JOHN BENNETT, MS, CSCS

ASSESSMENT BATTERIES—PRACTICAL DECISION MAKING FOR FIREFIGHTER FITNESS BATTERIES

Since fitness programs in the fire service are a support service to the primary job, it can be difficult to find experts or dedicate the time and resources to properly analyze these programs in order to ensure they are effective. In addition to standard health, fitness, or sport assessment guidelines, several guidelines exist for the fire service. Since each agency is different in its preferences, capabilities, and philosophies, applying a standard recommendation may not be possible. This article introduces basic decision-making considerations for fire service assessment batteries.

Given the varied types of training that may be necessary for firefighter performance, health, or injury risk reduction, there is a large quantity of possible end goals for each firefighter, crew, and department (2). As such, the rationale for choosing assessments can lead to an impractically lengthy assessment battery. Even after applying the most general needs analysis, narrowing down the optimal and necessary assessment types can be a very daunting task (Table 1).

A matrix can be shaped with more detail by specific muscle groups and joints, simplified by the primary motions (push, pull, squat, and rotate), globalized for the upper versus lower body, or any number of variables for the application (Table 2). Since a complete consensus is not always feasible due to evolving knowledge, leadership, organizational considerations, and different needs of individuals, the following should be considered when choosing a full assessment battery:

TABLE 1. FITNESS COMPONENT POTENTIAL ASSESSMENTS
Combining the body areas and fitness components would require several assessments to test every component individually in every anatomical area. This could be more detailed or more simplified based on philosophies and approach.

SAMPLE PERFORMANCE CATEGORIES TO ASSESS	BODY AREAS TO ASSESS
Posture	Shoulder
Flexibility	Elbow
Stability	Wrist
Balance	Lumbo-Pelvic
Coordination	Hip
Agility/Mobility	Knee
Strength	Ankle
Power	
Speed	
Local Muscular Endurance	
Cardiovascular Endurance/Aerobic Power	
Lactate Tolerance-Threshold-Recovery	
12 Measures	**7 Major Joints**
Assessments	
Many with anterior/posterior evaluations or various viable joint motions	

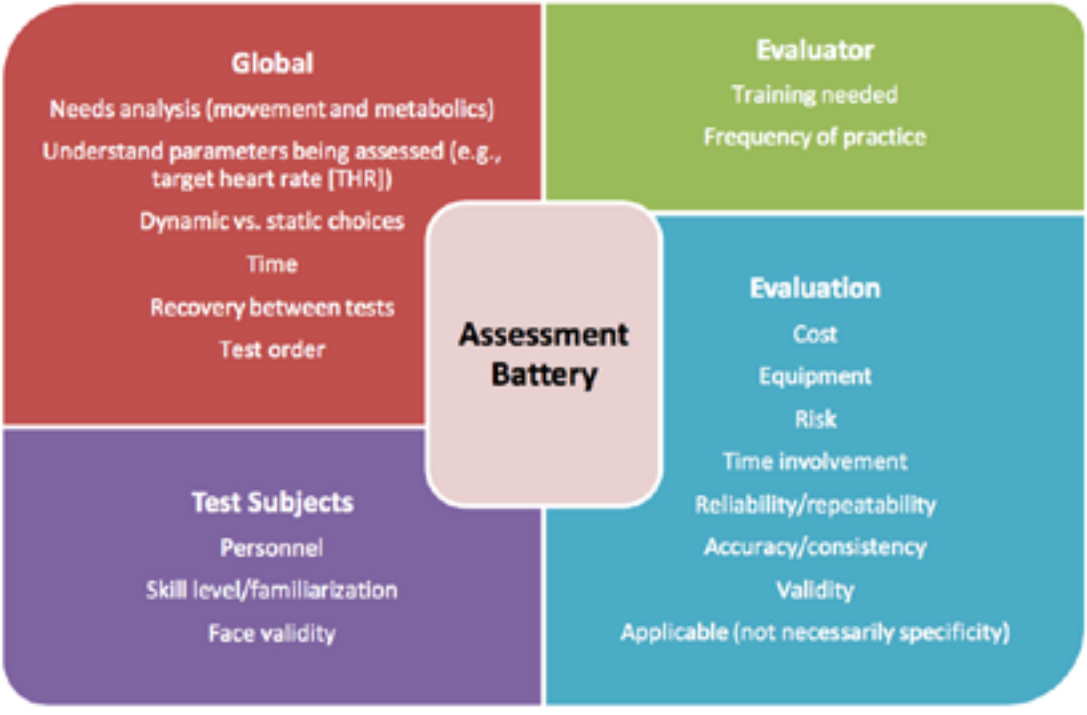
TABLE 2. ASSESSMENT BATTERY MATRIX SAMPLE

Table 2 is an excerpt of what Table 1 could look like in charted format. It demonstrates possible assessments using a matrix. While nearly any category can be justified, those with a dot (•) indicate the most globally applicable. There is little need to make it more detailed, but it can also be simplified.

