ABSTRACT

**Background:** Dual-task assessments can identify changes in postural control during recovery from a concussion. However, developing postural control in children presents a challenge when using adult balance assessments to examine children.

**Purpose:** The purpose of this study was to investigate the reliability of a cognitive dual-task postural control testing protocol among a youth sample with no history of concussion or exposure to head impacts.

**Study Design:** Reliability pilot study.

**Methods:** Testing comprised nine 120 second trials of standing on a force plate collecting data at 250 Hz. Test conditions included no dual-task, counting backwards by 2, counting backwards by 3, listening, and the Stroop test. Subjects completed each test with open and closed eyes, except for the Stroop test. The force plate was used to measure the subjects’ center of pressure (COP) trajectory.

**Results:** Nine healthy, youth subjects (average age: 11.6 ± 0.5 years) with no history of concussion or exposure to head impacts participated. Reliability was good (>0.6) or excellent (>0.75) for COP speed, sway, and sample entropy measures for several test conditions. The eyes open, no task condition produced the lowest COP measures. No differences were observed between the other dual-task conditions.

**Conclusion:** Given its high measures of reliability, this dual-task protocol might be able to detect postural control changes in concussed youth athletes.

**Evidence Level:** 2

**Keywords:** Balance, cognitive, force plate, pediatric

---

1 Department of Biomedical Engineering and Mechanics, Virginia Tech, Blacksburg, VA, USA
2 Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, USA

**Conflict of Interest:** The authors have no conflicts of interest to report.

**Acknowledgments:** Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number R01NS094410. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

**CORRESPONDING AUTHOR**
Eamon T. Campolettano
325 Stanger St.
Blacksburg, VA 24061
Phone: 516-306-7210
E-mail: eamonc@vt.edu
INTRODUCTION
Up to 1.9 million sports-related concussions occur annually within the United States' youth population.\textsuperscript{1,2} One of the most common symptoms associated with a concussion is balance dysfunction during quiet standing.\textsuperscript{3-7} These alterations are transient but serve as a key marker when evaluating a patient's neurologic health post-concussion.\textsuperscript{3,4,8} Developing postural control in children presents an additional challenge in using postural sway assessments in these instances.\textsuperscript{9} This continued development prevents the use of measures and protocols that have been examined for reliability in adults without first assessing their reliability in children.\textsuperscript{10-13}

Postural control assessments using a force plate traditionally consist of measures taken during quiet standing.\textsuperscript{14,15} Some researchers have made use of cognitive dual-tasks during postural sway assessments.\textsuperscript{16-20} The addition of a cognitive dual-task prevents the brain from allocating all of its resources towards balance and may provide a more sensitive testing protocol for assessing patient health post-concussion than quiet standing.\textsuperscript{21} In concussed adolescent and adult athletes, for whom cognitive and postural control deficits have been noted, the effect of performing cognitive dual-tasks is more pronounced than during quiet standing.\textsuperscript{18-20,22,23}

Several postural sway measures and cognitive dual-tasks were explored to identify a reliable combination for the potential assessment of youth football players immediately post-concussion, as well as during athlete recovery. The development of this protocol was predicated on two key outcomes: 1) test conditions must be reliable for individual subjects over time, and 2) variation in performance between subjects must be observed. Between-subject variation implies that the protocol is sensitive to differences in postural control. With these two outcomes met, the testing protocol would have the potential to differentiate COP measures of concussed athletes relative to their baseline values. The purpose of this study was to investigate the reliability of a cognitive dual-task postural control testing protocol among a youth sample with no history of concussion or exposure to head impacts.

METHODS
Nine youth male athletes (mean ± standard deviation age: 11.6 ± 0.5 years; height: 1.50 ± 0.07 m; mass: 40.4 ± 7.9 kg) were recruited to participate in this pilot study. Male subjects were included because the primary target application for the protocol developed in this study is youth football. The athletes were not actively involved in a contact sport during the study period and did not have a history of concussion. This study was approved by Virginia Tech's institutional review board, and all subjects provided verbal assent while their parents or guardians provided written consent.

