
Nathan Millikan, Dustin R. Grooms, Brett Hoffman, and Janet E. Simon

Context: Functional tests are limited primarily by measuring only physical performance. However, athletes often multitask, and deal with complex visual-spatial processing while being engaged in physical activity. Objective: To present the development and reliability of 4 new neurocognitive single-leg hop tests that provide more ecological validity to test sport activity demands than previous functional return to sport testing. Design: Cross-sectional. Setting: Gymnasium. Participants: Twenty-two healthy participants (9 males and 13 females; 20.9 [2.5] y, 171.2 [11.7] cm, 70.3 [11.0] kg) were recruited. Interventions: Maximum distance (physical performance) and reaction time (cognitive performance) were measured for 3 of the neurocognitive hop tests all testing a different aspect of neurocognition (single-leg central-reaction hop—reaction time to 1 central stimulus, single-leg peripheral-reaction crossover hop—reaction time between 2 peripheral stimuli, and single-leg memory triple hop—reaction time to memorized stimulus with distractor stimuli). Fastest time (physical performance) and reaction time (cognitive performance) were measured for the fourth neurocognitive hop test (single-leg pursuit 6m hop—requiring visual field tracking [pursuit] and spatial navigation). Main Outcome Measures: Intraclass correlation coefficients were calculated to assess reliability of the 4 new hop tests. Additionally, Bland–Altman plots and 1-sample t tests were conducted for each single-leg neurocognitive hop to evaluate any systematic changes. Results: Intraclass correlation coefficients based on day 1 and day 2 scores ranged from .87 to .98 for both legs for physical and cognitive performance. The Bland–Altman plots and 1-sample t tests (P >.05) indicated that all 4 single-leg neurocognitive hop tests did not change systematically. Conclusions: These data provide evidence that a neurocognitive component can be added to the traditional single-leg hop tests to provide a more ecologically valid test that incorporates the integration of physical and cognitive function for return to sport. The test–retest reliability of the 4 new neurocognitive hop tests is highly reliable and does not change systematically.

Keywords: functional performance, cognitive, visuomotor, FitLight

There is a high incidence of lower-extremity injuries in many sports. Ankle sprains, muscle strains, and ligamentous damage to the knee are all common injuries sustained in sports such as football, basketball, and soccer. Hootman et al. found that more than 50% of all injuries were sustained to the lower-extremity for both practices and games, with the knee and ankle accounting for the majority of those injuries. These injuries can have arduous recovery periods, involving rehabilitation and in some cases surgical reconstruction or repair. However, the risk for reinjury or another musculoskeletal injury is exponentially higher after a primary injury, highlighting the need for functional testing to determine injury recovery and return to play readiness to mitigate the high reinjury risk.

Functional testing comes in numerous forms. Lower-extremity functional testing has focused on agility runs, vertical jumps, hop tests, and balance tests. This form of testing provides an indication of an athlete’s physical capability regarding rehabilitation progress or return to play readiness. However, isolated physical performance is only 1 aspect of function that must be restored after musculoskeletal injury. An individual may be able to hop with adequate distance or symmetry, but it is unclear whether they have fully recovered all potential deficits and are prepared for return to training and competition. Currently, functional tests are limited primarily measuring only physical performance. However, athletes often multitask, divide their attention, and deal with complex visual-spatial processing while engaged in physical performances. Additionally, sport activities rarely allow full attention to the physical task, they typically require visual attention to the environment and challenge cognition in addition to physical abilities. Several laboratory-based biomechanical reports have confirmed increased injury risk movement strategies when under cognitive or visual load, however, typical functional testing does not incorporate such visual or cognitive aspects that increase injury risk.