	SHOULDER		KNEE	
POTENTIAL ASSESSMENTS CATEGORY	**Anterior Flexion/Horizontal Adduction**	**Posterior Extension/ Horizontal Abduction**	**Anterior Extension**	**Posterior Flexion**
Posture		•		
Flexibility/Mobility	•			•
Stability		•	•	•
Balance			•	•
Coordination				
Agility/Mobilization			•	
Strength	•		•	
Power	•	•	•	
Speed			•	•
Local Muscular Endurance	•			
Lactate Tolerance/Threshold/Recovery			•	•
Cardiovascular Endurance/Aerobic Power			•	•

TABLE 3. ASSESSMENT BATTERY DECISION MAKING CONSIDERATIONS FOR THE FIRE SERVICE

Among standard fitness battery considerations, specific factors apply to the fire service. Additional considerations must be added due to the nature and operation of the industry.



UNDERSTANDING SCIENCE AND RATIONALE

With many tactical facilitators entering the rapidly growing field of tactical strength and conditioning, it is worth consistently re-educating foundational principles of test selection, as well as practical ones. This can help leaders provide the best selections for their resources, as well as keep their personnel informed of testing measures.

MEASURE

When selecting test batteries, the facilitator needs to question whether the assessment directly measures the following:

- Performance—velocity, load, time, distance, etc.
- Physical response—heart rate, respiratory rate, etc.
- Physiological response—lactate, oxygen uptake, etc.

UTILITY OF FIELD ASSESSMENTS

Most field-usable assessments are about relationships—parallels of an activity (performance measure) to the physiological adaptations (physiological measure). The performance measure can provide an estimate of a more precise direct measure test without the complications and often cost of that criterion testing method.

These field tests show relationships with physiological capabilities of the body, and can highlight potential bottlenecks and improvement thresholds. As an example, a vertical jump, combined with other body measure information (biometrics) can provide an estimate of leg triple-extension power without directly measuring power (3,10,12). Similarly, heart rate or ratings of perceived exertion during a treadmill test can relate to anaerobic or lactate thresholds (20). Relationships are not exact, but do hold validity in what they are trying to show. When choosing assessments, results may vary between options of a certain testing/assessment mode (i.e., three different vertical jump tests), but can usually remain consistent within a qualified test.

BENEFITS AND DRAWBACKS

All assessment categories (not just specific tests) have benefits and drawbacks. While government entities may find finances as a pivotal point for test concern, even high profile professional sports have ongoing debates surrounding assessment choices. Three additional areas specific to fire service include the following:

1. **Time Involvement**: The time each test takes to administer, rest required between tests, total time of the battery, and reset time for the next firefighter can all add up quickly. Cardiovascular assessments usually consume the most time, followed by local muscular endurance (LME) tests and their recovery. Small organizations may find it worthwhile to conduct a longer 20 – 30-min cardiovascular assessment. Organizations consisting of several hundred or even several thousand personnel may want to use more efficient methods, such as a road test (open road or track run) or other approaches that shorten the test time commitment.

2. **Logistics**: The road test is often used in the military. However, it is not used because it is necessarily the best test by which to measure cardiorespiratory health, but because it is the most efficient method for the logistical variables that are at play. This test allows for the simultaneous testing of several hundred or even several thousand personnel in a day.

   The United States Marine Corps has long used an upper-body pull strength-endurance test (pull-ups). The primary reason this test is not used in most standardized military assessments is time management—hundreds of people can fit on an open road or field, but there are only so many pull-up bars and personnel to monitor. However, in the fire service, testing single or even several crews at a time (5 – 15 people on average), a pull-up is a very simple, low-cost, and easily supervised assessment that does not consume an excessive amount of time. Yet it is infrequently implemented, likely due to mixed reception and concerns by both field personnel and administrators as a passable assessment.

