Test Re-Test Reliability of Three Versions of the 3-Cone Test

Jason G. Langley
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
Langley, Jason G., "Test Re-Test Reliability of Three Versions of the 3-Cone Test" (2011). Graduate Theses, Dissertations, and Problem Reports. 3385.
https://researchrepository.wvu.edu/etd/3385

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Test Re-Test Reliability of Three Versions of the 3-Cone Test

Jason G. Langley

Dissertation submitted to
the College of Physical Activity and Sport Sciences
at West Virginia University
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy in
Kinesiology with an emphasis in
Physical Education Teacher Education

Sean M. Bulger, Ed.D. Co-Chair
Robert Chetlin Ph.D.
Kristen Dieffenbach Ph.D.
W. Guyton Hornsby Ph.D.
Robert L. Wiegand, Ed. D. Co-Chair

Department of Coaching and Teaching Studies

Morgantown, WV
2011

Keywords: Agility, Reliability, Test Construction, 3-cone

Copyright 2011 Jason G. Langley
ABSTRACT

Test Re-Test Reliability of Three Versions of the 3-Cone Test

Jason G. Langley

The primary purpose of this study was to examine the test re-test reliability of three versions of the 3-Cone Test (3CR). In addition, comparisons between the versions of the tests; and the relationships with leg dominance and body mass index (BMI) were to be determined. Forty students enrolled in classes in the Department of Physical Education at a mid-Atlantic university participated. On each of the testing days the participants performed ten total trials. In a random order, they performed three trials to the right (3CR), three to the left (3CL), and two modified trials (3CMR and 3CML), in which a visual cue was given to indicate which path was to be followed. Intra-class correlations (ICC) indicate a moderate to high reliability for the four tests, 3CR 0.81 (0.66-0.84 95%CI), 3CL 0.73 (0.54-0.85), 3CMR .83(0.71-0.91), and 3CML 0.80 (0.65-0.89). Small standard error of the measurement (SEM) were found and ranged from 0.25-0.34. Pearson correlations between tests were high (0.82-0.92) on day one as well as day two (0.84-0.89). Having only six left-footed participants did not allow for comparisons between dominant side and test direction; however, BMI was moderately correlated with each of the tests ranging from 0.70-0.73. These results indicate each version of the 3CR is reliable; however, further testing is needed with athletic populations. The high correlation indicates only one version of the test is needed to reliably test for agility. By definition only the 3CMR and 3CML should be considered tests of agility due to their inclusion of a reactive component. Future studies utilizing agility testing and training should incorporate technology including gate systems and video analysis. These tools will allow for the better design of open tests that simulate game conditions.
DEDICATION

This dissertation is dedicated to my parents:

John and Nancy Langley

And to my wife Nancy and son Carter

This is very much a shared accomplishment that would not have been

achieved without your support.
ACKNOWLEDGEMENTS

I would first like to recognize my outstanding committee for their guidance, not only during this process, but also throughout my academic preparation. I will carry the lessons I've learned throughout my career, and strive to achieve the high levels of professionalism each of you display. Thank you Dr. Bulger, Dr. Chetlin, Dr. Dieffenbach, Dr. Hornsby, and Dr. Wiegand for your many contributions.

I would also like to acknowledge the faculty, staff, coaches, and students in the Department of Physical Education at West Virginia University Institute of Technology. I sincerely believe the completion of my degree, while teaching full-time, would not have been completed as efficiently as it was without your support. Thank you Dr. Elmore, Mark Jones, Reginald Smith, both the men's and women's basketball, and baseball coaching staffs for the many “favors” that allowed my data collection to be possible. Good ideas remain good ideas without subjects, so thank you to all of the students who volunteered for my study.

I would individually like to acknowledge George King and Zsolt Murlasits, your mentorship has guided me through many personal and professional challenges, and means a great deal to me. I know I can always count on your advice and friendship.

To my graduate school “family”: we will always be bound by our alma mater, and I'm sure our friendships will continue to grow even as we settle in to our new homes across the country.

Lastly, I would like to acknowledge the support from my Grandma and Papa Hall, Grandma Langley (wish you could have seen me reach this goal), Uncle Jeff, and my sister Shannon (Tim, J.R., and Ben). You have probably bragged on me more than I deserve, but your love will always mean the most to me.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>4</td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
</tr>
<tr>
<td>Testing Procedures</td>
<td>5</td>
</tr>
<tr>
<td>Familiarization and test, re-test protocols</td>
<td>5</td>
</tr>
<tr>
<td>Instrument descriptions of 3-cone test</td>
<td>6</td>
</tr>
<tr>
<td>3-cone modified</td>
<td>7</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>8</td>
</tr>
<tr>
<td>Results</td>
<td>9</td>
</tr>
<tr>
<td>Discussion</td>
<td>9</td>
</tr>
<tr>
<td>Scientific foundation for agility</td>
<td>10</td>
</tr>
<tr>
<td>Limitations in the literature</td>
<td>13</td>
</tr>
<tr>
<td>Interpreting results</td>
<td>19</td>
</tr>
<tr>
<td>Limitations of the present study</td>
<td>24</td>
</tr>
<tr>
<td>Future directions</td>
<td>26</td>
</tr>
<tr>
<td>Conclusions</td>
<td>28</td>
</tr>
<tr>
<td>References</td>
<td>30</td>
</tr>
<tr>
<td>Appendices</td>
<td>48</td>
</tr>
<tr>
<td>Appendix A: Administration of three versions of the 3-cone test</td>
<td>48</td>
</tr>
<tr>
<td>Appendix B: Review of Literature</td>
<td>52</td>
</tr>
<tr>
<td>Appendix C: Additional Methods</td>
<td>85</td>
</tr>
</tbody>
</table>
Appendix D: Informed Consent & Subject Demographic Questionnaire........121
List of Tables

Table 1. Participant Characteristics……………………………………………………………..36

Table 2a and 2b. Balance Latin Square for Determining Test Block Order & Balance
Latin Square for Determining Trials of the 3-cone Modified …………………………37

Table 3. Performance Characteristics and Reliability Coefficients…………………………38

Table 4. Correlation matrix for the fastest trials on Day 1……………………………………39

Table 5. Correlation matrix for the fastest trials on Day 2…………………………………40

Table 6. Pearson's correlation (r) for BMI and fastest overall trial…………………………41
List of Figures

*Figure 1.* Model of Agility (Chelladurai, 1976)…………………………………………42

*Figure 2.* Conceptual Framework of Agility (Sheppard & Young, 2006)……………43

*Figure 3.* 3-cone right (3CR) ……………………………………………………………..44

*Figure 4.* 3-cone left (3CL) …………...…………………………………….45

*Figure 5.* 3-cone modified (3CM)………………………………………………………46

*Figure 6a. and 6b* Starting position of the cue for 3CM & position of cue for 3CML when
directing the runner to the tester's right ……………………………………………………..47
Among the sport science community consensus has yet to be reached regarding a definition for agility (Chelladurai, 1976; Sheppard, Young, Doyle, Sheppard, & Newton, 2006). Multiple definitions of agility have appeared in measurement texts, research studies, and training publications (Chelladurai, 1976; Chelladurai & Yuhasz, 1977; Clarke, 1950; Cureton, 1947; McCloy & Young, 1954; Sheppard & Young, 2006; Young, McDowell, & Scarlett, 2001). The definitional variations within the construct have significantly altered the psychometric quality of the tests designed to quantify agility for the purposes of skill acquisition (Beise & Peaseley, 1937; Mohr & Haverstick, 1956; Rarick, 1937), athletic success prediction (Carruth, 1952; Gates & Sheffield, 1940; Hoskins, 1934; Johnson, 1934; Larson, 1941; Lehsten, 1948), and possibly the maintenance of fitness/health (Barnett, Van Beurden, Morgan, Brooks, & Beard, 2008; Smyth & O'Keefe, 1999). As evidence of this disconnect, Craig (2004) noted a “gap between the applied and the scientific knowledge” (p. 13). In other words, practitioners and sport scientists do not always speak the same language. This results in the misclassification or application of tests. This gap may be due, at least in part, to the many different perspectives on agility that persist. Each definition is derived from the sub-disciplinary perspective (biomechanics, motor learning, physical education, or coaching) from which the measurement is conducted. Many of the tests fail to follow or report the procedures by which they arrived at the new test making replication difficult. Often the test was then used for another application improperly or not as the original author intended.

This issue is seen less frequently in the literature published since 2008. Chelladurai (1976) was first to criticize the definitional variations and inconsistencies,
followed by a further analysis from Sheppard & Young (2006). Combined their work has advanced agility as a definable construct that can be investigated for the purpose of talent identification as well as for analyzing progress following a training program.