Hop testing is one of the most common assessments to determine return to play readiness. Four foundational hop tests have existed in the literature since the 1990s. Noyes et al. documented that the 4 traditional hops are sensitive and specific to postinjury deficits. Reliability has also been established for all 4 hop tests by multiple authors. However, these 4 hop tests may not adequately capture injury risk related to neurocognitive deficits or the neurophysiologic dysfunction that persists after injury due to only assessing physical capability. Neurocognitive performance can include various constructs including visual attention, self-monitoring, fine motor performance, dual tasking, and reaction time. For the purposes of this study, we focused on reaction time under various cognitive stresses (working memory, visual field scanning, response inhibition, and peripheral vs central visual field processing) as such measures captured in isolation (no...
physical performance component) have been linked to injury risk.26–28 Specifically, individuals with slower reaction time, processing speed, and visual memory were associated with twice the risk of sustaining a lower-extremity injury.28 To capture a broad spectrum of neurocognition, each of the 4 hop tests challenge a specific aspect of neurocognitive processing within the reaction time measure, from simple response time (central-reaction hop), go-no-go response inhibition with varied distractors and working memory (memory hop), go-no-go with 1 distractor and peripheral visual processing (peripheral-reaction hop), and visual field awareness, scanning and timing (pursuit hop). Rarely has neurocognition been incorporated into functional test protocols with standardized measures beyond the research laboratory.15,29 There is a need for clinical tests that integrate neurocognitive and physical function to better simulate athletic challenges that lead to injury and thereby improve return to play readiness testing. Such athletic challenges include tracking opponents, processing the situation to determine when to execute an action, and keeping the peripheral visual field engaged during motor performance. However, these sport-specific abilities have traditionally been difficult to objectively quantify, therefore, the purpose of this study was to: (1) establish 4 new neurocognitive single-leg hop tests based on the 4 original single-leg hop tests proposed by Noyes et al8 who incorporate measurable reaction time components that mimic sport demands and (2) determine the test–retest reliability of the new neurocognitive hop tests.

Methods

Design

This study is a cross-sectional study investigating the reliability of newly developed single-leg neurocognitive hop tests. A power analysis was conducted using the approach from Walter et al30 with 2 observations (day 1 and day 2), significance at .05, power 0.8, acceptable reliability of 0.8 and expected reliability of 0.9. The sample size needed is 18. Expected reliability was based on Ross et al31 who demonstrated hop testing is highly reliable.

Participants

Nine male and 13 female (20.9 [2.5] y, 171.2 [11.7] cm, 70.3 [11.0] kg) college students volunteered. All participants were healthy, active for 3 days a week for at least 1 hour, and indicated at least a 7 on the Tegner activity scale (recreationally active in soccer, basketball, running, etc).32 Inclusion criteria included young adults, 18–25 years. Exclusion criteria included color blindness, lower-extremity surgery within the past 6 months or other injury within the past 6 weeks, ligament damage to the ankle, knee, or hip, any lower-extremity injury that limits participation, and any nervous system disorder such as Parkinson’s, multiple sclerosis, or degenerative disease. This study was reviewed and approved by the Ohio University Institutional Review Board. Participants provided informed and written consent, and demographics (sex, age, height, and weight) before participation in the study.

Procedures

Each participant performed 2 days of testing. Two weeks separated each testing day to allow for a proper washout period and eliminate a potential learning effect. Participants were instructed to not consume alcohol or caffeine within 24 hours of testing. Participants were also instructed not to participate in exercise 24 hours before testing. On day 1, participants performed the 4 new neurocognitive single-leg hops. The order of testing was randomized by having the participant pull 1 piece of paper out at a time and perform whichever single-leg hop was written on the piece of paper. On day 2, the participants performed the same 4 neurocognitive single-leg hop tests in a randomized order (the same way as day 1) to determine reliability. All testing was completed by the same examiner (N.M.).

The 4 neurocognitive single-leg hops were implemented using the FitLight system (FITLIGHT Sports Corp, Aurora, Canada). The FitLights function as visual stimuli as well as triggers or sensors that record time to contact with the sensor and record time from 1 sensor displaying a stimulus to contact with the same or other sensor. Participants were given 1 attempt to familiarize themselves with the neurocognitive single-leg hops.9 Similar procedures were implemented when instructing how to complete each hop as outlined by Noyes et al8 and Barber et al.22 Specifically, for every single-leg hop, participants were instructed to “hop as fast and as far as you can” and performed them until 3 successful trials on each leg were performed. Participants on average failed (did not stick the landing) 1.20 jumps for each leg during each single-leg hop test. Before each session, individuals were allowed to complete a warm-up of their choice for 15 minutes.