Subjects completed four identical testing sessions spaced a week apart. During each session, postural sway was measured in nine test conditions, each of which had a duration of 120 seconds. Subjects were given a minimum rest period of 30 seconds between trials and completed each task with either eyes open or eyes closed. During the eyes open conditions, subjects were instructed to look at a target on a computer screen placed at eye level at a distance of 1 m. The five dual-task conditions were no task, count backwards from 100 by two, count backwards from 100 by three, count the instances of a word in a passage read to the subject by the investigator, and the Stroop test. Each of the tasks were conducted with the subjects' eyes opened and closed, with the exception of the Stroop test, which was only conducted with the subjects' eyes open. The Stroop test involved reciting the color in which a word was written.\textsuperscript{24} The no task trials were conducted first each session, with the remaining trials randomized to account for possible effects due to test order. For all tests, subjects were instructed to stand facing forward with their feet together and touching, arms at their side, and to try to remain as still as possible.

During all trials, ground reaction forces and moments under the feet were sampled at 250 Hz using a force plate (AMTI, Watertown, MA). Center of pressure (COP) coordinates were then determined and used to compute eight COP-based measures of postural sway.

Four traditional sway measures were calculated. COP standard deviation in the anterior-posterior (AP) and medial-lateral (ML) directions were calculated along the respective axis. COP mean speed was calculated as the overall COP path length divided by...
the test duration (120 seconds). Additionally, the COP 95% confidence ellipse area was calculated as

\[ 2p F_{0.5,2,N-2} \left( SD_{AP}^2 SD_{ML}^2 - SD_{AP-ML}^2 \right) \]

where \( F_{0.5,2,N-2} \) represented the value of the F statistic for a bivariate distribution of \( N \) data points at a confidence level of 95%, \( SD_{AP} \) was standard deviation in the AP direction, \( SD_{ML} \) was standard deviation in the ML direction, and \( SD_{AP-ML} \) was the covariance between the AP and ML directions.\(^{14}\)

Four entropy measures were calculated. Sample AP and ML entropy assessed the variability of the COP trajectory along each direction. Sample entropy was determined by comparing a given data vector template from the COP trajectory to all other vectors within the trajectory and counting all those that were within a defined similarity range.\(^{25}\) Renyi Entropy and Shannon Entropy were also calculated using graphical representations of the COP data. More specifically, these involved developing a grid that is divided based on the standard deviation in the AP and ML directions.\(^{15,26}\) The number of points within each subunit of the grid were then summed, with the probability of a COP coordinate residing in a particular subunit serving as the input parameter for computing both Renyi and Shannon Entropy. Detailed background on the algorithm is available.\(^{27,28}\) In general, higher entropy values are associated with less repeatable or predictable COP trajectories, which would be representative of poorer postural control.

ANOVA was used to investigate the effect of session, test condition, and their interaction on COP measures with a significance level of 0.05. Tukey's Honest Significant Difference test was used to assess specific differences in these factors. ANOVA was also used on a subset of the data to determine the effect of test order on COP measures by test condition. Since the eyes open and eyes closed, no task conditions were always conducted first and second, they were not included in this subset analysis to assess the effect of test order.

Test-retest reliability over the four weeks of testing was defined by computing the intraclass correlation coefficient [ICC(3,4k)] for each measure and test condition. ICC values were interpreted using the following criteria: 0.00-0.39 poor, 0.40-0.59 fair, 0.60-0.74 good, and 0.75-1.00 excellent.\(^{29}\) A metric or test condition could have a high ICC value, or reliability, while failing to differentiate subjects from each other. Between-subject and within-subject variability, as well as the ratio of between-subject variability to within-subject variability, were computed to account for this. Between-subject variability was calculated as the standard deviation of mean subject performance for each measure within each test condition. Within-subject variability was calculated as the standard deviation of subject performance standard deviation for each measure within each test condition. Test conditions and measures with high ICC values, higher measures of between-subject variability and lower levels of within-subject variability are desirable for a healthy population. This would indicate that these experimental conditions and measures are reliable indicators of intra-subject balance performance and are sensitive to inter-subject differences.

RESULTS

None of the traditional sway measures exhibited a session by test condition interaction effect (\( p = 1.000 \)) or a main effect for session (\( p > 0.3099 \)). ML standard deviation (\( p = 0.0039 \)) and COP mean speed (\( p < 0.0001 \)) exhibited a main effect for test condition, but 95% confidence ellipse area (\( p = 0.0547 \)) and AP standard deviation (\( p = 0.3401 \)) did not (Figure 1). The eyes open, no task condition was associated with the lowest measures. In general, more sway was observed along the AP axis than the ML axis across all test conditions.