Chelladurai (1976) recognized the limitations of the available definitions of agility. He was critical that no definition included a response to a stimulus, which is a key component of any agile maneuver within a sport performance context. The result was the development of a new definition and corresponding model (see Figure 1) for classifying the various types of agility: “rapid and accurate movement of the total body in response to the perceived stimuli” (Chelladurai, 1976, p. 37). Chelladurai’s model included four classifications of agility: simple, temporal, spatial, and universal. Time (when) and space (where) determine which classification is chosen. An example of a simple task would be the long jump. The long jump is begun and executed without any cue or change in direction of the take-off board or landing pit. Based on his own definition this class would neither be considered agility nor change of direction speed (CODS). Temporal agility requires the reaction to a cue but lacks spatial uncertainty, for example, sprint events are begun on the firing of a pistol but the sprinters know the distance of the race. Spatial agility is the exact opposite of temporal, the event is not cued but there is spatial uncertainty. An example of this classification type is the receiving of a serve. The returner knows when the ball is going to be served but there is a level of uncertainty as to where the ball will be placed and/or the pace with which the ball was struck. The final class, universal agility, is a combination of the two previous sub-classes. It is both cued and spatially uncertain. This class is the most prevalent in sport. Goal keeping on a breakaway in the sport of hockey or soccer provides the best example.
The goalie is uncertain as to when the shot will be taken, the shots direction and initial velocity (Chelladurai, 1976).

Thirty years later, Sheppard and Young (2006) in a review of agility literature described the uniqueness of Chelladurai’s model, and its usefulness for coaches and sport scientists to better classify skills specific to a particular sport. This knowledge allows for the creation of tests and drills to be practiced that target the sub-components of a skill or the skill in total. However, Young and colleagues (2002) identified the limitations of Chelladurai’s model and proposed an alternative model that focused solely on “universal” agility (see Figure 2). This model identifies two major contributing factors to successful agility performance- CODS and perceptual and decision-making factors. Each major component is determined by a group of individual characteristics that serve as potential areas for research. Based on this merging of theoretical concepts, a new definition of agility was proposed: “a rapid whole-body movement with changing velocity or direction in response to a stimulus” (Sheppard & Young, 2006, p. 922). This definitional change effectively makes all but a few common assessment procedures for agility extinct, as they are only tests of CODS.

For example, the 3-Cone test (3CR) is used as a measure of agility during the National Football League Player Draft Combine (NFLPDC); however, little data has been published with concern to reliability of the test. Based on the Sheppard and Young's new definition of agility it cannot be considered valid on its face because it does not require any response to a stimulus. The test requires the participant to run full speed forward a total of 27.432 meters, while having to change direction a total of five times. The test is described in detail in the methods section. In addition to the 3CR drill, the study was
designed to re-design the test so that it met the new definitional requirements. Upon
determining the reliability, face validity may also be considered.

In view of the high-profile use of the 3CR in strength and conditioning profession,
and lack of data collected on the test; an investigation into the test’s reliability is
warranted. Therefore, the purpose of this study was to (a) determine if the 3CR is
reliable; (b) determine the reliability of three modifications (3-cone left (3CL) and 3-cone
modified right (3CMR) and 3-cone modified left (3CML)) of the 3CR; (c) establish
correlation coefficients amongst the four tests; and (d) ascertain the relationships between
the tests and two factors included in the Sheppard and Young (2006) model: body mass
index (BMI), and leg dominance.

Method

Participants

Data were collected on 49 male participants over the age of 18, however, 9 failed
to complete at least one of the required trials leaving the final sample size at 40.
Convenience sampling was used to recruit participants from the student body at a mid-
Atlantic university in the Department of Physical Education. An a priori power analysis
indicated a minimal sample size of \( n = 34 \) would be required to achieve 95% statistical
power with an \( r \geq 0.7 \). Each participant participated voluntarily, and was free of a lower
body injury that would limit involvement in the study. Prior to testing, all participants
read and signed an informed consent form approved by the university’s institutional
review board. They also completed a brief demographic questionnaire that included
information regarding their participation in sport as well as their history of injuries. At
the point of data collection participant height, weight, and leg dominance were recorded.
Descriptions of participant age, height, weight, BMI, dominant foot, and athletic participation levels are included in Table 1 (mean ± SD). Each participant acted as their own control in the test, re-test reliability study.

**Testing Procedures**

Three versions of the 3CR were administered. Each variation of the test required agility in the horizontal and transverse planes with varying requirements of perceptual reaction. The three tests included the 3CR, 3CL and 3CMR or 3CML. There has been no published data regarding the reliability of any of these test variations.

**Familiarization and test re-test protocols.** This study was conducted as a prospective test, re-test design to establish reliability. The components of the test battery included the 3CR test, and three modifications 3CL, 3CMR and 3CML. Each participant was tested on three occasions separated by at least 48 hours. Prior to each testing session, all participants completed a dynamic warm-up that included several stretching exercises designed to prepare the lower body for maximal sprinting and high velocity changes of direction.

Each participant performed a dynamic warm-up lasting approximately 10 minutes. Participants began by completing two laps around the gymnasium using a moderately paced jog and two sets of high knees and heel-ups. Dynamic stretching followed with one set of 10 repetitions on each side for hip flexion/extension, hip abduction/adduction, hip internal/external rotation, walking lunges, and walking toe touches. The warm-up was completed with one set of static stretching lasting 10 seconds each for the quadriceps, hamstrings, gastrocnemius/soleus complex, and tibialis anterior. Participants were also allowed to complete any personal stretch or to continue stretching as they
waited to be tested. The same member of the research team was present to supervise the warm-up and testing. The cool down consisted of a moderate jog for 2 laps around the gym and one set of static stretching for 10 seconds on each leg for hip flexion/extension, hip abduction/adduction, knee flexion/extension, and ankle plantar/dorsiflexors.

On day 1 (familiarization), all three tests were demonstrated by the lead researcher and performed by each participant as a means of introduction to test patterns, and to determine an acceptable rest interval during the actual trials on days 2 and 3. Each participant performed a total of 12 trials (four trials for each test). During the opening trial the participant was asked to run at a self-selected “half-speed,” while the second trial was to be at “three-quarter speed.” The final two trials were run at full speed. The order of the familiarization trials was the same for all participants 3CR, 3CL, and 3CM; and each trial was separated by a minimum of 50 seconds.

During testing on days 2 and 3, each participant completed 10 total trials; three 3CR, three 3CL, and four 3CM. The order of the three groups of tests was randomized using a Balanced Latin Square technique (see Table 2a) for each participant; and no participant completed the test groups in the same order on days 2 and 3. The four 3CM trials were also randomized to include two trials in each direction (see Table 2b). The fastest trial from each group was used for analysis. Each trial was separated by a minimum of 51.19 seconds. This time corresponds to an average trial time of 8.53 seconds during the familiarization trials and a 6:1 rest to work ratio.

**Instrument descriptions of 3-cone agility test (3CR and 3CL).** The 3CR is arranged by creating a triangle or an upside down L-shape setting up a start/finish line (A1 and A2), one cone is 4.57 meters (m) straight ahead (B), and a second cone 4.57m to
the right of B (C) (see Figure 3). The participants started in a two point stance (step 1) then sprinted to B where they touched the ground with their hand, (step 2) they then returned to A and touched with their hand before heading back, (step 3) toward B where they maneuvered around the outside of B (step 4) then weaved inside cone C, (step 5) and then back around the outside of C toward and around B (step 6) and finished through the line at A. The goal was to run as fast as possible without making contact with a cone. The time to complete the test was recorded to the nearest hundredth of a second by measurement of a stopwatch (Sportline 240). The participants started on their own; the watch was started on first movement and stopped upon crossing the line at A. The participant performed additional trials using the exact same procedure for the 3CL; except they went left at cone B and around cone D instead of right around cone C (see Figure 4).

3-cone modified (3CMR and 3CML). The 3CM test is a combination of 3CR and 3CL, having both the C and D cones as possible running patterns in the shape of a T. These modifications were necessary in order to have tests that require a response to a stimulus. The test variation included temporal reaction, and while this is not universal agility as described by Chelladurai (1976) it allowed for comparisons between tests. Had the tests been modified additionally to include a spatial modification, the test times could have increased enough to no longer be comparable based on the use of different energy systems. The modifications to the 3CM were specific to step 3 of the 3C (see Figure 5). During this step the runners reacted to a stimulus provided by the tester standing 4.11m directly in front of them or 1.829m from B (see Figure 5). During step 3, the runner’s turned to face cone B, as the runner passed the 2.29m mark on the floor the tester pointed
to either cone C or D. The participants then completed steps 4-6 in the same pattern as the 3CR or 3CL depending on the direction they headed. The directional signal used began with the fists of the tester together and both thumbs pointing upward directly in front of the chest of the tester (see Figure 6a). If the runner was assigned to go to their left, the tester pointed with his right thumb; and vice versa if the runner was to go right (see Figure 6b). The time to complete the test was recorded to the nearest hundredth of a second by measurement of a stopwatch.

**Data Analysis**

All data were analyzed using SPSS (Version 19 Chicago, IL) and Microsoft Excel (Version 2010). Statistical significance was set at p<.05. Reliability was determined via ICC model 3,1 for between-day analysis using the best trial each day for 3CR, 3CL, 3CMR and 3CML. The ICC 3,1 was chosen “because this investigation was an intrarater design with a single tester representing the only tester of interest” (Beekhuizen, Davis, Kolber, & Cheng, 2009, p. 2169). The ICC is a reliability coefficient that generates a ratio ranging from 0.00 to 1.00 to estimate the consistency of performance on repeated trials (a score of 0.00 indicates the measure was 100% unreliable) (Drouin, 2003). In calculating the ICC, the ratio determines accurate interpretation of how much variability in the observed measure is due to a change in the participant or the result of measurement error. Standard error of measurement (SEM) was also determined in conjunction with the ICC. Descriptive statistics were also used including means, and standard deviations for other variables such as age, height, weight and BMI. In addition, Pearson’s correlation coefficient was determined specifically between fastest time and BMI. Pearson
correlations were not tabulated in relation to leg dominance due to the imbalance of self-reported right (n=34) and left (n=6) foot dominance among participants.