Session 1. Participants performed the following 4 single-leg neurocognitive hops (Figure 1) until 3 successful trials were recorded:

Single-leg central-reaction reaction hop: The participant stood on 1 leg, and waited for the stimulus FitLight at eye level, 360 cm away from the participant, to flash the correct color (green—hop, red—do not hop). When the correct color flashed, the participant jumped as far as possible with a single hop. A timing FitLight was placed 25 cm in front of the participant on the ground, and synced with the FitLight at eye level to measure reaction time (time from eye level FitLight flashing green until the participant executed the hop signaled by crossing the timing FitLight placed in front of the participant’s foot) (Figure 2). The participant had to land and maintain balance on the same leg without touchdown of the other foot, as well as trigger the timing FitLight for a successful trial. The maximum distance (in centimeters) and fastest reaction time (in seconds) were recorded for each leg and used for statistical analysis.

Single-leg memory triple hop: The participant stood on 1 leg with a FitLight 360 cm in front of them and 60 cm to the right at eye level, and 1 FitLight on the ground 25 cm distance away from their foot to measure reaction time (time from eye level FitLight flashing “correct color” until the participant executed the hop signaled by crossing the timing FitLight placed in front of the participant’s foot). The FitLight flashed 1 of 6 colors (red, dark blue, light blue, yellow, green, and purple) at random. Each trial, a random color was selected as the “hop” color, while the remaining 5 were assigned as “do not hop.” When the “hop” color flashed, the participant would single-leg hop forward in 3 consecutive hops. A successful trial was defined as getting a single “hit” by jumping over the FitLight on the ground while performing 3 consecutive hops without falling or touching the opposite foot to the ground. The maximum distance (in centimeters) and fastest reaction time (in seconds) was recorded for each leg and used for statistical analysis.

Single-leg peripheral-reaction crossover hop: The participant stood on 1 leg, and looked directly in front of them. Two stimuli FitLights, 1 each side, were placed 60 cm away from the center line
and on the edge of the participant’s peripheral vision at a maximum angle of 200° to 220° horizontally. This was done by manually moving the FitLight as far into the participant’s peripheral vision as possible and still see each light. One of the FitLights blinked red (do not hop) or green (hop); the respective red or green color was flashed at random. A FitLight timing gate was set up 60 cm ahead of the participant, to measure reaction time (time from peripheral FitLight flashing green until the participant executes the hop.

Table 1  Reliability of Neurocognitive Hops Between Day 1 and Day 2 L and R Legs

<table>
<thead>
<tr>
<th></th>
<th>ICC$_{3,k}$; 95% CI</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-leg central-reaction hop R physical performance</td>
<td>.89; .70–.96</td>
<td>.91</td>
</tr>
<tr>
<td>Single-leg central-reaction hop L physical performance</td>
<td>.90; .72–.96</td>
<td>.92</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop R physical performance</td>
<td>.87; .71–.94</td>
<td>.88</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop L physical performance</td>
<td>.90; .75–.96</td>
<td>.91</td>
</tr>
<tr>
<td>Single-leg memory triple hop R physical performance</td>
<td>.89; .76–.95</td>
<td>.89</td>
</tr>
<tr>
<td>Single-leg memory triple hop L physical performance</td>
<td>.92; .82–.97</td>
<td>.92</td>
</tr>
<tr>
<td>Single-leg pursuit 6 m hop R physical performance</td>
<td>.91; .79–.96</td>
<td>.90</td>
</tr>
<tr>
<td>Single-leg pursuit 6 m hop L physical performance</td>
<td>.91; .80–.96</td>
<td>.91</td>
</tr>
<tr>
<td>Single-leg central-reaction hop R cognitive performance</td>
<td>.95; .88–.98</td>
<td>.95</td>
</tr>
<tr>
<td>Single-leg central-reaction hop L cognitive performance</td>
<td>.93; .83–.97</td>
<td>.93</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop R cognitive performance</td>
<td>.98; .71–.99</td>
<td>.99</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop L cognitive performance</td>
<td>.97; .94–.99</td>
<td>.98</td>
</tr>
<tr>
<td>Single-leg memory triple hop R cognitive performance</td>
<td>.96; .94–.99</td>
<td>.97</td>
</tr>
<tr>
<td>Single-leg memory triple hop L cognitive performance</td>
<td>.97; .94–.99</td>
<td>.97</td>
</tr>
<tr>
<td>Single-leg pursuit 6 m hop R cognitive performance</td>
<td>.89; .75–.95</td>
<td>.89</td>
</tr>
<tr>
<td>Single-leg pursuit 6 m hop L cognitive performance</td>
<td>.88; .73–.95</td>
<td>.88</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; L, left; R, right.
### Table 2  Mean Difference Between Day 1 and Day 2 L and R Legs