3. **Dynamic versus Static Tests**: Static tests have inconsistent arguments on their translation to real-life movement. Common strength and power tests in the fire service use a static maximal squat extension and a static maximal biceps curl tests. Using populations who are not consistently trained for safety and who are highly competitive can pose a significant risk. Some static exercises such as the static plank seem to have become more acceptable in therapeutic settings and subsequently, performance as an option for improving spinal musculature health and more recently noticed as an assessment preferred over most other abdominal assessments (19,21). Lumbar load may be questionable when planking and should be considered cautiously when used competitively in assessments (19). They can be related to a version of anatomical core strength endurance but are position and contraction type (i.e., isometric) specific.

SPECIFIC COMPONENT CONSIDERATIONS

Arguments can be made for any assessment category, but well-supported rationale is needed for test choices. For example, while agility (high speed changes of direction) could show relationship and relativity to balance fatigue and other performance elements, it is not directly a specific action needed by firefighters (4,17). Since it is not feasible to test every possible task that a firefighter can experience on the job, the most major influencers can provide a global picture of performance or health.

CARDIORESPIRATORY

Most cardiovascular field tests including running, walking, stepping, biking, or rowing tasks. Firefighters often question the use of this type of assessment as structural firefighters typically do none of these on the job. More importantly, the needs analysis of firefighters is absolutely absent of any ongoing, repetitive, cyclical activity of any particular gross muscle group. Considering this, these measures are important to the firefighter. They are not only health markers, but cardiorespiratory fitness is required for the ongoing elevated heart rate (a physical response measure related to physiological processes) demand in the loaded work of intense fire and rescue operations. Cardiovascular fitness is also a foundation of the metabolic system's ability for anaerobic replenishment during ongoing operations (8,9).

Many field firefighters would suggest a practical evolution. A measurable way to identify the cardiovascular health component is through an ongoing cyclical test. Stair climbing, cycling, rowing, and swimming are also relatable. These, again, are tied into a relationship and are one piece of the entire battery.

- **Graded Treadmill Protocols**: Common treadmill fitness assessments that increase speed and grade over set periods of time require one-on-one supervision and can be time intensive for fire service needs, especially if conducted on duty. The Gerkin protocol can be a useful test for the fire service; its goal is to provide a shorter assessment that is easy to administer and addresses agency time and logistical demands. The evaluation violates some quality treadmill protocol standards, such as minimum stage length and progression rate (20). Real-world experience supports evidence that the VO_2 estimation from the results are not a realistic representation across individuals; however, the numbers mean little to most people (7,14). Fit individuals who have a higher tolerance threshold may have underestimated results while those in poor condition have overestimated results. Nonetheless, practical experience suggests the results are consistent from year to year and, therefore, provide an accurate measurement for fitness improvements or decrements.
- **Lactate Testing**: Possibly one of the most accurate relationships to the "wall" that firefighters often hit during heavy operations that can lead to injury and hazards, lactate testing may initially seem appealing to organizations. The equipment is inexpensive relative to others and may be easy for small departments. However, the test is not practical for most agencies to use regularly due to time constraints and constant training for accuracy. While most firefighters are familiar with the sampling method similar to blood glucose analysis, the protocol can be sensitive to inconsistencies. Lactate sampling can be performed during the aforementioned graded treadmill protocol, another standardized graded protocol, or in practical evolutions. If chosen and applied in the right mode, lactate threshold, tolerance, and recovery can provide a picture of strength, LME, and cardiorespiratory capabilities.

