

Comparison of three agility tests with male servicemembers: Edgren Side Step Test, T-Test, and Illinois Agility Test

Michele A. Raya, PhD, PT, SCS, ATC;¹ Robert S. Gailey, PhD, PT;^{1-2*} Ignacio A. Gaunaud, PhD, MSPT;² LTC Daniel M. Jayne, MPT, OCS;³ MAJ (Ret) Stuart M. Campbell, MPT;⁴ Erica Gagne, DPT;¹ Patrick G. Manrique, DPT;¹ Daniel G. Muller, DPT;¹ Christen Tucker, DPT¹

¹Department of Physical Therapy, Miller School of Medicine, University of Miami, Coral Gables, FL; ²Miami Department of Veterans Affairs Healthcare System, Miami, FL; ³Womack Army Medical Center, Fort Bragg, NC; ⁴Center for the Intrepid, Brooke Army Medical Center, San Antonio, TX

Abstract—Performance-based outcomes such as the T-Test, Edgren Side Step Test (ESST), and Illinois Agility Test (IAT) have been used to assess agility in athletes and nonathletes; however, the reliability and validity of these tests have not been established. The purpose of this study was to establish the reliability and convergent construct validity of the ESST, T-Test, and IAT in young, nondisabled, physically active male servicemembers (SMs). Ninety-seven male Active Duty U.S. Army SMs completed the study. Statistically significant differences were not found between the ESST ($p = 0.10$), T-Test ($p = 0.09$), and IAT ($p = 0.23$) when administered twice within a 24 to 48 h period. These tests were found to have excellent interrater reliability and moderate to good test-retest reliability. A good positive relationship exists between the IAT and T-Test ($r = 0.76$, $p < 0.001$) and a moderate negative relationship exists between the ESST and both the T-Test ($r = -0.69$, $p < 0.001$) and IAT ($r = -0.65$, $p < 0.001$). The results suggest that these tests are valid measures of agility that uniquely assess movement in different planes, thus providing a comprehensive assessment of high-level mobility.

Key words: agility, Edgren Side Step Test, Illinois Agility Test, military, minimal detectable change, reliability, servicemembers, standard error of measurement, T-Test, validity.

INTRODUCTION

Agility has been defined as the ability to maintain a controlled body position and rapidly change direction

without a loss of balance, body control, or speed [1–2]. The components of agility have been defined as balance, coordination, power, and speed [3]. It is essential for elite, recreational, and “tactical” athletes who require the ability to rapidly change directions in all planes for sport or work. Benefits from improved agility include increased body control during fast movements, increased intramuscular coordination, and decreased risk of injury or reinjury [2,4–5]. A variety of tests have been designed to assess agility; however, few have been established as reliable or valid measures for young adult males, and the relationship between tests has not been established.

Abbreviations: AD = Active Duty, BAMC = Brooke Army Medical Center, CI = confidence interval, ESST = Edgren Side Step Test, HUC = Human Use Committee, IAT = Illinois Agility Test, ICC = interclass correlation coefficient, IRB = Institutional Review Board, MDC = minimal detectable change, SD = standard deviation, SEM = standard error of measurement, SM = servicemember, VA = Department of Veterans Affairs, WRAMC = Walter Reed Army Medical Center.

*Address all correspondence to Robert S. Gailey, PhD, PT; Department of Physical Therapy, Miller School of Medicine, University of Miami, 5915 Ponce de Leon Blvd, Plummer Bldg, 5th floor, Coral Gables, FL 33146; 305-284-4535; fax: 305-378-4107. Email: rgailey@miami.edu
<http://dx.doi.org/10.1682/JRRD.2012.05.0096>

Current measures assess one or multiple components of agility, including unidirectional, bidirectional, and multidirectional movements, which may or may not include retropulsion and rapid acceleration and deceleration transitions when changing direction [4,6]. Examples of agility tests include the Hexagon Test, 505 Agility Test, 40 yd Shuttle Run, and T-Test [5,7]. There appears to be inconsistency with regards to administration of many agility tests, including the T-Test [2,6,8–10]. Paoule et al. examined the T-Test, as described by the National Strength and Conditioning Association [11], concluding that it was a reliable and valid measure of sprinting speed, lower-body power, and agility and predictor of sports performance in nondisabled college-aged men and women [4]. The Edgren Side Step Test (ESST) assesses an individual's agility in the lateral direction [12–17]. The ESST was first introduced in the literature in 1932 by Harry D. Edgren to assess lateral mobility in basketball players; however, statistical metrics were not established. The ESST has been found to be a valid measure of off-ice lateral agility in hockey players [13–14]. Because the ESST is a popular test of agility, there is a need to establish its consistency and validity with specific groups [11]. The Illinois Agility Test (IAT) was introduced in 1942 as a test of motor ability, particularly running and dodging agility in nondisabled individuals [18–19]. Distinct characteristics of the IAT include a prone start position followed by a rapid transition to standing, followed by a combination of multidirectional maneuvers around obstacles. Excellent within-day reliability ($r = 0.98, p < 0.001$) was established with a variation of the IAT in semiprofessional rugby players, using a shorter course, unknown performance surface, and without the prone to standing start [20]. Numerous authors have described the IAT as a measure of multidirectional agility for a variety of sports without establishing the prerequisite performance metrics [2,12,20–26].