**Results**

Table 3 identifies the performance characteristics (M±SD) of the fastest trials used for each of the four tests, as well as the ICC, SEM, mean time differences, and 95% confidence intervals CI's. Small SEM's (range 0.25-0.29) indicate that the measurement error was minimal. The ICC’s for the four tests were slightly lower than the .85 predicted at the outset of the study; however, according to Portney and Watkins (1992) >.75 should still be considered good. A summary of the correlations between the best times by test on day 1 are shown in Table 4, and day 2 in Table 5. Correlations between tests by days were high (0.82-0.92). Pearson correlations (0.70-0.73) demonstrate moderate to strong relationships between the fastest times on each of the four tests each day and BMI (see Table 6).

**Discussion**

In the sport science literature, several unanswered questions remain with regard to the testing and training of athletic agility. However, the theoretical framework proposed by Young et al. (2002), and modified by Sheppard and Young (2006), in combination with a clear definition of the agility construct (Sheppard & Young, 2006) aided in this investigation. The two primary components of agility have been classified as CODS and perceptual and decision making factors. Research has begun to identify which of the related secondary or tertiary factors contribute the most to agility (Hoffman, Rattamess, Klatt, Faigenbaum, & Kang, 2007; Wheeler & Sayers, 2010; Young et al. 2002), but several components of the framework remain uninvestigated. The primary purpose of
this study was to establish the reliability of a commonly used test of CODS (3CR), in addition to three modifications of the test, two of which could be considered agility, based on having the participant react to a stimulus. A secondary purpose was to determine the contribution of BMI (the anthropometry component in Sheppard and Young’s framework) and leg dominance (right/left muscle balance) to both agility and CODS. The need for this study is based on the lack of a scientific foundation; the limitations in the literature, and the inconsistent application of psychometric principles. These needs have been previously identified by Sheppard and Young (2006), reiterated by Haff (2007), and were emphasized even more following an extensive investigation of the literature as far back as 1932.

**Scientific Foundation for Agility**

The Sheppard and Young review (2006) only addressed a small set of testing protocols covering the last 35 years in their review of literature related to agility. The review conducted for the current study confirmed the findings presented in Sheppard and Young (2006), but extended the review to include findings from the 1930's until the present. This is important because some of the tests developed during this longer time period are still referenced and used for evaluating athletes. When combining the findings of both reviews of the literature it can be concluded that assessment of agility has historically been arbitrary, and lacked a solid scientific foundation.

The first component of a solid scientific foundation lies in the need for a comprehensive definition of agility. Sheppard and Young (2006) did an excellent job providing a limited view of the lack of consensus for a definition prior to suggesting the definition used for this investigation. An extended review for the current study found 24
unique definitions including three researchers, O’Conner & Cureton (1945), Cureton (1947) and Clarke (1950), who did include reaction when describing agility. Unfortunately, these early definitions were not adopted and several CODS tests were developed prior to, and even after, Chelladurai (1976) critiqued prior research while also developing an agility classification scheme that included both simple and complex agility (see Figure 1). However, this classification scheme is also limited in that, it still classified CODS tests as “simple” agility. Nearly 30 years later, and with no significant changes in the tests being developed, Sheppard and Young (2006) indicated the need for “a simpler definition of agility [that] could be established by using an exclusion criterion, rather than an inclusion criterion” (p. 922). The result is the framework used for this investigation (see Figure 2). It is hypothesized that this framework constitutes a substantial piece of the foundation for future agility research. However, the foundation is not complete without a second technical component involving motor learning principles.

The second foundational component of agility as a construct is testing and training based on motor learning theory. “If agility can be improved through practice, then agility should be subject to the same skill acquisition theories as any other sports skill” (Jeffreys, 2006, p. 72). Jeffreys suggested using a target based approach. Fundamentally this approach requires sport scientists to intensively examine the basic movements in every sport skill to identify target movement patterns for which tests can be created. The target movements are then broken down into discrete parts to be taught as a part of a whole complete movement used in competition. This approach fits nicely into the Sheppard and Young framework under technique. Jeffreys stated that movements should be described based on their function (initiation, transition, or
actualization). Initiating movements might include: cross-steps, first-step starts, drop-steps, or cut steps. Transition movements allow for reaction while maintaining balance and proper body position. These movements include: side shuffles, backpedals, jump cuts and chop steps. Transition movements are often executed over a short distance before the athlete moves to the actualization movement. Historically, transition movements have played a large role in testing of CODS. The CODS tests are closed skills in which the athlete must only learn the pattern of the test and complete the skills required never moving toward an actualized movement that closely mimics a crucial skill, or as often is the case a sprint (Jeffreys, 2006). An example of this process is a running back in American football. The back will begin in a parallel stance but immediately after the ball is snapped will take a drop or false step to initiate forward movement toward the line of scrimmage. Just before, during, and after the ball is handed to the back he will stutter-step, shuffle or jump cut in a transition movement while using his vision to react to the oncoming defense. Once a weakness is found he will accelerate in full sprint until the process is re-initiated by approaching another defender. Jeffreys focused on developing a training system that included each of the individual movement types as well as skill specific movements, but the same should be considered when coaches and sport scientists are developing new agility tests. Craig (2004) noted the central nervous system (CNS) governs over all of these movements by firing the muscles in the appropriate sequence required to perform the skill. Adding a reactive component to future drills and tests will require the CNS to improve the coordination of the basic motor pattern based on proprioceptive feedback (Craig, 2004). The current study was the first to use a manual cueing system to incorporate a reactive component. The cue was
simple to apply, but further research is necessary to determine the point at which the cue is initiated to better simulate the game action being represented.

**Limitations in the Literature**

Other areas of concern in the literature that this study intended to address are published inaccuracies, lack of specificity, and incomplete test administration. The first example of an inaccuracy comes from a recent study by Hoffman et al. (2007). The 3CR in Figure 1 of that paper (p. 129), indicates the runner proceeds to the left (clockwise) when circling cone C, when, in fact, the runner should be routed around the cone to the right (counterclockwise). It is likely an error in the design of the figure, but, as the only published model of the test, it could be confusing to an uninformed coach. In addition to the possible incorrect path being used in the Hoffman study, comparisons between this data and those of Hoffman et al. (2007) should be cautioned, due to the differences in the test, re-test protocols, and no mention of the statistical model used to calculated reliability. Another sample inaccuracy involves the Edgren side-step test, which is also referred to, without citation, in the *Essentials of Strength Training and Conditioning* text (National Strength and Conditioning Association, [NSCA] 2008). An original study by Edgren (1932), titled “An experiment in the testing of the ability and progress in basketball,” described, a “coordination” test that required the players to shuffle side to side across an eight foot lane with the score being the time to complete ten “shifts” across the lane. The test credited to Edgren in the *Essentials of Strength Training and Conditioning* text (NSCA, 2008) uses a 12-foot wide section and requires the participant to shuffle back and forth for ten-seconds; the number of three-foot sections completed in ten seconds is given as the score. No account for these discrepancies in the test protocol
can be found in the literature. Complicating matters further, two additional tests have similar descriptions; published in the New York State Physical Fitness Test (1972), side shuffling employs an eight-foot wide section, with the number of passes completed in 10 seconds used as the score; and the North Carolina Motor Fitness Battery uses a 12-foot section with the number of passes in 30 seconds being counted (Carolina, 1977). The current study has provided an accurate account of the original 3CR as well as two previously undescribed versions (3CL and 3CM). From this point forward, there should be no confusion by coaches and sport scientists as to how to appropriately apply these tests to the population they are studying.

The second limitation in the current literature, especially prior to the review done by Sheppard and Young (2006), involves the lack of sport-specificity. The four factors identified lacking specificity in the current CODS literature are stance, patterns run, skills used, and test duration. The starting stance is the first element that appears to lack specificity. Some tests begin with the athlete in a 3-point stance with one hand on the ground. This stance is only specific to starting positions in American football, most notably offensive and defensive lineman, who may use a variation of the stance and rarely move beyond a meter before making contact with an opponent. The two primary CODS tests used to evaluate players at the NFL Combine, the 3CR (Hoffman et al., 2007) and pro-agility tests (Sierer, Battaglini, Mihalik, Shields, & Tomasini, 2008), fail the specificity test due to using a 3-point stance. Based on the unfamiliarity of the stance to a mixed group of attendees in this study, and to limit any potential confounding results based on the ability to start from a 3-point stance, participants were permitted to use a 2-point stance during these trials. The same stance had to be used during all trials. Future
research should include stance variations; however, the recommendation of the researcher is to initiate the test using the stance used most often by the population being tested.