MUSCULAR ASSESSMENTS

- **Muscular Strength**: Muscular strength is unarguably one of the most necessary tasks in the fire service; contemplating what to prioritize for muscular strength assessment is not as simple. Traditional multi-joint exercises (e.g., squat, chest press, etc.) are most realistic to demanding tasks, and are time efficient as opposed to completing multiple single-joint exercises. Most agencies, however, do not have enough equipment to complete these assessments, nor the time resources. These tests work well in organizations that have a culture of weight training due to the technique training required for safety. Three to five repetition assessments for the purposes involved in the fire service are as reliable as one-repetition maximum assessments while reducing injury risk and saving time (6).
- **Power**: Untrained persons often believe power is simply another term for strength. Clinically, power is work (force x displacement) measured over time, and in performance it is specific to rapid and usually explosive movements. Lighter weight than those used for strength is necessary for most explosive power-based movements (5). Heavy lifts and moves are common in the fire service, but rapid, explosive lifts are not common. Especially considering how encumbered with load firefighters are under during operations, it can be nearly impossible to conduct explosive movements. Power production is often judged with Olympic-style lifts or jumps and is related to practical power performance. It can also have a strong relationship to strength (3,10,12). Power training (e.g., power lifts, plyometrics, etc.) can improve performance and reduce injury risk (5,15). A regular fitness measure of power is not needed in the fire service; it can, however, be used as a simpler and possibly safer method relating to strength.
- **Local Muscular Endurance (LME)**: LME may be the most related to fire service demands due to the ongoing nature of operations; constant shift between body areas; very high anaerobic demands; extreme demands of intensity, volume, and time during firefighting performance; and limited recovery periods (2). LME can also adequately reflect relative strength capabilities in the upper body, and more importantly, strength endurance (repeatable strength) (6,11). It does not measure absolute strength; for example, a 160-lb individual may be able to perform 80 push-ups (LME) and bench press 225 lb (strength), while a 200-lb individual may only be able to perform 50 push-ups but bench press 300 lb. Push-ups, pull-ups, sit-ups, and their variations are the most common LME assessments. Lower body LME evaluations are less common, but are offset by the cardiorespiratory evaluations. Carry-type assessments (e.g., unilateral farmer's carry) with accountability for postural mechanics may provide excellent

subjective trunk evaluations. While there is little research on the assessment methodologies and protocols, they are applied heavily in training and applicable to fire and rescue activities.

MOBILITY

Joint health and safety can be judged not just by their force capabilities, but by their mobility. It is important for tactical facilitators to judge and relate them together as opposed to strictly independently. As an example, the traditional sit-and-reach test may or may not judge lumbar mobility depending on the technique used and individual limitations (1). This is because basic range of motion (ROM) assessments measure extensibility as opposed to pliability, and may not be important to the purpose at hand (11).

Biomechanical mat tests and other more advanced assessments require extensive mentoring and experience to learn mechanical norms and recognize pathological variances; as such they are not practical in a non-clinical environment. Movement assessments, once limited to biomechanical experts, have been popularized and more acceptable through summarized systems that make learning the skill more feasible. These evaluations can highlight ROM issues that flexibility evaluations alone may not show. Some scrutiny has been given to movement screens due to the test subject learning adaptation, subjective scoring, inter-rater consistency, and ability to predict performance outcomes (13,16,18). However, nearly all health and fitness mobility assessments are subjective, and quantification is limited by the uniqueness of every individual's anatomy and specific demands.

CONCLUSION

There is no perfect assessment battery for every circumstance. This is even more apparent for tactical athletes due to variable demands which are based not only on the specific situation, but also on jurisdictional needs, operational strategies, and governing support. Tactical facilitators must remember to not worry about all the smaller details of the precision of task and assessment type, and keep the big picture in mind. As an example, if a battery was modified several times over the course of several years, covering similar factors each time with different tests, there would likely not be much variance in the ranking of each individual compared to the group. It is more important to begin with a few simple, acceptable, evaluations than it is to begin perfectly at a much later stage. As with all other fire operations, improvements will evolve as time and philosophy provide feedback.

REFERENCES

1. Baltaci, G, Un, N, Tunay, V, Besler, A, and Gerceker, S. Comparison of three different sit and reach tests for measurement of hamstring flexibility in female university students. *British Journal of Sports Medicine* 37(1): 59-61, 2003.

2. Bennett, J. What types of fitness training do firefighters actually need? Models to adapt training preference. *TSAC Report* 42: 6-9, 2016.

3. Carlock, JM, Smith, SL, Hartman, MJ, Morris, RT, Ciroslan, DA, Pierce, KC, et al. The relationship between vertical jump power estimates and weightlifting ability: A field-test approach. *Journal of Strength and Conditioning Research* 18(3): 534-539, 2004.

4. Cortes, N, Quammen, D, Lucci, S, Greska, E, and Onate, J. A functional agility short-term fatigue protocol changes lower extremity mechanics. *Journal of Sports Sciences* 30(8): 797-805, 2012.

5. de Vos, NJ, Singh, NA, Ross, DA, Stavrinos, TM, Orr, R, and Singh, MAF. Optimal load for increasing muscle power during explosive resistance training in older adults. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 60(5): 638-647, 2005.

6. Desgorces, FD, Berthelot, G, Dietrich, G, and Testa, MSA. Local muscular endurance and prediction of the 1 repetition maximum for bench press lift in different athletic populations. *The Journal of Strength and Conditioning Research* 24(2): 394-400, 2010.