Although the ESST, T-Test, and IAT have been described in the literature, discrepancies do exist with regard to the test layout, administration, and scoring, suggesting a need to establish a standardized means of testing that is also proven to be reliable. Moreover, the question of whether multiple agility tests for comprehensive assessment are needed or a single test is sufficient because few differences exist between tests has never been addressed in the literature. In other words, would measuring movement in one or two planes suffice for agility testing or must all planes be assessed to ensure

proficiency of mobility? Finally, agility tests are commonly used in sports for athletes but rarely used in the military or in clinical situations. Because servicemembers (SMs) are largely considered tactical athletes, there is a need for a set of selected agility tests appropriate for this population in the clinical environment.

The establishment of a series of reliable and valid tests of agility capable of assessing three distinct patterns of movement in the uniplanar, biplanar, and multiplanar directions would provide a comprehensive assessment of high-level mobility. Furthermore, a performance-based set of measures enables clinicians to establish baseline high-level mobility capabilities; quantify changes in performance over time; and in the event a pre-morbid test was performed, assist in determining readiness to return to sport or activity.

The primary purpose of this study was to establish the interrater and test-retest reliability of the ESST, T-Test, and IAT in a nondisabled population of male Active Duty (AD) U.S. Army SMs and determine convergent construct validity of the ESST, T-Test, and IAT by examining the relationship between these three tests. We hypothesized that (1) the tests would present with excellent interrater and test-retest reliability when administered by two raters simultaneously and independently and when performed twice within a 24 to 48 h period, and (2) the ESST, T-Test, and IAT would prove to be valid but different measures of lower-body agility. A moderate correlation was expected between the ESST, T-Test, and IAT because they measure similar but different agility competencies. The secondary purpose of the study was to determine the standard error of measurement (SEM) and minimal detectable change (MDC) for the ESST, T-Test, and IAT for this population of nondisabled athletic male SMs.

METHODS

Subjects

A convenience sample of participants was recruited from U.S. Army base Fort Bragg, North Carolina. Individuals were included in the study if they were male, between 18 and 40 yr old, and on AD. Participants were excluded if they had a physical profile at the time of enrollment (a physical profile is defined as any condition associated with an illness, injury, or surgery that requires the SMs to be restricted from returning to full-functional

duty). Specifically, participants were asked whether they had recently experienced any increase or decrease in body weight or difficulty sleeping, in addition to other relevant health-related questions.

Participants provided demographic information regarding age, education, ethnicity, race, and weekly exercise and physical fitness routines. Anthropometric measurements such as height, weight, and waist circumference were obtained by the same investigator.

Testing Procedures

The tests were performed at an indoor gymnasium with wood flooring during the normal working day hours of 7:30 a.m. to 4:30 p.m. Water stations were located in the staging and testing area for hydration purposes. The participants were given 10 min to prepare for the tests by performing exercises from a printed dynamic warm-up protocol specially designed by military medical and exercise experts familiar with military fitness programs. Subjects were asked to repeat the same warm-up on both testing days. The order of administration of the three tests was ESST, T-Test, and IAT. This order was based on test direction, progressing from the uniplanar movement of the ESST to the biplanar movements of the T-Test and concluding with the multidirectional movements of the IAT. The participants were randomly assigned to two trained raters who simultaneously observed and independently recorded all testing results. Rater 1 read the scripted instructions for each test, while rater 2 walked the test to ensure that each participant understood the test pattern and how to perform the test but did not witness any possible movement strategies.

Raters did not verbally motivate or offer performance enhancement suggestions. Both raters timed or scored the participant's performance separately and were always blinded from each other's recording sheet. Participants were asked to perform each test twice, with the best score of the two trials selected for data analysis. In the event a participant was unable to complete a test in two trials because of a disqualification or a fall, a third trial was permitted. All participants recorded a score for all tests. A 60 s rest period was given between each trial, and a 2 min rest period was given between each test. According to Wathen and Roll, a relief ratio of 1:3 is suggested during anaerobic interval training [27]. We established a range from 1:3 to 1:6 between trials (60 s) and a range from 1:6 to 1:12 between tests (120 s). The entire testing session, including rest periods, was less than 9 min. This

amount of physical exertion fell well below the 60 to 90 min of daily physical training customarily performed by the SMs [28]. A rest period of 24 to 48 h between trials was also consistent with the training regimen of AD SMs [28]. Participants were tested twice within a 24 to 48 h period by the same rater team.

Outcome Measures

Prior to testing, all tests were reviewed to identify any variations found with respect to test description and administration found within the literature. The team met to review all discrepancies and to standardize the testing procedures, including course layout, specific verbal instruction scripts, and scoring. The following description of each test discusses the modifications that were introduced with the goal of standardization of tests while maintaining the original authors' intent for each test.