After the stance, the second criticism related to specificity addresses the patterns and skills used during the tests. The T-test, 3CR, and pro-agility tests are three examples of closed, non-reactive (un-cued or a lacking spatial or temporal challenge) CODS speed tests used by various high school (Duehring, Feldman, & Ebben, 2009) and professional strength coaches in the National Football League (Ebben & Blackard, 2001), the National Hockey League (Ebben, Carroll, & Simenz, 2004), Major League Baseball (Ebben, Hintz, & Simenz, 2005), and the National Basketball Association (Simenz, Dugan, & Ebben, 2005). In these tests, the athletes cover the various set patterns primarily using a forward sprint, with the exception of the T-test, which requires sprinting, side shuffling and back pedaling. These movements do not represent a comprehensive list of movements used in sport. Another component to each of these tests is touching the ground with the lead hand when a change in direction occurs. During no competitive athletic skill performance are athletes required to reach down and touch the ground, especially when changing direction in response to an opponent’s movement. Thus, these “traditional,” but arbitrary, test requirements lack specificity to any sporting task these tests purport to measure. The 3CM examined in the current study required the participants to make a transitional movement as described by Jeffreys (2006) that is more like the movement required during game play of any sort. Future investigations should focus on greater levels of specificity.

The final criticism lies in the specificity of the test duration. Iosia and Bishop (2008) found the average college football play lasts 5.23 ± 1.7 seconds, which is shorter
than the average 3CR run by Division III football players by nearly three full seconds. The average time of the pro-agility at the NFL Combine in 2004 and 2005 (Sierer et al., 2008) was over a second faster than the average time (5.23 s) for the skill players and slightly less for linemen. Therefore, this incongruity also contributes to the capricious nature of these assessments. For the purposes of this study, time specificity was not a consideration because the population used was not being tested for any specific event. However, the average time for the 3CM was approximately 8.6 seconds, which is significantly greater than the average college football play. In order to meet the specific physiological demands of many athletes the 3CM could be shortened by eliminating the first 2 steps in the test (up to cone B and back to the start line). This modification with the inclusion of a reactive component would be similar to the 'L' drill described by Gabbett et al. (2008), be more appropriate for college football athletes.

The primary purpose of many early investigations was to create tests with strong psychometric properties. Test construction is detailed in numerous test and measurement texts, and has not changed much since Bovard & Cozens (1938). However, many sport scientists, coaches, and physical educators have failed to go through the formal process when creating new tests. This is especially problematic when new agility tests are needed. Without both of the foundational components (agility construct definition, motor learning), agility testing will remain capricious when compared to components of protocols used to measure health-related fitness (i.e. VO\textsubscript{2max}, 1-repition max, body fat percentage, modified sit-n-reach). The process used to evaluate the literature for the 3CR for this study, and partially used for the creation of the 3CM, was described in Johnson and Nelson (1986). In the eight-step process of Johnson and Nelson, the investigator
must consider the following: (1) analyze the game the athletes-to-be-tested are playing; (2) select tests that measure the major skills used in the game; (3) create detailed administration and scoring procedures; (4) determine reliability of each test; (5) compute objectivity; (6) establish validity; (7) revise the test, as needed, based on the findings, and; (8) develop test norms. Many, if not all, of these steps have been ignored in the majority of published articles available to coaches. For example, the T-test was first described by Semenick (1990) in the National Strength and Conditioning (NSCA) Journal. The author addressed only Step #3, establishing exact procedures, but fails to provide any information on the other seven steps. Fry et al. (1991) also compared starters and non-starters after an off-season program that included the T-test as the measure of agility without providing any information beyond Step 3 as well. Yet, this test is provided as an example for agility testing in the both the first and second editions of the NSCA text, Essentials of Strength Training and Conditioning. Pauole, Madole, Garhammer, Lacourse, and Rozenek (2000) determined the T-test to be reliable (0.98) however the method could be questioned based on the lack of detail of the model used to calculate reliability. Pauole et al. (2000) indicate, “Intraclass reliability coefficients were computed for each test using ANOVA”. This method of analysis may result in inaccurate values and will be discussed in the next section. Pauole et al. (2000) also chose to validate the T-test with the hexagon which had not been shown to be reliable prior to their investigation; and therefore is a questionable choice as a criterion measure to determine validity. The T-test and hexagon remain in the third edition of the Essentials of Strength Training and Conditioning (NSCA, 2008) even after Sheppard and Young's review demonstrated that both T-test and hexagon should be more correctly identified as
CODS tests. Thus, acceptance of such tests has been legitimized professionally, but appears to be erroneously validated. These publications also either fail to provide complete procedures or use questionable choices when describing rest intervals.

Absent or inappropriate rest intervals also limit the psychometric value of the currently available tests. Many are missing any sort of suggested rest interval (Hoffman et al., 2007; Semenick, 1990; Wheeler & Sayers, 2010), have exceeding long rest intervals (Pauole et al., 2000; Sassi et al., 2009), or seemingly too short to apply to most sports applications (Beekhuizen et al., 2009). Pauole et al. (2000) and Sassi et al. (2009) both allowed the subjects they tested minutes between each test equivalent to ~17:1 rest to work intervals; while Beekhuizen et al. (2009) used a nearly 1:1 rest to work ratio. In the present study a 6:1 ratio was used based on the average time for the participants during the familiarization trials. This ratio is also appropriate for football athletes running 5-second plays with a 40-second play clock.

While the current study did not attempt create an entirely new test of agility for a specific sport thus missing the completion of steps 1 (game analysis) and 2 (movement selection) the remaining steps of test construction were applied except for the development of norms (step 8). It was the intent of the author to develop a test that met Sheppard and Young’s operational definition of agility while also providing Steps 3 (exact procedures, see appendix A) and 4 (reliability) in the test construction process. In addition, Steps 6 (validity) and 7 (item revision) were also covered on the basis of the 3CM now having face validity with the inclusion of reaction, and also being a revision of 3CR. After completing this process several practical implications were developed for design future agility testing methods.
Interpreting Results

Only one investigation has been published using the 3CR test (Hoffman et al., 2007), and while ICC's were reported, the determination of reliability was not the primary purpose of the investigation, as it is in the current study. This fact is important, because as Beekhuizen et al. (2009) noted,

Test-retest reliability is critical in establishing the reproducibility of a test and is a prerequisite for establishing validity in future investigations. Moreover, the reliability of a test is essential when being utilized as a predictive variable for decisions such as return to sport or activity. (p. 2170)

Similarly, the present study, as well as a future line of research designed to evaluate agility performance, should more accurately predict transfer of agility skills to game performance (validity). In terms of reliability, several different statistical models for calculating reliability have been used in the studies that actually determined it. Several do not provide any detailed evidence of the statistical model used (Gabbett, Kelly & Sheppard, 2008; Hoffman et al., 2007; Pauole et al., 2000). Weir (2005) addressed this problem, as well as, the importance of including the standard error of the measurement. Based on Weir (2005), Portney and Watkins (2008); and Beekhuizen et al., (2009) the current study applied the 3,1 model used for a single rater. Using this model, the primary finding of this study was that the 3CR test and its multiple variations were determined to be moderately reliable. Intraclass correlation coefficient values ranged from .73 to .83, with SEM values between 0.25 and 0.34. The ICC values observed in this study fall below the reported values of similar tests of CODS and agility, and; well below the ICC of 0.90 (no SEM available) reported by Hoffman et al. (2007). When examining the
results of reactive tests (i.e. 3CMR and 3CML), the best comparison may be the findings of Serpell et al. (2009). They found an ICC of 0.82 for their reactive test. Additionally, Gabbett et al. (2008) reported ICC 0.92 for their version of a reactive agility test (RAT). Nonetheless, any measure of true agility will be difficult to objectively validate without a gold standard to compare it to. However, if coaches and sport scientists follow the test construction and statistical procedures outlined in the current study soon there will be a larger set of highly reliable tests to choose from. After using the appropriate procedures up to this point, the final step is to interpret the results and suggest adjustments for future testing.

The interpretation of the results can be divided into three primary aspects of any skill analysis: participants, environment, and task. This is also step 7 of the test construction process. Revisions should be based on multiple factors, including, but not limited to: (a) statistical analysis of reliability and validity; (b) ease of administration and scoring, and; (c) advances in technology with relationship to measurement (Johnson & Nelson, 1986).

Having addressed reliability, and satisfying the elements of the definition proposed by Sheppard and Young (2006) (face validity), ease of administration and scoring will be described as part of the analysis of the test administration process including participants, the environment and the task itself. In addition limitations of the study as well as future directions utilizing advances in technology will be suggested.

The first area for discussion is the administration of these tests, as well as other CODS and agility tests. Comparisons between the findings of this study and those of Hoffman et al. (2007) should be cautioned for several reasons. While not the specific
purpose of the Hoffman et al. (2007) investigation, “test re-test correlation coefficients were R > 0.90 for all power and agility tests” (p. 128) were reported with no explanation of the statistical methods used to calculate them. Secondly, the populations examined were not congruent. Their sample included sixty-two National Collegiate Athletic Association Division III (DIII) American football players on the first day of pre-season practice, while the current study used a subject population comprised of otherwise-healthy college-aged students taking courses in a Department of Physical Education. Though 65% of the sample were college athletes, none of them were actively preparing for the immediate start of their season.