None of the entropy measures exhibited a session by test condition interaction effect (\( p > 0.2082 \)) or a main effect for session (\( p > 0.0888 \)). Entropy measures did not differ between test conditions or session (\( p > 0.05 \)), with the exception of Shannon Entropy (\( \text{Condition: } p < 0.0001 \); \( \text{Session: } p = 0.858 \)). For this measure, the eyes closed, no task condition was associated with higher entropy than each of the eyes open conditions, with the exception of the eyes open, no task condition, and the eyes open, listening condition was associated with lower entropy than each of the eyes closed test conditions (Figure 2).
or excellent reliability and was observed to vary by test condition (Table 2). The eyes closed, no task test condition was associated with good or excellent test-retest reliability (ICC > 0.7) for each of the five measures, as well as a ratio of between-within subject variability exceeding 3.0 for all measures. Among the measures and test conditions investigated here, AP Entropy and ML Entropy during the Stroop test exhibited the highest reliability (> 0.90) and ratio of between-within subject variance (> 8.0).

**DISCUSSION**

This study assessed the reliability of a dual-task postural control testing protocol among a group of youth male athletes without a history of concussion or recent exposure to head impacts. Traditional COP measures and COP entropy measures were investigated during different cognitive dual-tasks to identify a reliable combination for the potential use in clinical and research settings.

Further, test order was not associated with performance (p > 0.05) for any of the measures or test conditions.

Reliability varied across measures when including all test conditions (Table 1). Renyi Entropy, Shannon Entropy, and 95% ellipse area exhibited poor to fair reliability (ICC < 0.6). These measures were also associated with lower ratios of between-subject variability to within-subject variability when compared to the other measures investigated in this study. AP Entropy (ICC = 0.761 [0.683-0.827]) and ML Entropy (ICC = 0.809 [0.745-0.864]) were associated with excellent test-retest reliability, while AP standard deviation, ML standard deviation, and COP mean speed were observed to have good test-retest reliability (Table 1).

Test-retest reliability for each test condition was assessed for the five measures that exhibited good or excellent reliability and was observed to vary by test condition (Table 2). The eyes closed, no task test condition was associated with good or excellent test-retest reliability (ICC > 0.7) for each of the five measures, as well as a ratio of between-within subject variability exceeding 3.0 for all measures. Among the measures and test conditions investigated here, AP Entropy and ML Entropy during the Stroop test exhibited the highest reliability (> 0.90) and ratio of between-within subject variance (> 8.0).
exhibited excellent reliability for the highest number of sway measures. AP Entropy and ML Entropy were observed to have the highest reliability (ICC > 0.90), as well as the highest ratio of between-subject variability to within-subject variability, for the Stroop test condition. This combination of reliability and variability suggests the capacity for this test condition to be sensitive to differences between subjects, which could potentially include concussion-related differences.

Postural control testing with youth populations has largely assessed subjects during quiet standing without the addition of a cognitive task. During eyes open and eyes closed testing with youth subjects, it has been observed that healthy and concussed populations have differing measures of postural control. Renyi Entropy values have ranged from 4.5-5.5, which was consistent with values reported here. COP mean speed values have also been observed to vary from 1.5-3.5 cm/s in a healthy, youth

### Table 1. Intraclass correlation coefficients for each metric.

<table>
<thead>
<tr>
<th>Metric</th>
<th>ICC [95% Confidence Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>COP Mean Speed</td>
<td>0.707 [0.619 – 0.785]</td>
</tr>
<tr>
<td>AP Standard Deviation</td>
<td>0.678 [0.585 – 0.762]</td>
</tr>
<tr>
<td>ML Standard Deviation</td>
<td>0.626 [0.525 – 0.720]</td>
</tr>
<tr>
<td>95% Ellipse Area</td>
<td>0.550 [0.419 – 0.639]</td>
</tr>
<tr>
<td>Renyi Entropy</td>
<td>0.151 [0.052 – 0.251]</td>
</tr>
<tr>
<td>Shannon Entropy</td>
<td>0.533 [0.423 – 0.642]</td>
</tr>
<tr>
<td>AP Entropy</td>
<td>0.761 [0.683 – 0.827]</td>
</tr>
<tr>
<td>ML Entropy</td>
<td>0.809 [0.745 – 0.864]</td>
</tr>
</tbody>
</table>