Edgren Side Step Test

The ESST was administered using a version standardized from previous literature [8–9,11]. The units of measurement were changed from feet to meters, resulting in a slightly longer course of 4 m in length with four 1 m (3.28 ft) increments. The starting position was changed from the center cone to the far left cone in order to standardize the starting position, and each 1 m increment was marked with a strip of tape and a cone, which allowed for easier scoring of the test (**Figure 1**). The raters were positioned in front of and behind the participant. The participant began in a standing position behind the far left cone and was instructed not to cross his feet while sidestepping. On the "go" command, the participant sidestepped to the right until his right foot had touched or crossed the outside cone or tape mark. The participant then sidestepped to the left until his left foot had touched or crossed the left outside cone or tape mark. The participant sidestepped back and forth to the outside cones as rapidly as possible for 10 s. The participant is given one point per completion of each 1 m increment marked by a cone or tape mark. If the far end lines were not reached, those points were not awarded. A subject was given a score of 0 if he failed to keep his trunk and feet pointed forward at all times, crossed his legs, or did not complete the course successfully.

T-Test

The T-Test was administered using a version standardized from previous literature [2,11]. The units of measurement were changed from yards to meters, creating a 10 ×



Figure 1.
Edgren Side Step Test.

10 m course. The course procedure of having the participant touch each cone is not standardized in the literature; therefore, the task was eliminated. The directions adopted for this study were based on Miller et al. [2]. On the “go” command, the participant (1) ran or moved as quickly as possible forward to the center cone, (2) sidestepped to the right 5 m to the right cone, (3) sidestepped to the left 10 m to the far left cone, and then (1) sidestepped back to the right to the center cone (**Figure 2**). The participant then ran or moved backward as quickly as possible to cross the finish line. The raters began the stop watch on “go” and stopped when the participant broke the plane of the finish line. The time to complete each trial was recorded in seconds. Disqualification was determined if the participant failed to run the course as instructed, failed to reach the finish line or complete the course, moved any cones, did not keep his trunk and feet pointed forward at all times, or crossed his legs more than once when sidestepping. If a participant did not complete a trial successfully, a score of 0 was given.

Illinois Agility Test

The IAT was administered using a version standardized from previous literature [1,18–19,23–24]. The length of the IAT was originally set at 30 ft, which was increased slightly to 10 m for ease of test administration. The IAT course was marked by cones, with four center cones spaced 3.3 m apart and four corner cones positioned 2.5 m from the center cones (**Figure 3**). The participant began the test lying prone on the floor behind the starting line with his arms at his side and his head turned to the side or facing forward. On the “go” command, the participant ascended to his feet and ran or moved quickly forward to the first tape mark. Participants were required to touch or cross the tape mark with their foot. The participant turned around and moved back to the first center cone, where he weaved up and back through the four center cones. The participant then ran or moved as quickly as

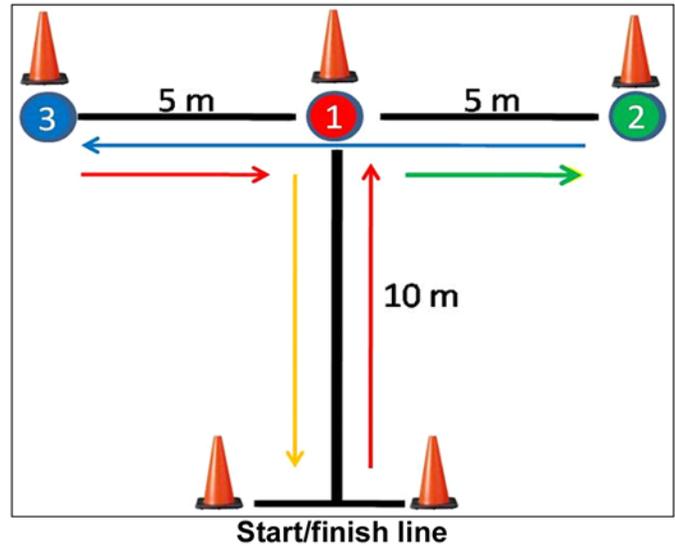


Figure 2.
T-Test.

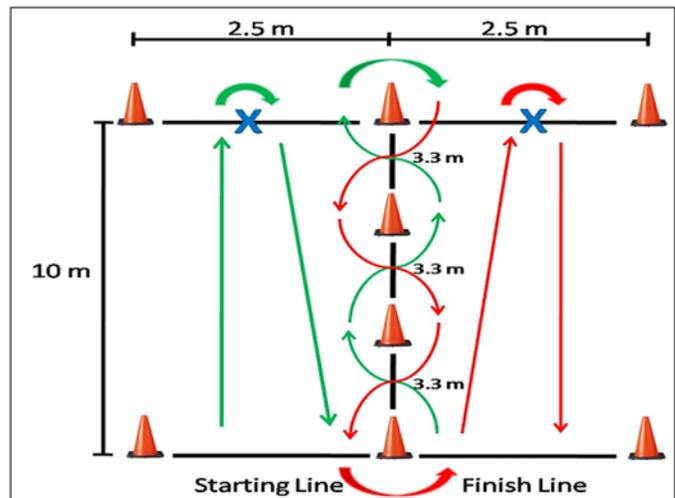


Figure 3.
Illinois Agility Test.

possible to the second tape mark on the far line. Again, participants were required to touch or cross the end-line tape marks with their foot. Lastly, the participant turned around and ran or moved as quickly as possible across the finish line. The time to complete each trial was recorded in seconds. Disqualification was determined if the participant failed to run the course as instructed, failed to reach the end lines, failed to complete the course, or moved any cones.