The subjects in the Hoffman et al. study (2007) were all tested in one day and ICC”s were calculated based on three trials. The current protocol included 10 total trials (4 reactive and 6 non-reactive) each day separated by a minimum of 48 hours. The present study protocol also included a familiarization phase to minimize any potential learning effect, which may have skewed the data. This was appropriate for the population of students tested in the current study, but was not deemed necessary in a population of DII football players studied by Hoffman et al. (2007). Based on the reported ICC of > 0.90, the participants in Hoffman et al. (2007) were likely very familiar with the test. The fact that they were tested on the first day of pre-season practice may also indicate a certain level of comfort with the 3CR, perhaps due to practicing it during off-season training sessions. That is contrasted to the present subject population of students comprised of athletes and non-athletes. Very few had any knowledge of the test and the rest were first introduced to this test during the present study.
The inter-test correlations in the current investigation were high (r=0.84-0.91), indicating a close relationship between the tests and, therefore, no need to complete all three versions of the 3CR. However, agility testing should be done only with the 3CMR and 3CML; these assessments require a reaction to a visual cue, unlike the 3CR and 3CL, which test CODS. The average time on the reactive tests was 0.2 seconds greater than the CODS test on both days. While the difference may not be statistically significant, this may have a practical impact during competition. The current study was not designed to determine the impact reaction has on participant technique, but such a finding could be further evaluated using a segmented calculation of the total time. This could be determined using the split function on a stop watch or through the use of video analysis. If using a stopwatch, the tester would need to press the lap function at each COD point and then stop the watch at the end of test. Using video analysis software such as Dartfish™ would allow for frame-by-frame time referenced analysis of the video.

This is the first study to make direct comparisons between identical tests, with and without reaction. With no other available comparisons, it may be concluded that reaction remains a critical element of agility; however, the method of examining this difference was not measurable using the 3CMR or 3CML. A possible reason for similar results between the reactive and non-reactive version of the 3CR is the cue being given too early allowing the participant to react and change direction at nearly the same velocity they were running at during the non-reactive test. In addition, more than a single reactive component may be necessary. It also may be necessary to calculate segment times (including reactive time) instead of total test time to hone in on exactly where time was lost between trials. For example, a runner may get a bad start or lose balance during
another change of direction unrelated to the reactive portion of the test. In that case total
time would not be representative of the impact reaction has on executing the test as fast
as possible. Lastly, the use of a more sophisticated cueing system timed more closely to
a specific game situation may be required. The manual cue given in the current study
was at a distance equal to 2.29 meters from the required change in direction. This
distance may have contributed to the lack of difference between the reactive (3CMR and
3CML) and non-reactive tests (3CR and 3CL). A computerized cueing system like is
being used in the new generation of reactive tests could randomly provide the visual cue
at various distances from the point of change of direction.

Only six subjects (15%), who self-selected left-foot dominance, participated in the
present study. Thus, it was problematic to formulate any conclusions regarding closed-
chain reaction involving the dominant leg versus the non-dominant leg. The percentage
of left-sided subjects in the current study coincides with published estimates of left-
handedness (Porac & Coren, 1981); unfortunately, insufficient statistical power did not
permit comparisons between right-handed and left-handed subjects. Without statistically
equivalent groups, it would be difficult to analyze the effect of right or left dominance on
agility especially on certain movements that are dominated by one side of the body. For
example, the 3CR or 3CMR requires four-out-of-five closed-chain directional changes
from the right leg. In sport, agility is not so heavily sided; this is recognized as right and
left leg balance in Sheppard and Young's model (see Figure 1), in addition to reactive
and concentric strength and power (Sheppard & Young, 2006). Potentially, these
variables could have been examined in this study, but Hoffman et al. (2007) previously
found that “bilateral power differences do not significantly affect direction-specific

23
movement patterns” (p. 130). This finding indicates further research is needed on the other variables in the model, possibly using a using a factor analysis approach. One such factor is anthropometry.

In this study, anthropometrics was quantified as the ratio of the participant weight to height squared (kg/m$^2$), or BMI. BMI is used to estimate body fatness, and, in the current study, a moderate to strong relationship was observed between BMI and the best trial from each test (see Table 6). The values represent 49-53% of the variance in the fastest trial of each test. Chaouachi et al. (2009) reported an even higher correlation (r=0.80) to the T-test when estimating body fat percentage using a four site skin-fold equation. The remaining factors (visual scanning, knowledge, pattern recognition, anticipation, and technique) in Sheppard and Young's model will also need to be tested. In fact, Wheeler & Sayers (2010) examined running technique during pre-panned versus reactive testing. They have reported “reactive conditions … altered agility running technique” (Wheeler & Sayers, 2010, p.450). Technological advances in instrumentation may permit more precise outcomes of these perceptual and decision-making factors.

**Limitations of the Present Study**

The first limitation associated with this study is the use of a convenience sample which is a threat to external validity and ultimately generalizability. The findings are not likely to be generalizable to a sample of athletes or to female participants. Unfortunately, coaches today often ignore the specificity principle and apply an agility test to a population of athletes other than that for which the protocol was designed.
The second limitation was the consistency of maximum effort. The participants were asked to refrain from any strenuous activity the day prior to testing; however, the participants were tested during the weeks leading up to and during final exam week; the possibility exists that subjects were not well-rested. This is only speculative, because no participants displayed any outward signs of fatigue or complained of being tired. However, it became obvious once the data were analyzed that a few individuals provided less than maximal effort levels on certain trials. In five of the 40 subjects (12.5%), test times decreased greater than one second from trial 1 to trial 2 for at least one test version. The test results from these participants had a significant impact on the ICC values. These types of large discrepancies may necessitate re-testing or potential removal from the data set, due to an apparent lack of consistent effort.

In the absence of a "gold standard" for agility, validation of the tests in the current study is limited to face validity. The determination of validity was not the primary focus of the present study; however, based on a required agility construct that “a rapid whole-body movement with changing velocity or direction in response to a stimulus” (Sheppard & Young, 2006, p. 922), the 3CR and 3CL are deemed reliable, but not valid agility tests. Instead these should be labeled as reliable CODS tests. The two variations of the test, 3CMR and 3CML, do appear to be reliable options for measuring agility. The next step will be to validate the tests.

The final limitation of this study was the choice of a 12-inch cone. This is a tall cone that was easily dislodged as the participant navigated around the 3C pattern. While the current investigator does not expect this impacted the results, any repeated trial of a modified test would unintentionally reveal the direction to the participant. To the
researcher’s knowledge, no participant used this as a strategy to try to gain an advantage during any retrial. Failed tests due to a disrupted cone did occur often enough that using a flatter disc type cone is recommended for future investigations. The use of cones could be eliminated all together with the development of programmable gate systems.

**Future directions**

The current list of test options, inclusive of reaction, is quite limited. The majority of these reactive agility tests (Gabbett et al., 2008; Serpell et al., 2009; Sheppard et al., 2006; Wheeler & Sayers, 2010) employ expensive timing gate systems and no investigation has compared an existing test to a modified agility test version that uses a manual source of reaction, which was the purpose in the present study. Only in the last five years has there been any advancement in the use of technology when timing or analyzing tests of agility. Sheppard et al. (2006) were the first to develop a reactive test, i.e. RAT that utilized technology; however, the test has a low degree of specificity. The apparatus investigated involved the tester standing on a timing mat. The participant ran toward the tester and, at a certain point, the tester made one of four movements that required the participant to react and run either to the right or left. Timing gates were set to the right and left of the tester and the time was recorded from when the tester stepped off of the timing mat until the runner passed through the gate. This test appears to be reliable based the results, and was able to discern between senior league and reserve groups of rugby players, which is a form of concurrent validity. This test, and the use of appropriate technology, appears simple to use and adaptable to any sport setting. In fact, a second study by Gabbett et al. (2008) also determined reliability of the RAT, along with three other non-reactive measurements (i.e. 505, modified 505, and the 'L' test).
Interestingly, only the RAT was able to determine differences between first and second grade rugby league players demonstrating concurrent validity.

Serpell et al. (2009) has developed the most technologically impressive methodology to measure sport-specific agility under game-like conditions. Using a very similar arrangement to that of Sheppard et al. (2006) and Gabbett et al. (2008), Serpell exchanged the tester on the timing mat for a live video image of a rugby player displayed on a screen. Participants are instructed to run toward the screen, and, at a certain point, break the beam on a laser gate system, which, in turn, starts a video clip of a would-be tackler. The participant then changes direction, as they would normally do in a game situation, to avoid the on-screen “opponent.” The system is able to calculate stages (start to reaction, reaction to completion) of the test.

Galpin, Li, Lohnes, & Schilling (2008) also used a technologically novel approach to measure agility. Using the Quick Feet Board, by The Quick Board, LLC, active, but non-agility trained, men and women were tested. The board was connected to a power source control device that illuminated a set pattern to be followed. The subjects were pre and post-tested four weeks apart. This agility test was found to be highly reliable (ICC=0.89), and the subjects lowered their testing times significantly from pre to post. This method of testing follows the construct of agility; however, it fails the specificity requirement because it was stationary mat in which the subjects moved their feet rapidly touching 10 random spots.