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Measurement Type</th>
<th>Mean Difference; 95% CI</th>
<th>t value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-leg central-reaction hop R</td>
<td>Physical performance</td>
<td>1.25; −0.52 to 3.02</td>
<td>1.47</td>
<td>.16</td>
</tr>
<tr>
<td>Single-leg central-reaction hop L</td>
<td>Physical performance</td>
<td>0.25; −1.59 to 2.09</td>
<td>0.28</td>
<td>.78</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop R</td>
<td>Physical performance</td>
<td>−1.00; −3.69 to 1.69</td>
<td>−0.77</td>
<td>.45</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop L</td>
<td>Physical performance</td>
<td>−1.27; −2.84 to 0.29</td>
<td>−1.69</td>
<td>.12</td>
</tr>
<tr>
<td>Single-leg memory triple hop R</td>
<td>Physical performance</td>
<td>1.32; −0.81 to 3.45</td>
<td>1.29</td>
<td>.21</td>
</tr>
<tr>
<td>Single-leg memory triple hop L</td>
<td>Physical performance</td>
<td>−0.23; −2.57 to 2.11</td>
<td>−0.21</td>
<td>.84</td>
</tr>
<tr>
<td>Single-leg pursuit 6m hop R</td>
<td>Physical performance</td>
<td>0.01; −0.07 to 0.09</td>
<td>0.24</td>
<td>.81</td>
</tr>
<tr>
<td>Single-leg pursuit 6m hop L</td>
<td>Physical performance</td>
<td>−0.01; −0.09 to 0.08</td>
<td>−0.24</td>
<td>.82</td>
</tr>
<tr>
<td>Single-leg central-reaction hop R</td>
<td>Cognitive performance</td>
<td>−0.002; −0.029 to 0.025</td>
<td>−0.15</td>
<td>.88</td>
</tr>
<tr>
<td>Single-leg central-reaction hop L</td>
<td>Cognitive performance</td>
<td>−0.017; −0.054 to 0.021</td>
<td>−0.99</td>
<td>.37</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop R</td>
<td>Cognitive performance</td>
<td>−0.007; −0.014 to 0.011</td>
<td>−1.14</td>
<td>.44</td>
</tr>
<tr>
<td>Single-leg peripheral-reaction crossover hop L</td>
<td>Cognitive performance</td>
<td>−0.007; −0.027 to 0.011</td>
<td>−0.84</td>
<td>.41</td>
</tr>
<tr>
<td>Single-leg memory triple hop R</td>
<td>Cognitive performance</td>
<td>0.008; 0.002 to 0.021</td>
<td>0.17</td>
<td>.86</td>
</tr>
<tr>
<td>Single-leg memory triple hop L</td>
<td>Cognitive performance</td>
<td>−0.005; −0.023 to 0.013</td>
<td>−0.59</td>
<td>.56</td>
</tr>
<tr>
<td>Single-leg pursuit 6m hop R</td>
<td>Cognitive performance</td>
<td>0.007; 0.002 to 0.024</td>
<td>0.16</td>
<td>.87</td>
</tr>
<tr>
<td>Single-leg pursuit 6m hop L</td>
<td>Cognitive performance</td>
<td>0.005; 0.001 to 0.025</td>
<td>0.82</td>
<td>.42</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; L, left; R, right.