Data Analysis

Data was entered into Excel worksheets (Microsoft; Redmond, Washington), and analysis was performed on SAS version 9.13 (SAS Institute Inc; Cary, North Carolina). Descriptive statistics were used to characterize the study participants via their obtained demographic and anthropometric information. The best performance of the participants was chosen from among the completed trials for each test using rater 1 data from sessions 1 and 2. This was used to determine the mean \pm standard deviation (SD) and range of scores for each test. Student paired *t*-test was used to determine whether test familiarity occurred when participants performed the agility outcome measures twice within a 24 to 48 h period. Rater 1 data for sessions 1 and 2 were used to determine whether significant differences were found between performances of the agility tests ($p \leq 0.05$).

The interclass correlation coefficient (ICC) was calculated to determine the interrater and test-retest reliability for the ESST, T-Test, and IAT using Model 2.1. We reported 95 percent confidence intervals (CIs) for all ICC values. For interrater reliability, comparison of agreement was analyzed between raters 1 and 2 for session 1. For test-retest reliability, comparison of agreement was analyzed between sessions 1 and 2 for rater 1 of each rater team.

The SEM is defined as the determination of the amount of variance or spread in the measurement error of a test [29]. The SEM for each of the agility tests was calculated using the test-retest reliability coefficient (*r*) and SD of the agility tests for rater 1, session 2 data for the study participant (**Equation (1)**):

$$\text{SEM} = \text{SD} \sqrt{1-r} \quad (1)$$

The MDC is a statistical approach to determine the responsiveness of a test. The MDC for the ESST, T-Test, and IAT was calculated using the following formula (**Equation (2)**):

$$\text{MDC} = z\text{-score}_{\text{level of confidence}} \times \text{SD} \times \sqrt{(2[1-r_{\text{test-retest}}])} \quad (2)$$

The *z*-score represents the CI from the normal distribution. We used 95 percent CI for the study population, making the *z*-score = 1.96. The SD was used from session 2 data. The ICC was used from the test-retest reliability results. A multiplier of the $\sqrt{2}$ was used to account for the increased error in calculating the test-retest reliability with two different data sets at two different points in time [30–31].

The Pearson product-moment correlation coefficient was used to determine the convergent construct validity between ESST distance, T-Test time, and IAT time. Fair degree of relationship is an *r*-value between 0.26 and 0.50. A moderate to good degree of relationship is an *r*-value between 0.51 and 0.75. A good to excellent degree of relationship is an *r*-value between 0.76 and 1.00.

RESULTS

A convenience sample of 97 participants completed both testing sessions. **Table 1** describes the baseline characteristics of the study participants. The mean \pm SD age, height, weight, and number of days trained each week were 26.2 ± 5.5 years, 177.92 ± 6.47 cm, 82.26 ± 10.54 kg, and 5 ± 0.8 days, respectively. Approximately half of the participants were married (51%) and a greater number had completed at least 1 yr of college (60%) and were white/Caucasian (80%). The majority of the participants were enlisted soldiers (91%) with the minority being warrant and commissioned officers (9%).

Table 2 describes the mean results for ESST distance, T-Test time, and IAT time. Significant differences were not found between agility test performance when comparing sessions 1 and 2. These results suggest that test familiarity or a learning effect did not occur when these participants performed these agility tests twice within a 24 to 48 h period. Of the 97 participants tested, 92 performed the tests twice within a 48 h period. Five participants, due to a conflict in scheduling, rested 24 h before performing the tests the second time. These five participants were part of a graduating Special Forces class.

The ICC values for interrater reliability for the agility tests ranged from 0.92 for the ESST to 0.99 for the IAT. The ICC for test-retest reliability of the agility tests ranged from 0.62 for the ESST to 0.83 for the T-Test. (**Table 3**)

The SEMs for the ESST, T-Test, and IAT were determined using each measure's test-retest reliability coefficient (**Table 3**) and SD (**Table 2**, session 2 results). The SEMs for the ESST, T-Test, and IAT were 1.41 m, 0.40 s, and 0.65 s, respectively. The MDCs for the ESST, T-Test, and IAT were 3.91 m, 1.10 s, and 1.80 s, respectively (**Table 4**).

The Pearson product-moment correlation coefficient indicated that a good positive relationship existed between T-Test and IAT performance ($r = 0.75$, $p < 0.001$) and a moderate negative relationship existed between T-

Table 1.
Characteristics of Active Duty U.S. Army servicemembers.