Based upon these studies and the present investigation, one may hypothesize that the combination of Sheppard and Young's model (2006), advanced programmable gates, and video-based skill analysis of sport-specific agility movements, a new generation of
scientifically-based testing, teaching, and coaching of agility may soon come to fruition. In microcosm, such evolution may provide the impetus for coaches and trainers to use instrumentation based on science, not simply on tradition.

There are many paths this line of research can take, but the final destination should be a set of valid, sport/position specific, developmentally and possibly gender appropriate tests of agility. Recent evidence supports the use of increasing levels of technological inclusion. The use of technology allows for limitless variations in test patterns that challenge both the temporal and spatial components described by Chlelladurai (1976) and Sheppard and Young (2006). Video from actual game action can be replicated for testing and training purposes, as Serpell (2009) demonstrated using practice clips. It is likely that, in the not too distant future, virtual training centers could be developed where athletes interact with 3-dimensional images as part of their testing and training programs, much like a live-action video game. On the other end of the spectrum, alternative “field” accessible tests of agility must also be developed for practitioners at the youth sport and physical education level. These tests must include the same characteristics described above and should ultimately be derived from the valid tests that utilize appropriate technology.

Conclusions

Agility has many interlocking pieces and, as researchers begin to build from each corner of the puzzle (i.e. Young's framework), the necessary factors for developing agility should become evident. Once the “puzzle” is completed, the next step will be to determine the relative impact of each piece to the puzzle. In this study, two factors from Young’s conceptual framework were examined anthropometry (BMI) and visual
scanning (manual cuing). While neither of these were the primary purpose of the study, they do provide useful insight into future test designs. It is an exciting time to be involved in this area of research. However, according to Stone, Sands and Stone (2005), there are very few opportunities for the development of sport scientists who would be interested in basing their career in part to researching sport-specific assessment. This can often be daunting task, as Stone et al. (2005) discuss, with the “disincentives” (i.e. few jobs, lack of grant availability, limited access to elite or near-elite subject populations, and the difficulty publishing sometimes unconventional study designs) few graduates of kinesiology programs choose sport science and fewer specializing in testing and measurement. In order to validate sport science in terms of assessment, antiquated and generally applied tests should be replaced with sport or position-specific performance assessments. The development of such tests will also create an opportunity for examining different training methodologies for the purpose of improving an athlete’s in game agility performance. The current study provides the methodological guidelines for proper test construction and, with a few revisions, will represent a protocol researchers can use as a template to create the valid, sport-specific performance tests of the future.
References


trained, but active men and women. *Journal of Strength and Conditioning Research, 22*(6), 1901-1907.


New York Department of Education (1972). *New York State Physical Fitness Test for boys and girls grades 4-12.*


### Table 1

*Participant Characteristics*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>20.97</td>
<td>± 1.67</td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>85.07</td>
<td>± 18.27</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>181.26</td>
<td>± 6.11</td>
</tr>
<tr>
<td>BMI (kg·m⁻²)</td>
<td>25.88</td>
<td>± 5.39</td>
</tr>
<tr>
<td>Foot kick with</td>
<td>Right (n=34)</td>
<td>Left (n=6)</td>
</tr>
<tr>
<td>College athlete</td>
<td>Yes (n=26)</td>
<td>No (n=14)</td>
</tr>
<tr>
<td>Varsity High School athlete</td>
<td>Yes (n=38)</td>
<td>No (n=2)</td>
</tr>
</tbody>
</table>

M= Mean, SD= Standard Deviation
Table 2a

*Balance Latin Square for Determining Test Block Order*

<table>
<thead>
<tr>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>L</td>
<td>M</td>
<td>R</td>
</tr>
<tr>
<td>M</td>
<td>R</td>
<td>L</td>
</tr>
</tbody>
</table>

Table 2b

*Balance Latin Square for Determining Trials of the 3-cone Modified*

<table>
<thead>
<tr>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>L</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>R</td>
<td>R</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>R</td>
<td>R</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>
Table 3

Performance Characteristics and Reliability Coefficients

<table>
<thead>
<tr>
<th>Tests</th>
<th>Best Trial 1 M (secs) ± SD</th>
<th>Best Trial 2 M (secs) ± SD</th>
<th>Mean Difference (secs) ± SD</th>
<th>ICC (95% CI)</th>
<th>SEM Trial 1</th>
<th>SEM Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3CR</td>
<td>8.41±0.65</td>
<td>8.38±0.64</td>
<td>-0.03±0.40</td>
<td>0.81 (0.66-0.84)</td>
<td>0.287</td>
<td>0.279</td>
</tr>
<tr>
<td>3CL</td>
<td>8.42±0.66</td>
<td>8.35±0.56</td>
<td>-0.07±0.45</td>
<td>0.73 (0.54-0.85)</td>
<td>0.343</td>
<td>0.293</td>
</tr>
<tr>
<td>3CMR</td>
<td>8.63±0.61</td>
<td>8.58±0.64</td>
<td>-0.04±0.36</td>
<td>0.83 (0.71-0.91)</td>
<td>0.250</td>
<td>0.260</td>
</tr>
<tr>
<td>3CML</td>
<td>8.61±0.58</td>
<td>8.63±0.59</td>
<td>0.02±0.37</td>
<td>0.80 (0.65-0.89)</td>
<td>0.262</td>
<td>0.266</td>
</tr>
</tbody>
</table>

M=Mean, SD= Standard Deviation, ICC= Intraclass correlation coefficient, SEM= Standard error of the measurement.
Table 4

*Pearson's Correlation (r) Matrix for the Fastest Trials on Day 1.*

<table>
<thead>
<tr>
<th></th>
<th>3CR</th>
<th>3CL</th>
<th>3CMR</th>
<th>3CML</th>
</tr>
</thead>
<tbody>
<tr>
<td>3CR</td>
<td>1.00</td>
<td>0.91**</td>
<td>0.92**</td>
<td>0.90**</td>
</tr>
<tr>
<td>3CL</td>
<td></td>
<td>1.00</td>
<td>0.82**</td>
<td>0.86**</td>
</tr>
<tr>
<td>3CMR</td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.90**</td>
</tr>
<tr>
<td>3CML</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.05 level (2-tailed)**
Table 5

*Pearson's Correlation (r) Matrix for the Fastest Trials on Day 2*

<table>
<thead>
<tr>
<th></th>
<th>3CR</th>
<th>3CL</th>
<th>3CMR</th>
<th>3CML</th>
</tr>
</thead>
<tbody>
<tr>
<td>3CR</td>
<td>1.0</td>
<td>0.89**</td>
<td>0.87**</td>
<td>0.87**</td>
</tr>
<tr>
<td>3CL</td>
<td>1.0</td>
<td>0.84**</td>
<td>0.85**</td>
<td></td>
</tr>
<tr>
<td>3CMR</td>
<td>1.0</td>
<td></td>
<td>0.84**</td>
<td></td>
</tr>
<tr>
<td>3CML</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.05 level (2-tailed).
Table 6

*Pearson’s Correlation (r) for BMI and Fastest Overall Trial*

<table>
<thead>
<tr>
<th>BMI</th>
<th>3CR</th>
<th>3CL</th>
<th>3CMR</th>
<th>3CML</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.71</td>
<td>0.70</td>
<td>0.73</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Model of Agility (Chelladurai, 1976)
Figure 2. Conceptual Framework of Agility (Sheppard & Young, 2006)
Figure 3. 3-cone right (3CR)

- Step 1
- Step 2
- Step 3
- Step 4
- Step 5
- Step 6

Tester
A1 A2

4.57 meters
Figure 4. 3-cone left (3CL)

- Step 1
- Step 2
- Step 3
- Step 4
- Step 5
- Step 6

4.57 meters
Figure 5. 3-cone modified (3CM)

--- Step 1

--- --- Step 2

--- --- --- Step 3
Figure 6a. Starting position of the cue for 3CM.

Figure 6b. Position of cue for 3CML when directing the runner to the tester’s right.
Appendix A

Test Re-Test Reliability of Three Versions of the 3-Cone Test
Administration of Three Versions of the 3-Cone Test
Test Administration of the 3CR, 3CL, and 3CM

I. Equipment needed
   a. 5, 12-inch cones or discs
   b. Stop watch (1/100th)
   c. A 12.8m x 15.24m flat surface
   d. Tape measure
   e. One timer and one recorder