Figure 3 — Right leg physical performance Bland–Altman plots for 4 neurocognitive single-leg hops.
signaled by crossing the timing gate of the FitLight placed in front of them). The participant stood on a center line and either hopped right, left, right forward, and across the tape measure if the right FitLight flashed green, or left, right, left forward, and across the tape measure if the left FitLight flashed green. A successful trial consisted of setting off the timing gate, and completing the 3 hops in the correct directions without falling or placing the other foot down. The maximum distance (in centimeters) and fastest reaction time (in seconds) was recorded for each leg and used for statistical analysis.

Single-leg pursuit 6m hop: The participant stood on 1 leg, with one stimulus FitLight 6 m away directly in front of the participant, another stimulus FitLight 6 m away and at a 45° angle to left of the participant, and one more stimulus FitLight 6 m away and at a 45° angle to the right of the participant. At a random time, 1 FitLight would blink either red (do not hop) or green (hop). The participant had to wait for the correct color (green), then hop to the respective light as quickly as possible. Participants needed to hop on 1 foot to the correct light, and trigger it for a successful trial to be recorded. The participant’s minimum time for completion or quickest trial (in seconds) and fastest reaction time (in seconds) (time from eye level FitLight flashing green until the participant executed the hop signaled by crossing the timing FitLight placed in front of the participant’s foot) was recorded for each leg and used for statistical analysis.

Session 2. At session 2, the participants were instructed again on the 4 neurocognitive single-leg hop tests, and given 1 practice trial. The participants followed the procedures for the neurocognitive hops as stated above.

Statistical Analyses
Test–retest reliability was assessed for the 4 new neurocognitive single-leg hops using an intraclass correlation coefficient (ICC3, 1). Additionally, 1-sample t tests were conducted for the mean difference between the 2 time points for each single-leg hop test with the test value set at 0 (indicates no change in scores between the 2 time points) with an alpha level set at P < .05. Finally, Bland–Altman plots were constructed for each single-leg hop test to evaluate any systematic changes. The previous was completed for each leg and

![Figure 4](image-url) — Left leg physical performance Bland–Altman plots for 4 neurocognitive single-leg hops.
for the performance measure (distance or time to complete the hop) and cognitive measure (reaction time).

**Results**

Reliability was assessed for each leg by conducting ICC between the 2 test days for each neurocognitive single-leg hop. The ICC values and Cronbach’s alphas are listed in Table 1. For physical performance, all ICC values were above .85 (ranging from .87 to .92) and all Cronbach’s alpha values were above .8 (ranging from .88 to .92) indicating good to excellent reliability and consistency. The same was seen for the cognitive performance (reaction time) with the ICC values ranging from .88 to .98 and Cronbach’s alphas ranging from .88 to .98 indicating excellent reliability and consistency. Additionally, all 1-sample t tests were not significant (physical performance and cognitive performance; \( P > .05 \)), indicating that there was no change in scores between the 2 time points (Table 2). The Bland–Altman plots indicated that all 4 single-leg neurocognitive hop tests did not change systematically (physical performance and cognitive performance; Figures 3–6).

![Figure 3](image1)

**Figure 3** — Right leg physical performance Bland–Altman plots for 4 neurocognitive single-leg hops.

**Discussion**

The purpose of this study was 2-fold. Primarily, we proposed to create 4 new neurocognitive single-leg hop tests based on the 4 traditional hop tests created by Noyes et al. Additionally, we established reliability for the 4 new neurocognitive hop tests. Our results indicate that the 4 neurocognitive single-leg hop tests had good to excellent reliability for physical performance and cognitive performance and were consistent across time points.