Characteristic	<i>n</i> (%)	Mean ± SD	Range
Age (yr)	97 (100)	26.2 ± 5.5	18–39
Height (cm)	97 (100)	177.92 ± 6.47	163.83–193.04
Weight (kg)	97 (100)	82.26 ± 10.54	56.25–109.77
Waist Circumference (cm)	97 (100)	86.10 ± 6.97	69.22–104.14
Days of Regular Fitness (<i>n</i>)	97 (100)	5.0 ± 0.8	3–7
Marital Status			
Single	45 (46)	—	—
Married	49 (51)	—	—
Divorced	3 (3)	—	—
Education			
High School Graduate or Equivalent	39 (40)	—	—
1–3 Years of College	39 (41)	—	—
Graduate with 4-Year Degree	13 (13)	—	—
Some Graduate School	6 (6)	—	—
Race/Ethnicity			
White/Caucasian	78 (81)	—	—
Black/African American	12 (12)	—	—
Asian	4 (4)	—	—
Native Hawaiian or Pacific Islander	3 (3)	—	—

SD = standard deviation.

Table 2.
Description of best performance for Edgren Side Step Test (ESST), T-Test, and Illinois Agility Test (IAT) for Active Duty U.S. Army servicemembers from rater 1 (*n* = 97).

Score	ESST (m)			T-Test (s)			IAT (s)		
	Session 1	Session 2	<i>p</i> -Value	Session 1	Session 2	<i>p</i> -Value	Session 1	Session 2	<i>p</i> -Value
Mean ± SD	23.89 ± 2.85	24.27 ± 2.29	0.10	12.27 ± 0.91	12.19 ± 0.97	0.09	18.26 ± 1.04	18.18 ± 1.14	0.23
Minimum	15	20	—	15.53	15.01	—	16.28	23.37	—
Maximum	31	30	—	10.82	10.13	—	22.17	15.09	—

SD = standard deviation.

Table 3.
Reliability of Edgren Side Step Test, T-Test, and Illinois Agility Test for Active Duty U.S. Army servicemembers (*n* = 97) by intraclass correlation coefficient (*r*) and 95 percent confidence interval (95% CI).

Test	Interrater Reliability, <i>r</i> (95% CI)	Test-Retest Reliability, <i>r</i> (95% CI)
Edgren Side Step Test	0.92 (0.88–0.94)	0.62 (0.48–0.73)
T-Test	0.98 (0.97–0.99)	0.83 (0.75–0.88)
Illinois Agility Test	0.99 (0.98–0.99)	0.68 (0.55–0.78)

Test and ESST performance ($r = -0.69$, $p < 0.001$). The ESST and IAT demonstrated a moderate negative relationship ($r = -0.65$, $p < 0.001$). As expected, faster T-Test times were associated with faster IAT times. In addition, faster T-Test and IAT times were associated with greater distance covered laterally in the ESST.

DISCUSSION

The purpose of this study was to establish a standardized means of testing the ESST, T-Test, and IAT as reliable measures of agility in a nondisabled population of SMs and to determine the need for multiple tests when

Table 4. Standard error of measurement (SEM) and minimal detectable change (MDC) for Edgren Side Step Test, T-Test, and Illinois Agility Test.

Test	SEM	MDC
Edgren Side Step Test (m)	1.41	3.91
T-Test (s)	0.40	1.10
Illinois Agility Test (s)	0.65	1.80

assessing agility or high-level mobility. Many different agility tests are used as clinical tools to measure anticipation, pattern recognition, and change of direction and speed [16]. The series of agility tests specifically selected for this study were low-cost, clinically friendly, and each assessed different agility movement characteristics. The ESST assesses frontal plane agility by measuring the amount of distance covered within a 4 m course in 10 s. The T-Test assesses biplanar (frontal and sagittal) agility by assessing how quickly an individual completes a course shaped as a “T” that includes forward, backward, and lateral movements and four changes of direction over a 40 m course. The IAT assess multiplanar agility by assessing how quickly an individual is able to complete a course consisting of a prone to standing transfer, multiple forward movements, three 180° turns, and forward weaving through stationary cones over a 60 m course.

The study results found that significant differences did not exist between ESST, T-Test, and IAT performance when administered twice in a single session and repeated a second time in a 24 to 48 h period, indicating that a learning effect did not occur with this population. All participants were read precise verbal instructions by rater 1, were shown a picture of the course, and had a chance to view rater 2 walk each test to ensure that they understood the test pattern without gaining an advantage for potential movement strategies. We did not include a practice run or test run of the tests because we wanted to eliminate accumulative effects of fatigue from becoming a factor on later trials. Therefore, we felt the tests were explained thoroughly and were simple to follow, making test familiarity a nonconcern. Additionally, there were no differences found between the first and second trials on day 1 or between days 1 and 2, eliminating the concern of test familiarity.

The interrater reliability of the ESST, T-Test, and IAT was excellent, suggesting that different clinicians, with a standardized set of written verbal instructions and consistent training, can administer these tests with confidence that their scores are precise and accurate. The implemen-

tations of the tests require simple equipment such as cones, tape, stopwatch, and counter. However, we found that the instructions, implementation, and interpretation of these tests were not consistent in the literature. The standardization of the test instructions and training of the raters are common practice with all performance-based outcome measures and, in conjunction with the design of each test, most likely contributed to the high inter-rater reliability.