II. Protocol
   a. Cone arrangement (3CR, (see Figure 3)
      i. The 3CR is arranged by creating a triangle or an upside down L-shape- setting up a start/finish line (A1 and A2), one cone is 4.57 meters (m) straight ahead (B), and a second cone 4.57m to the right of B (C). If placed in the correct spot C should be 6.675m from A. Participants start in a two point stance (step 1) sprint to B where they touched the ground with their hand, (step 2) they then return to A and touch with their hand before heading back, (step 3) toward B where they maneuvered around the outside of B (step 4) then weaved counter-clockwise around cone C, (step 5) and then back toward and around the outside of B (step 6) and finishing through the line at A.
   b. Cone arrangement (3CL, see Figure 4)
      i. The 3CR is arranged by creating a triangle or an upside down L-shape- setting up a start/finish line (A1 and A2), one cone is 4.57 meters (m) straight ahead (B), and a second cone 4.57m to the left of B (D). If placed in the correct spot D should be 6.675m from A. Participants start in a two point stance (step 1) sprint to B where they touched the ground with their hand, (step 2) they then return to A and touch with their hand before heading back, (step 3) toward B where they maneuvered around the outside of B (step 4) then weaved clockwise around cone D, (step 5) and then back toward and around the outside of B (step 6) and finishing through the line at A.
   c. Cone arrangement (3CMR or L, see Figures 5, 6a and 6b)
      i. The 3CM test is a combination of 3CR an 3CL, having both the C and D cones as possible running patterns in the shape of a T. If placed in the correct spots C and D should be 6.675m from A. The modifications to the 3CM are specific to step 3 of the 3CR. During this step the runners react to a stimulus provided by the timer standing 4.11m directly in front of them or 1.829m from B.
During step 3, the runner’s turn to face cone B, as the runner passes the 2.29m mark on the floor the tester will point to either cone C or D. The participants then completed steps 4-6 in the same pattern as the 3CR or 3CL depending on the direction they headed. The directional signal used began with the fists of the tester together and both thumbs pointing upward directly in front of the chest of the tester (see Figure 6a). If the runner was assigned to go to their left, the tester pointed with his right thumb; and vice versa if the runner was to go right (see Figure 6b).

d. General warm-up ~10 minutes
   i. 2 laps around gym
   ii. Second lap high knee to half-court, heel-ups back
   iii. Walking toe touches
   iv. Walking ER
   v. Walking IR
   vi. Walking lunges
   vii. Leg swings-hip flexion/extension
   viii. Leg swings- hip abduction/adduction
   ix. 10-sec static hamstring
   x. 10-sec static quadriceps
   xi. 10-sec static gastrocnemius
   xii. 10-sec static soleus
   xiii. 10-sec static tibilais anterior
   xiv. Personal stretch

e. Specific warm-up ~6 minutes
   i. 1 each (3CR,3CL,3CM) at 50%
   ii. 1 each (3CR,3CL,3CM) at 75%
f. Rest interval
   i. 6:1 rest to work ratio for full recovery
g. Scoring
   i. The timer will start the watch on the participant’s first movement and stop the watch as the participant passes the finish line.
   ii. During the 3CR and 3CL, the timer is stationed near the start finish line, opposite the direction the participant is planning to go. This will provide a physical reminder to run the most direct line to the finish after rounding cone B.
   iii. During the 3CMR and 3CML, the timer begins stationed on the cueing spot 1.829m directly in front of B. When the participant is between cones B and C or D the timer will move to the finish line opposite the direction the participant is cued toward.
   iv. A participant should be disqualified if they:
1. Fail to touch the ground with their hand when required at cone B in step 1 and at cones A1 and A2 after step 2.
2. Touch the ground with their hand at any other point during the test.
3. A cone is upset resulting in it moving from its assigned location.
4. During the 3CM versions, fail to go the cued direction or do not go around the outside of cone B during step 3.

III. Tips
   a. Indoor testing
      i. Trace a corner of each cone on the floor with a permanent marker. This will allow for a more efficient and accurate replacement of the cones if they are upset during testing.
      ii. Using the volleyball lines on a basketball court are helpful when setting up the cones C and D. The court is 30 feet wide so it is easy to replace the cone to the spot marked on the line if B is in the center of the court.
   b. Outdoor testing
      i. Trace a corner of each cone on the grass or turf with a temporary volleyball corner. Cones can also be secured on a grass surface with stakes.
      ii. The marked lines on a football field can be used to mark the distances between b and C or D.
Appendix B

Test Re-Test Reliability of Three Versions of the 3-Cone Test
Review of Literature
REVIEW OF LITERATURE

Definitions of Agility

Prior to the construction of any evaluation technique there must be a commonly agreed upon definition of the construct that will be tested. To date no universal definition of agility has been reported in the scientific or coaching literature. More than fifteen definitions have been published over the last 60-years (Table 1). Each of these definitions is unique in its components as well as its choice of descriptors; however, there are many obvious similarities as well. Four components are found in the majority of the definitions (1) precision/accuracy, (2) change of direction (COD), (3) body (parts), and (4) quickness/rapidity (Table 2). It would appear that while there is distinctness in all of the definitions listed above; the common components of the definition are not in dispute. Therefore, the author has selected the proposed definition provided by Sheppard and Young (2006) for the purpose of evaluating previous tests as well as to consider when constructing a new test.

These commonalities should not be seen as all inclusive of the construct of agility though. Ellis et al. (2000) identified that the definitions (testing protocols) prior to 2000 are limited to horizontal movements, which require no “visual processing, timing, reaction time, perception, and anticipation.” Response to unexpected circumstances was a part of Clarke's (1950) definition as far back as 1950, but it took another 25 years for Chelladurai (1976) to highlight the fact that as of late 1970’s no attempts were made to develop tests that included the “perceptual and decision-making factors” that the majority of athletic movements require during actual game play (Sheppard & Young, 2006, p. 921). Based on his evaluation Chelladurai developed a more inclusive definition and
followed that up with a test of agility to confirm his classification scheme. Based on a review of the literature, including the addition of new components published in the 1980’s and 1990’s, Sheppard and Young (2006) agreed with Chelladurai's (1976) conclusions. With the addition of change of velocity Sheppard and Young (2006) has put forth the most complete definition to date. Without this definition the impact of agility has likely been greatly understated, and the currently available tests and drills for improving agility do not allow for valid assessment, identification of deficiencies, or task transfer to game situations.

**Conceptual Framework for Agility**

Recognizing the impact agility can have on motor skill development as well as athletic performance a model has recently been developed. This model, developed by Sheppard and Young (see Figure 2), is based on the previous classification scheme proposed by Chelladurai (see Figure 1). The hierarchal classification scheme (Table 3) described by Chelladurai is related to Knapp's (1963) continuum of skills and Robb's (1972) definitions. Knapp and Robb described the characteristics of an open or closed skill; these can be considered synonymous with Chelladurai's description of simple and complex skills. The key to Chelladurai's system of classification is reaction to a stimulus. It can be either simple or complex which is further divided into temporal or spatial responses. Temporal reaction requires the performer to only respond to a variably timed signal to initiate movement. Spatial reaction is a response to COD or distance. The absence of either of these reactive components is classified as simple. The presence of both stimuli is classified as universal. An example of a complex agility task based on the classification scheme is the pass rush of a defensive lineman. The player begins in
his stance, waits to react to the ball, and based on the defensive call will attack the quarterback on a predetermined path to his right, left, or straight ahead in an attempt to beat the offensive lineman across from him (temporal). Upon successful completion of this step the defensive lineman finds the quarterback and reacts to the distance and direction of his location in the pocket (spatial). The complete execution of rushing the passer is a universal agility skill that requires both components of temporal and spatial reaction. Chelladurai used this scheme to design and test a set of agility tasks that will be described in a future section.

The 2006 review of agility by Sheppard and Young detailed Chelladurai’s classification scheme as well as their proposed framework, which was first published in 2002. The authors note that while Chelladurai’s own definition of agility contains the phrase “response to a stimulus,” his “simple” skills do not require the reaction to a stimulus but rather are closed skills. The critique of the scheme was it might be confusing to coaches and sport scientists because it was too inclusive. Tasks such as the long jump and shot put were classified as simple agility when in actuality they should not be included. This led to a similar yet more exclusive definition “rapid whole-body movement with change of velocity or direction in response to a stimulus” (Sheppard & Young, 2006, p. 922), and a slightly modified conceptual framework (see Figure 2). It will be this definition that will be used for comparison purposes and in the new developed of tests.

As the definition implied, the newly proposed framework contained two major arms: CODS and perceptual and decision-making factors. Each of these factors has additional attributions that affect an athlete’s ability to demonstrate high levels of agility.
For example, perceptual and decision-making factors require situational knowledge, pattern recognition, visual scanning and anticipation. Likewise, CODS relies on anthropometrics (body shape), technique, straight sprinting speed, and leg muscle qualities (Young, McDowell, & Scarlett, 2001). These factors more adequately cover how agility manifests itself in sport. Knowing that both components are essential for superior agility it is the task of a coach or physical educator to evaluate deficiencies and then design tasks that will improve it. Unfortunately, while there are many testing options in both physical education and sport very few of the currently available tests include reactive components but rather focus singularly on evaluating CODS.

**Common Agility Assessments**

**Physical Education**

In the early 1900's the focus of many investigations was on motor ability. Many of the theses and dissertations of the graduate students in master’s and doctoral programs were designed to develop batteries of tests that evaluated student abilities. Often included was a test of COD. Cozens' (1928) dissertation entitled “The measurement of general athletic ability in college men”, which includes seven elements that “52 physical educators … deemed most important”. Two elements of interest are the combination of body coordination, agility, and control and secondly, speed of legs. The tests used for each of these elements were the bar snap for distance and the dodging run (see Figure 3).