*COP = Center of pressure; AP = Anterior-posterior; ML = Medial-lateral*
observed in this study. Lastly, test order was determined to not be a significant factor in the measures of COP used in this study. With nine trials lasting 120 seconds, the potential existed for fatigue to affect those test conditions towards the end of the testing protocol. By giving subjects time to rest between each test condition, the risk for fatigue was mitigated. Moreover, the future implementation of these measures and test conditions should be limited to those that exhibited good to excellent reliability, and thus reduce the number of trials required.

In addition to test-retest reliability, between-subject and within-subject variation were also computed for each COP measure and test condition combination (Table 2). A viable measure would be one in which within-subject variation was low, especially when compared to between-subject variation. This would mean that individual subjects’ performance varied from each other while each individual subject’s performance did not vary over time. If individual postural control were compromised due to concussion, population during quiet standing. Though postural control is still developing in the youth population, previous research with this population has shown good to excellent measures of reliability for static postural control testing. The COP and reliability values measured in the present study for the no task test conditions were in the same range as previous research utilizing quiet standing protocols.

At least one COP measure for each test condition was observed to vary from the eyes open, no task test condition (Figures 1 and 2). The addition of a cognitive task and the elimination of visual feedback independently or combined resulted in elevated measures for the COP measures utilized in this study. Performance was not tied to session for this subject pool. These youth subjects were not exposed to head impacts during the testing timeframe and it would not be expected for their postural control, as measured by static standing trials, to change. As subject performance did not consistently improve over time for any test conditions, no learning effect was observed in this study. Lastly, test order was determined to not be a significant factor in the measures of COP used in this study. With nine trials lasting 120 seconds, the potential existed for fatigue to affect those test conditions towards the end of the testing protocol. By giving subjects time to rest between each test condition, the risk for fatigue was mitigated. Moreover, the future implementation of these measures and test conditions should be limited to those that exhibited good to excellent reliability, and thus reduce the number of trials required.

In addition to test-retest reliability, between-subject and within-subject variation were also computed for each COP measure and test condition combination (Table 2). A viable measure would be one in which within-subject variation was low, especially when compared to between-subject variation. This would mean that individual subjects’ performance varied from each other while each individual subject's performance did not vary over time. If individual postural control were compromised due to concussion,
we would expect the subject to have higher postural sway measures relative to a baseline because between-subject variability suggested good sensitivity. Over time, we would also expect the subject's postural sway measures to return to baseline values, which would represent within-subject variability. A protocol with specific sway measures and test conditions that maximize between-subject variability relative to within-subject variability presents the best opportunity to detect postural control changes in this population.

Several limitations of this study should be noted. This study investigated the reliability of measures and dual-task test conditions among a healthy male youth sample with no history of concussion. The validity of these measures and test conditions in regards to the effects of concussion must still be evaluated. The reliability values reported in this study may not be generalizable to a different study population, including youth female athletes. As postural control is still developing in youth males, reliability of the measures and tests investigated in this study may actually be lower over longer periods of time during which further neuromusculoskeletal control development may occur. Because only nine athletes participated in this study, the results are not meant to be generalized but to serve as a basis for a larger analysis involving youth football players who are exposed to repetitive head impacts and may incur concussions.

**CONCLUSIONS**

At present, postural control testing within youth populations largely utilizes assessments that were validated for adults. The reliability and viability of using these tools with a population that is still developing balance and postural control has not been addressed, which limits the generalizability of the results from these studies. Dual-task assessments have become increasingly used as part of the return-to-activity protocol in instances of sports-related concussion, as the addition of a dual-task environment increases sensitivity to either cognitive or postural deficits. A subset of the dual-task assessments utilized in this study resulted in good or excellent levels of test-retest reliability. This protocol will be applied within a group of youth football players, some of whom may experience a clinically-diagnosed concussion.

**REFERENCES**

14. Prieto TE, Myklebust J, Hoffmann R, Lovett E, Myklebust B. Measures of postural steadiness:


