assessments can be challenging. Even if key tasks are identified, these tasks could vary within and by region. For example, a study by Orr et al. (64) found that police officers from an Australian state police force performed tasks either more or less frequently depending on whether their station was in a metropolitan, suburban, or rural region. Furthermore, even within the same region, a common task such as attending to a domestic incident was found to range from 2 to 94 minutes (64).

A further complication occurs when the same task itself varies depending on an individual’s role during that task. Soldiers from different units in the same military force have been found to carry different external loads for the same foot patrol, depending on what their military occupational specialty was, which determined the equipment required (71). Given the sheer myriad of tasks that can be performed by tactical personnel, the number of assessments (and commensurate time and equipment) needed to represent these tasks could be unsustainable, and thus one assessment is often used to assess multiple tasks. A farmer’s carry-style assessment could be used to replicate moving storeys, carrying stretchers, and dragging an injured person, yet the distances, loads, and speeds for the carry may represent none of the tasks individually. The downstream impact of reducing a variety of tasks to a single task designed to represent a group of tasks would likely create low face validity, and without an understanding of the assessment, concerns that an assessment does not meet typical occupational requirements can ensue.

A notable confounder comes from the use of general fitness assessments to predict performance on job-related assessments. In essence, this refers to a general fitness measure being compared with an occupational fitness measure, with the occupational fitness measure itself a generic compilation of measures derived to replicate physical requirements of occupational tasks. In a study investigating the relationships between fitness measures and performance on a WSBT, Lockie et al. (47) found that pull-ups accounted for 49% of the variance in solid wall climb ability. Likewise, research by Orr et al. (68) found several relationships between measures of fitness, marksmanship, and defensive tactics. However, while these fitness measures may be closely aligned to a solid wall climb, marksmanship assessment, or defensive tactics assessments, whether these representative tasks are valid occupational tasks may be questionable. Thus, 2 degrees of separation may occur, whereby a fitness measure is used to assess an occupational task measure that may (or may not) relate to an actual task and may have been collapsed down to represent multiple tasks.

Differences in occupational task standards should only exist when there are differences in task requirements and thus creating the need for modular style physical fitness assessments that account for role-specific differences. The Australian Army PESA standards provide an example, whereby, based on the Corps of service (e.g., infantry, other combat arms corps, and combat services support corps), the load carriage requirements differ in both weight carried and distance (19). Thus, although multiple challenges exist, once the physical requirement to perform a given task or group of tasks is established and a physical assessment developed, the assessment benchmarks remain extant.

**ENCAPSULATION AND PRACTICAL APPLICATION CONSIDERATIONS**

If physical fitness assessments are to be valid when used in tactical populations, the rationale behind the assessments needs to be well defined. For assessments that are used as a tool to predict risk of injury or training failure, the assessment standards need to be contextualized to the specifics of the organization and their training or occupational requirements rather than drawing on findings from other organizations or research into other populations. Furthermore, the organization itself must decide on the level of risk they are willing to accept.

General health and fitness assessments, which may augment medical screening assessments, are of use to monitor the tactical workforce and to identify those personnel at risk of poor health outcomes. These measures increase in importance when personnel may have to physically engage with the general population or, due to the nature of their work, are at a higher risk of health-related concerns (e.g., cardiovascular disease). Although these measures may be benchmarked against population normative data, the level of requirement (e.g., “average,” “above average,” or “excellent”) must be established and justifiable based on researched evidence pertaining to the given subpopulation.

Occupational fitness measures, although potentially collapsed down to a few measures to assess multiple tasks, can be used to ensure that personnel entering and retained in the workforce are able to perform expected tasks. Given that the tasks do not change because of individuality, these fitness measures should be age and sex neutral. As such, they may also be used as return-to-work performance indicators or for job realignment, whereby an individual who cannot meet a job-task requirement may be reallocated to another area with different tasks and hence potentially
different requirements. Implementation of an occupational standard could follow a traffic light system, rather than a binary pass/fail system, whereby personnel are deemed as acceptable (i.e., fit enough), uncertain (i.e., below desired standard but may reach standards), and unacceptable (i.e., below desired standard and is not expected to be able to reach standards) (89). The inclusion of the middle tier (i.e., uncertain) provides the opportunity for a qualified professional (e.g., strength and conditioning coach) to assist the member to increase their level of fitness to the required level.

Finally, it should be noted that all 3 uses of physical fitness assessments, be they injury risk identification, ensuring general health and well-being, or occupational task performance, are relevant to an organization. If a trainee or qualified individual is injured, their work capability and that of the organization can be affected. Similarly, if the person is of poor physical health, they will miss work because of illness and affect organizational outcomes. If an individual cannot complete a work task, then there are again downstream impacts to the organization. As such, all 3 uses of fitness assessments can in essence be considered occupationally specific.

It should also be noted that these 3 factors are not mutually inclusive. An individual can be healthy and physically able to perform occupational tasks but be working at near maximal efforts due to poor fitness and as such, at high risk of injury. Subsequently, an individual can be very physically fit and able to perform all required work tasks but unhealthy. Considering these potential mutual exclusivity relationships, an organization may use several fitness assessment frameworks—one to assess potential risk of injury or training failure, one to ensure general health and fitness commensurate with the general population, and one to ensure personnel can perform required work tasks. An example of this integrated approach is found in the Australian Army that uses an initial fitness assessment at the commencement of army training where the PSRT is used to inform trainee injury risk, a Basic Fitness Assessment as a measure of general health and fitness, and the PESA assessment for job tasks.

**CONCLUSION**

Physical assessments can be used in tactical occupations for various reasons, all of which can either directly or indirectly affect tactical outcomes. Although an integrated assessment framework can be used and include benchmarks that do and do not take sex or age into account, it is vital that the purpose behind these selected physical fitness assessments and any benchmark standards be well understood and scientifically valid. Failure to understand the intent of the physical fitness assessment can lead to personnel disgruntlement and potential legal action.

**Conflicts of Interest and Source of Funding:** The authors report no conflicts of interest and no source of funding.

**REFERENCES**

4. Beck AO, Clasey JL, Yates JW, et al. Relationship of physical fitness measures vs. occupational physical ability in campus
Use of Physical Fitness Assessments


Use of Physical Fitness Assessments