The test-retest reliability of ESST, T-Test, and IAT was moderate to good for this population, indicating that all three tests were stable when no appreciable change over time has occurred. It is highly unlikely that an appreciable change in function could have occurred between a 24 to 48 h period in this group of nondisabled active males. The tester team, procedures, time of day, and environment remained consistent for each participant on both testing days. Fatigue was an unlikely source of error considering this population consisted of young, fit, highly trained SMs who perform physical training at least 5 d/wk for 60 to 90 min [28]. Personal factors such as motivation to improve performance, which is common in this highly competitive population, were not accounted for and may have affected the test-retest reliability.

The SEM for the ESST, T-Test, and IAT was small because of the homogeneity of the group with respect to age, sex, fitness level, and physical occupational demands. The responsiveness or MDC of the ESST, T-Test, and IAT represents the minimum amount of performance improvement required to detect or demonstrate true change in agility capabilities for nondisabled fit males between the ages of 18 and 40 years. These results provide both a criterion and numeric goal for this population who are interested in measuring or enhancing agility for sport or tactical performance as well as quantifying changes in the clinic.

The moderate correlation between the ESST, T-Test, and IAT was expected because all three tests measure the construct of agility with components such as balance, posture, coordination, power, and speed. If two or all three tests' *r*-values were greater than 0.90, we would have to consider the value of having a person perform more than one test of similar construct. But since each test assesses a different aspect of the construct of agility with regard to direction, distance, timing, and planes of motion, the assumption would be that each test evaluates different muscle groups and movement strategies. As a result, each test could be considered a unique stand-alone

measure of agility, or when performed together as a series, the combination of performance could be considered a total assessment of higher-level mobility.

For the purposes of clinical application, the order in which the clinician administers the agility tests could follow the suggested rehabilitation protocol that sequentially advances toward high-level activity and athletic performance [32–33]. Sherry and Best found that a rehabilitation protocol progressing from frontal and transverse plan motion in the first phase of healing to transverse and sagittal movements at velocities and intensities similar to competition during the second phase of healing was more effective to returning injured athletes to sport after a hamstring injury [34]. Therefore, we believe that our test sequence where the movement direction progresses from frontal plane to sagittal and frontal plane and concludes with sagittal and transverse plane movements promotes a comprehensive assessment of agility without redundancy. The ordering of agility tasks may be of interest in future studies.

The homogeneity of our subject population could be considered a study limitation. Future studies should include female SMs. Another study examining athletes from a variety of sports to determine the value of these tests with respect to the sport, position, and possible training and conditioning programs would be of interest to players and coaches. Likewise, sports medicine clinicians and trainers could use these tests to quantify recovery after injury in athletes if a baseline performance has been established. The results of this study may also help with goal-setting and treatment approaches that more closely relate to a specific sport or activity for the tactical athlete during the rehabilitation intervention. In addition, we believe these tests will provide clinicians with a battery of outcome measures that assess agility in different planes of motion that could help to determine readiness to return to a particular sport or activity. These tests could establish benchmarks of agility for the SM or athlete and help identify where deficits in movements in different planes of motion may exist as compared with their peers in order to help guide their strength and conditioning program. Lastly, future research should examine the use of these reliable agility outcome measures as clinical tools to help determine return to AD and/or return to sport following the rehabilitation of a lower-limb injury.

CONCLUSIONS

To date, the T-Test has been one of the few performance-based functional measures in sports medicine with established psychometric properties. We were able to establish the reliability of the T-Test, ESST, and IAT in a population of male SMs by developing standardized instructions and implementation methods. The T-Test, ESST, and IAT were found to be reliable measures of agility in nondisabled fit AD male SMs between the ages of 18 and 40 yr. Establishing the reliability of these additional measures provides the clinician with a clinical tool that assesses uniplanar, biplanar, and multiplanar movements, thus creating a comprehensive assessment of agility. Finally, the measurable values presented for each test may be used as reference data for clinicians and other fitness professionals who work with this population.

ACKNOWLEDGMENTS

Author Contributions:

Study concept and design: M. A. Raya, R. S. Gailey, I. A. Gaunaurd.

Acquisition of data: S. M. Campbell, D. M. Jayne, E. Gagne, P. G. Manrique, D. G. Muller, C. Tucker.

Analysis and interpretation of data: M. A. Raya, R. S. Gailey, I. A. Gaunaurd.

Drafting of manuscript: M. A. Raya, R. S. Gailey, I. A. Gaunaurd, E. Gagne, P. G. Manrique, D. G. Muller, C. Tucker.

Critical revision of manuscript for important intellectual content: M. A. Raya, R. S. Gailey, I. A. Gaunaurd.

Statistical analysis: I. A. Gaunaurd.

Study supervision: R. S. Gailey, I. A. Gaunaurd.

Financial Disclosure: The authors have declared that no competing interests exist.

Funding/Support: This material was based on work supported by the Military Amputee Research Program and Telemedicine and Advanced Technology Research Center (Prime Award No. W81XWH-06-2-0073). The U.S. Army Medical Research Acquisition Activity is the awarding and administering acquisition office. It was administered by the Henry M. Jackson Foundation for the Advancement of Military Medicine and the South Florida Veterans Affairs Foundation for Research and Education.