The neurocognitive single-leg hop tests were designed for the participant to process cognitive and visual-spatial information and execute the hop based on an external signal instead of being able to jump when they choose. Each hop provided a different challenge to the participant. The single-leg central-reaction hop was designed as the most basic by adding 1 FitLight creating a “hop, no-hop” scenario using green and red colors. The single-leg memory triple hop created a “hop, no-hop” scenario with multiple no-hop stimuli, increasing the cognitive processing to suppress hop action due to varied no-go stimulus with only 1 go stimulus. The single-leg peripheral-reaction crossover hop incorporated testing peripheral...
vision reaction capability and increased the number of FitLights to 2. Finally, the most complex hop was the single-leg pursuit 6m hop, which had 3 FitLights for the participant to scan across their field of vision. The increased level of difficulty in each single-leg hop progressively adds standardized visual and cognitive challenges. The added layer of cognition has external validity to provide a better assessment of athletic performance under the visual and cognitively loaded athletic environment.

Current return to play decision-making has relied on functional measures only assessing physical performance under little or no cognitive load. However, sports-specific activities are more challenging than landing from a planned hop in a controlled environment, and thus the deficits seen in single-legged hop performance may be magnified in sports-specific activities. The traditional hop tests have been shown to predict successful short-term return to sport after nonoperative management. However, this classification algorithm has later been shown to have poor ability to predict return to sport at a 1 year following injury. Furthermore, even when hop performance (symmetry) is recovered, reinjury remains high when returning to high competitive sport. Thus, the dependence on the traditional single-leg hop testing using limb symmetry to determine readiness for return to play may be inadequate to provide sufficient data to make return to play decisions, as the testing does not incorporate sufficient cognitive or visual challenge of athletic play.

The neurocognitive hops incorporate neurocognitive demand overlaid on the traditional functional performance testing to quantify both physical and cognitive/visual-motor capabilities. Neurocognitive testing has typically only captured deficits in reaction time, processing speed, and visual and verbal memory in isolation and traditionally for concussion specifically, when in fact these neurocognitive deficits may influence neuromuscular control in the lower-extremity via changes in situational awareness, arousal, and attention. Injury tends to occur when physical and cognitive variables interact, as video analysis indicates many noncontact anterior cruciate ligament injuries occur almost immediately after initial contact during landing, cutting, and movement with the knee in full extension, and typically with many players or visual distractors in close proximity, causing combined cognitive-motor perturbations that influence movement patterns.
The evidence indicating that poor neurocognitive performance is associated with an elevated risk of injury is further supported by studies demonstrating higher injury risk movement mechanisms in those with poorer neurocognitive performance and after a concussion. Thus, the mechanisms by which neurocognition influences injury risk are theorized to be due to influencing neuromuscular joint control when under cognitive or visual load. This is with athletes who have experienced a recent concussion or have lower neurocognitive performance that demonstrated altered joint mechanics during dual tasks or when required to respond to unanticipated cues. These data combined with poorer neurocognitive processing increasing risk of subsequent anterior cruciate ligament tear, indicate the need for clinical return to play tests that assess motor performance under standardized neurocognitive loads. These neurocognitive single-leg hop tests are able to reliability assess the combined integration of motor and neurocognitive performance and may better detect subtle degrades in either system.

Limitations

There are several limitations to this study. The level of athletic ability in the participants is a limitation for generalizability as they were recreational athletes, and the results may vary if all participants were Division 1 or professional athletes. Additionally, neurocognition includes many aspects not directly assessed with the neurocognitive hops; however, we selected key components of visual processing, reaction time, and working memory found to be predictive of future injury risk. Future research should investigate incorporating other measures of neurocognition into functional testing. Future research should also attempt to confirm the intrarater reliability of the neurocognitive tests and determine interrater reliability in other populations and postinjury. Additionally, prospective studies should be done to determine if physical–cognitive integrative performance can improve return to play decision-making or predict injury risk.

Conclusions

This study determined the test–retest reliability of 4 new neurocognitive single-leg hop tests. The findings of this research showed good to excellent reliability and scores did not change between time points. This indicates that the neurocognitive hops are repeatable between days as a measuring tool for the interaction between physical and cognitive athletic performance. Finally, neurocognitive hop tests could be a potential functional test for those recovering from lower-extremity injury and be more ecologically valid in return to play decision-making.

References

Development of Neurocognitive Hop Tests

9