7. Dolezal, BA, Barr, D, Boland, DM, Smith, DL, and Cooper, CB. Validation of the firefighter WFI treadmill protocol for predicting VO_2max. *Occupational Medicine* 65(2): 143-146, 2015.

8. Gledhill, N, and Jamnik, VK. Characterization of the physical demands of firefighting. *Canadian Journal of Sport Sciences* 17(3): 207-213, 1992.

9. Henderson, ND, Berry, MW, and Matic, T. Field measures of strength and fitness predict firefighter performance on physically demanding tasks. *Personnel Psychology* 60(2): 431-473, 2007.

10. Impellizzeri, FM, Rampinini, E, Maffiuletti, N, and Marcora, SM. A vertical jump force test for assessing bilateral strength asymmetry in athletes. *Medicine & Science in Sports & Exercise* 39(11): 2044-2050, 2007.

11. Jones, AM. Running economy is negatively related to sit-and-reach test performance in international-standard distance runners. *International Journal of Sports Medicine* 23(1): 40-43, 2002.

12. Kraska, JM, Ramsey, MW, Haff, GG, Fethke, N, Sands, WA, Stone, ME, and Stone, MH. Relationship between strength characteristics and unweighted and weighted vertical jump height. *International Journal of Sports Physiology & Performance* 4(4): 461-473, 2009.

13. Lockie, RG, Schultz, AB, Jordan, CA, Callaghan, SJ, Jeffriess, MD, and Luczo, TM. Can selected functional movement screen assessments be used to identify movement deficiencies that could affect multidirectional speed and jump performance? *The Journal of Strength and Conditioning Research* 29(1): 195-205, 2015.

14. Mier, CM, and Gibson, AL. Evaluation of a treadmill test for predicting the aerobic capacity of firefighters. *Occupational Medicine* 54(6): 373-378, 2004.

15. Myer, GD, Ford, KR, Palumbo, JP, and Hewett, TE. Neuromuscular training improves performance and lower extremity biomechanics in female athletes. *The Journal of Strength and Conditioning Research* 19(1): 51-60, 2005.

16. Onate, JA, Dewey, T, Kollock, RO, Thomas, KS, Van Lunen, BL, DeMaio, M, and Ringleb, SI. Real-time intersession and interrater reliability of the functional movement screen. *The Journal of Strength and Conditioning Research* 26(2): 408-415, 2012.

17. Pauole, K, Madole, K, Garhammer, J, Lacourse, M, and Rozenek, R. Reliability and validity of the T-test as a measure of agility, leg power, and leg speed in college-aged men and women. *The Journal of Strength and Conditioning Research* 14(4): 443-450, 2000.

18. Shultz, R, Anderson, SC, Matheson, GO, Marcello, B, and Besier, T. Test-Retest and interrater reliability of the functional movement screen. *Journal of Athletic Training* 48(3): 331-336, 2013.

19. Snarr, RL, and Esco, MR. Electromyographical comparison of plank variations performed with and without instability devices. *The Journal of Strength and Conditioning Research* 28(11): 3298-3305, 2014.

20. Tanner, R, and Gore, C (Eds.). *Physiological Tests for Elite Athletes* (2nd ed.). Champaign, IL: Human Kinetics; 2013.

21. Tong, TK, Wu, S, and Nie, J. Sport-specific endurance plank test for evaluation of global core muscle function. *Physical Therapy in Sport* 15(1): 58-63, 2014.

22. Vaara, JP, Kyröläinen, H, Niemi, J, Ohrankämmen, O, Häkkinen, A, Kocay, S, and Häkkinen, K. Associations of maximal strength and muscular endurance test scores with cardiorespiratory fitness and body composition. *The Journal of Strength and Conditioning Research* 26(8): 2078-2086, 2012.

ABOUT THE AUTHOR

John Bennett is the co-developer and chair of the Florida Firefighter Fitness Collaborative, an organization that focuses on agency-to-agency cooperation and creates training, education, and resource networking opportunities. Additionally, he has experience in a sports performance lab, in sports and fitness businesses, and teaching collegiate sports sciences and pre-professional development. He has been honored to present for the National Strength and Conditioning Association (NSCA), United States of America Triathlon (USAT), and commercial athletic conferences. Bennett has a Master of Science degree in Exercise Physiology from the University of Central Florida and is a Certified Strength and Conditioning Specialist® (CSCS®) through the NSCA.