Additional Contributions: The authors would like to thank COL (Ret) Charles Scoville, PT, DPT; Alison A. Linberg, DPT, ATC; SGM (Ret) Brad Halling, CP; Carlos Gomez-Orozco; Orlando Gomez-Marin, MSc, PhD; Kathryn E. Roach, PhD, PT; Erick Harada, DPT; Emily Lo, DPT; Megan Mannikko, DPT; Jeremiah Randall, DPT, ATC; Emily Roberts, DPT; and Jarrod Schechla, DPT, for their dedication and countless hours of work contributed to make this project a success. The authors also thank the staff at the Frederick Physical Fitness Center and Womack Army Medical Center at Fort Bragg, North Carolina; Center for the Intrepid at Brooke Army Medical Center

(BAMC), San Antonio, Texas; Military Advanced Training Center at Walter Reed Army Medical Center (WRAMC), Bethesda, Maryland; Miami Department of Veterans Affairs (VA) Healthcare System; and South Florida Veterans Affairs Foundation for Research and Education for their generous support of this project. Mr. Muller is now with the Department of Physical Medicine and Rehabilitation, Charlie Norwood VA Medical Center, Augusta, Georgia; Ms. Tucker is now with the Department of Physical Therapy, Novacare Rehabilitation, Philadelphia, Pennsylvania; Mr. Manrique is now with ATI Physical Therapy, Chicago, Illinois, and a physical therapist with the Minnesota National Guard; and Ms. Gagne is now with the Specialized Manual Orthopedic Clinic at Holy Cross Orthopedic Institute, Ft. Lauderdale, Florida.

Institutional Review: Approved by the Womack Army Medical Center Institutional Review Board (IRB), WRAMC Department of Clinical Investigation, WRAMC Human Use Committee (HUC), BAMC Department of Clinical Investigation, BAMC HUC, Army Clinical Investigation Regulatory Office, and Miami VA Healthcare System Human Studies Subcommittee. A research investigator reviewed and explained to each participant the eligibility criteria, methodology, confidentiality, and potential risks involved in study participation. The participants' questions were asked prior to signing IRB-approved informed consent and protected health information form.

Participant Follow-Up: The authors plan to inform participants of the publication of this study.

Disclaimer: The views expressed in this manuscript are those of the authors and do not reflect the official policy or position of the Department of the Army, Department of the Navy, Department of Defense, or U.S. Government. The content of the information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

- Roozen M. Action-reaction: Illinois Agility Test. *NSCA Perform Training J*. 2008;3(5):5–6.
- Miller MG, Herniman JJ, Ricard MD, Cheatham CC, Michael TJ. The effects of a 6-week plyometric training program on agility. *J Sports Sci Med*. 2006;5(3):459–65.
- Angeli C. Training for agility and balance. In: Nyland J, editor. *Clinical decisions in therapeutic exercise: Planning and implementation*. Upper Saddle River (NJ): Pearson Education; 2006. p. 360–61.
- Paule K, Madole K, Garhammer J, Lacourse M, 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. *J Strength Cond Res*. 2000;14(4):443–50.
- Gabbett TJ. Physiological characteristics of junior and senior rugby league players. *Br J Sports Med*. 2002;36(5):334–39. [PMID:12351330] <http://dx.doi.org/10.1136/bjism.36.5.334>
- Sassi RH, Dardouri W, Yahmed MH, Gmada N, Mahfoudhi ME, Gharbi Z. Relative and absolute reliability of a modified agility T-test and its relationship with vertical jump and straight sprint. *J Strength Cond Res*. 2009;23(6):1644–51. [PMID:19675502] <http://dx.doi.org/10.1519/JSC.0b013e3181b425d2>
- Ortiz A, Olson SL, Roddey TS, Morales J. Reliability of selected physical performance tests in young adult women. *J Strength Cond Res*. 2005;19(1):39–44. [PMID:15705043]
- Harman E, Pandorf C. Principles of test selection and administration. In: Baechle TR, Earle RW, editors. *Essentials of strength training and conditioning*. 2nd ed. Champaign (IL): Human Kinetics; 2000. p. 275–311.
- Chaouachi A, Brughelli M, Chamari K, Levin GT, Ben Abdelkrim N, Laurencelle L, Castagna C. Lower limb maximal dynamic strength and agility determinants in elite basketball players. *J Strength Cond Res*. 2009;23(5):1570–77. [PMID:19620905] <http://dx.doi.org/10.1519/JSC.0b013e3181a4e7f0>
- McMillian DJ, Moore JH, Hatler BS, Taylor DC. Dynamic vs. static-stretching warm up: The effect on power and agility performance. *J Strength Cond Res*. 2006;20(3):492–99. [PMID:16937960]
- Semenick D. Testing protocols and procedures. In: Baechle T, editor. *Essentials of strength training and conditioning*. 1st ed. Champaign (IL): Human Kinetics; 1994. p. 258–73.
- Pearson DT, Naughton GA, Torode M. Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports. *J Sci Med Sport*. 2006;9(4):277–87. [PMID:16844415] <http://dx.doi.org/10.1016/j.jsams.2006.05.020>
- Farlinger CM, Kruisselbrink LD, Fowles JR. Relationships to skating performance in competitive hockey players. *J Strength Cond Res*. 2007;21(3):915–22. [PMID:17685681]
- Farlinger CM, Fowles JR. The effect of sequence of skating-specific training on skating performance. *Int J Sports Physiol Perform*. 2008;3(2):185–98. [PMID:19208927]
- Edgren HD. An experiment in the testing of ability and progress in basketball. *Res Q*. 1932;3(1):159–71.
- Sheppard JM, Young WB. Agility literature review: classifications, training and testing. *J Sports Sci*. 2006;24(9):919–32. [PMID:16882626] <http://dx.doi.org/10.1080/02640410500457109>
- Sheppard JM, Young WB, Doyle TL, Sheppard TA, Newton RU. An evaluation of a new test of reactive agility and its relationship to sprint speed and change of direction speed. *J Sci Med Sport*. 2006;9(4):342–49. [PMID:16844413] <http://dx.doi.org/10.1016/j.jsams.2006.05.019>
- Cureton TK. *Physical fitness workbook: Fit for democracy—Fit to fight*. Champaign (IL): Stipes; 1942.
- Cureton T. General motor fitness characteristics and strength of champions. In: *Physical fitness of champion*

- athletes. Urbana (IL): University of Illinois Press; 1951. p. 67–69.
20. Gabbett TJ. Influence of physiological characteristics on selection in a semi-professional first grade rugby league team: A case study. *J Sports Sci.* 2002;20(5):399–405. [PMID:12043829] <http://dx.doi.org/10.1080/026404102317366654>
 21. Raven PB, Gettman LR, Pollock ML, Cooper KH. A physiological evaluation of professional soccer players. *Br J Sports Med.* 1976;10(4):209–16. [PMID:1009297] <http://dx.doi.org/10.1136/bjism.10.4.209>
 22. Keogh JW, Weber CL, Dalton CT. Evaluation of anthropometric, physiological, and skill-related tests for talent identification in female field hockey. *Can J Appl Physiol.* 2003;28(3):397–409. [PMID:12955867] <http://dx.doi.org/10.1139/h03-029>
 23. Wilkinson M, Leedale-Brown D, Winter EM. Validity of a squash-specific test of change-of-direction speed. *Int J Sports Physiol Perform.* 2009;4(2):176–85. [PMID:19567921]
 24. Caldwell BP, Peters DM. Seasonal variation in physiological fitness of a semiprofessional soccer team. *J Strength Cond Res.* 2009;23(5):1370–77. [PMID:19620929] <http://dx.doi.org/10.1519/JSC.0b013e3181a4e82f>
 25. Vescovi JD, Brown TD, Murray TM. Descriptive characteristics of NCAA division I women lacrosse players. *J Sci Med Sport.* 2007;10(5):334–40. [PMID:16962826] <http://dx.doi.org/10.1016/j.jsams.2006.07.010>
 26. Vescovi JD, McGuigan MR. Relationships between sprinting, agility, and jump ability in female athletes. *J Sports Sci.* 2008;26(1):97–107. [PMID:17852692] <http://dx.doi.org/10.1080/02640410701348644>
 27. Wathen D, Roll F. Training methods and modes. Baechle T, editor. *Essentials of strength training and conditioning.* 1st ed. Champaign (IL): Human Kinetics; 1994. p. 403–14.
 28. U.S. Department of the Army. *U.S. Army fitness training handbook: The official U.S. Army physical readiness training manual.* Huntsville (AL): Red Bike; 2010. p. 31–49.
 29. Harvill LM. Standard error of measurement. *Educ Meas.* 2005;10(2):181–89.
 30. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther.* 2006;86(5):735–43. [PMID:16649896]
 31. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. *J Clin Epidemiol.* 2001;54(12):1204–17. [PMID:11750189] [http://dx.doi.org/10.1016/S0895-4356\(01\)00407-3](http://dx.doi.org/10.1016/S0895-4356(01)00407-3)
 32. Beam JW. Rehabilitation including sport-specific functional progression for the competitive athlete. *J Bodyw Mov Ther.* 2002;6:205–19.
 33. Kegerreis S. The construction and implementation of functional progressions as a component of athletic rehabilitation. *J Orthop Sports Phys Ther.* 1983;5(1):14–19. [PMID:18806433]
 34. Sherry MA, Best TM. A comparison of two rehabilitation programs in the treatment of acute hamstring strains. *J Orthop Sports Phys Ther.* 2004;34(3):116–25.

Submitted for publication May 30, 2012. Accepted in revised form November 16, 2012.

This article and any supplementary material should be cited as follows:

Raya MA, Gailey RS, Gaunaurd IA, Jayne DM, Campbell SM, Gagne E, Manrique PG, Muller DG, Tucker C. Comparison of three agility tests with male servicemembers: Edgren Side Step Test, T-Test, and Illinois Agility Test. *J Rehabil Res Dev.* 2013;50(7):951–60. <http://dx.doi.org/10.1682/JRRD.2012.05.0096>

ResearcherID/ORCID: Robert S. Gailey, PhD, PT: I-3458-2013; Ignacio A. Gaunaurd, PhD, MSPT: I-3490-2013