In 1932, Alden, Horton, and Caldwell (1932) followed up Cozen (1928) with a similar ability test for women. The specific purpose was to “set up a battery of tests which can be used to classify freshman college women according to their motor ability”. Overall, eleven elements of bodily control and 14 objective tests were identified through
a survey of 20 physical educators. The “through window ladder” was selected as representative of whole body agility. This unique test involved climbing up though three windows and back down and past the starting line three feet from the ladder. The test is started with a “GO!” command and total time is the measurement. No required technique was described. Using 131 participants, a reliability coefficient of 0.69 and a correlation coefficient of 0.37 were found. The test was correlated with the composite score of all 14 tests. This correlation was an attempt to assess the validity of the test. The authors also correlated the score on the through window ladder with two course scores from beginning tennis (r=0.61) and dancing (r=0.33). A replication of the study was completed with a new set of students the following year and the reliability coefficient was identical r=0.69 but the correlation increased to r=0.64. The second year a 40-yard obstacle race was added the reliability (0.68) and correlation (0.69) were similar to the through window ladder. The tests were again correlated with activity course scores: modern sports and stunts (0.81 with 40-yard maze and .40 with through window ladder), basketball (0.56 with 40-yard maze and 0.84 with through window ladder) and dancing (0.75 with 40-yard maze and 0.61 with through window ladder). While climbing through a ladder had a moderate reliability and correlation to the composite score it appears that it would not be the best measurement of agility, even with a 0.84 correlation with basketball. The 40-yard maze had multiple changes of direction and is designed more like the dodging run and other more common measures of agility.

Five years later, at the State University of Iowa Kathryn E. Young was completing her Master's Thesis entitled “An analytic study of the tests of COD” (1937). The focus of her investigation was a factor analysis of 16 tests “of the elemental
components of physical activities” (Young K. E., 1937). Based on the factor analysis run and the author’s interpretation, four of the tests were classified under the classic definition of COD. The Burpee (see Figure 4), zigzag run (see Figure 5), potato race (see Figure 6), and Howe exercise (see Figure 7) all loaded under factor IV (agility). The potato race and Howe exercises were described in this investigation for the first time. The potato race is not a potato sack race, but rather a shuttle test that involves running between two lines 10 yards apart. A pair of blocks is taken from one line to the other and then return to their original placement. This test is a longer version of the AAHPERD shuttle run (see Figure 8) used in the fitness battery developed 20 years later in which the blocks were only moved from one line to the other and not back to the original line. The Howe test was is used to measure whole body agility in a relative small area (six feet) in both the vertical and horizontal planes. There is a lack of information related to this test as well as the potato race. No reliability coefficients are provided for the tests which does not allow for a more detailed analysis.

Johnson (1960) was also interested in determining the fundamental motor skills of elementary children in a physical education environment. His dissertation examined a large variety of skills and across grades 1 to 6. His version of an agility test (see figure 9) was a third generation modification of test first described in 1934 (Johnson L. W., 1934) as a basketball dribbling test (see figure 10). McCloy and Young (1954) then changed the test into one that measured running agility in junior high children. The test consisted of running around a set of “four hurdles chairs spaced six feet apart, with the first hurdle placed 12 feet from the starting line” (Johnson R. D., 1960, p. 2). The score was based the time to complete one trip. Based on pilot studies and the recommendations of
teachers in the schools he would be testing he replaced the hurdles with folding chairs, placed the first chair only six feet from the start line instead of 12, and placed an “X” four feet high on the wall, six feet from the last chair. Students were required to zig-zag through the chairs touch the “X” and return back to the start line running again between the chairs (see Figure 5) as fast as possible. He found that three trials were most appropriate for this age group after determining subsequent trials were slower than either trials two or three. The following findings are based on a total of 4,654 boys and girls in grades 1-6. Test-retest reliability was determined by an unspecified method. The authors failed to describe the statistical model used as well as the time frame between the examined trials. The test proved highly reliable for both boys and girls for grades 2-6, with r values ranging from 0.79 to 0.95. The values for grade 1 were low (0.56-0.60, boys and girls, respectively), but all were statistically significant at p=0.01. Concurrent or criterion validity was also calculated based on comparison between time to run the test and a subjective ranking of agility by a teacher. The author fails to identify which specific correlations were significant, but appears that none of them achieved significance which would invalidate the test based on the criterion used. With the large number of subjects the author was able to create normative tables by grade and gender.

**Sport Performance**

In 1932, Edgren published, *An experiment in the testing of ability and progress in basketball*. As a basketball coach, Edgren was interested in devising a set of tests that would predict playing ability. Thirty members of a beginning basketball class were run through a battery of eight skill specific tests, four tests of general athletic ability, which included: agility, endurance, speed and coordination and the Brace Motor Ability Test.
Game playing ability was used to validate the items used. The group of skill tests correlated highly (0.77) with playing ability, and the group of general athletic ability tests (0.52). Correlations between the individual skill tests and playing ability ranged from -0.13 and 0.64. Likewise, between individual general athletic ability tests and playing ability ranged from 0.50 and 0.72. In contrast, the Brace Test had a low correlation with general ability and playing ability (0.14 and 0.16, respectively). Each of the general athletic ability tests is of interest. The “speed” test, Cozen's dodging run (see Figure 3), was classified as a test of agility and described above, the “agility” test was Sargent’s test (vertical jump), which measures lower body power, and the “coordination” test was the side-step test, often used to measure agility. Presently, Edgren is credited (National Strength and Conditioning Association, 2000, p. 300) with a side-step test (see Figure 11) however it is different from the test described in this early paper (see Figure 12). Edgren appears to be referencing Cozen’s (1928) classification of the dodging run as a test of the “speed of legs”.

At the University of Michigan, Beise and Peaseley (1937) were also interested in the speed, agility, and reaction time (S-A-R Test) for the sport skills of tennis, golf, and archery. They used a very unique custom reactive apparatus in order to measure each of the characteristics. A two-plate system was attached to a kymograph that recorded the reaction time to a light stimulus, the time between plates I and II (speed), and time to cover a path around three high stools back from plate II back to plate I (agility). This is an extreme departure from the other tests being devised and evaluated during this time in history. No other test had a reactive component. No other test was timed using an objective recording device that measured to the hundredth of second by using a
A kymograph equipped with an “electronic tuning fork tuned to the correct number of vibrations per second” (Beise & Peaseley, 1937, p. 135). Unfortunately the small sample sizes allowed for only a single statistical comparisons were prepared between skilled VS unskilled (t=6.79 significant difference). The authors only present mean times for each of the components between the three sporting activities. Apparently the method used to collect data was either extremely difficult to set-up or to produce reliable data because no further uses of the S-A-R Test are found in the available literature.

Draper and Lancaster (1985) were the first to attempt to create a test that would isolate acceleration and minimize the impact of velocity in the measurement of agility in a horizontal plane in a dynamic state. Two tests were designed to answer this question. In order to determine the effect of speed and acceleration the 18 participants performed a 20-meter (m) sprint. In addition they completed the up and back test (UAB) (see Figure 13) and during the same run a 505 (see Figure 14) measurement was calculated using video strobe analysis with a built in timer. For comparison purposes the participants ran the Illinois agility test (see Figure 15). Participants ran the tests in T-shirts and shorts after a five-minute warm-up jog and stretching session. Following the completion of all of the tests a correlation matrix was created. A significant (P<.05) moderate correlations were found between the 20m test and both the UAB test (r=0.495) and the Illinois agility test (r=0.472). These tests were also correlated with each other. The development of the 505 test was the focus of this investigation, and it correlated with the UAB test (0.506) from which it was derived, as well as the acceleration into (-0.888) and out of the turn (-0.967) at the 0 line; not with the velocity going into and out of the turn. The conclusion the authors present is that “The 505 is a valid test of agility since it is
significant correlated with acceleration in the turning phase, but not with velocity” (Draper & Lancaster, 1985, p. 18). This conclusion is limited based on the definition used for agility and its specificity of a single turn test. The impact of acceleration should not be ignored. Many definitions of agility include rapid COD which can be a quantified by calculating the rate or change in velocity or acceleration.

Gabbett et al (2008) also compared the 505 to both a modified 505 (see Figure 16), a COD „L” run (see Figure 17) and reactive agility test (RAT) (see Figure 18) using 42 rugby first and second grade players. The primary purpose of the study was to “investigate the relationship between among speed, COD speed, and reactive agility in first and second grade rugby league players” (Gabbett et al., 2008, p. 175) and to determine construct validity. Reliability was also determined via repeated measurements 48 hours following the initial test. The authors concluded that the all four tests were highly reliable with correlations of 0.90 (505 test), 0.92 (505 modified and reactive agility), and 0.95 (L run). In addition the tests had Pearson Product correlations between the test ranged between 0.46 and 0.73 (p<0.01). Most interesting to the proposed investigation was the failure of construct validity in the COD tests. This finding was determined by examining the differences in the results between the two grade levels. Neither the 505, 505 modified nor „L” run found any significant differences between the two levels of players. It would be expected that players of different playing skill levels would differ and therefore this results in a failure of construct validity. The movement time of the reactive test was, however, significantly different between the two levels. Leading the authors to conclude preplanned COD tests will not be able to discriminate between higher and lesser skilled rugby players (Gabbett et al., 2008). It was also noted
that the test was more defensive in design, which may misrepresent the capabilities of a strong offensive player.

The reactive test utilized by Gabbett et al. (2008) was first described and evaluated by Sheppard et al. (2006). The purpose of their examination was very similar to Gabbett et al. (2008). The primary purpose was to determine the reliability of a new test of agility in 38 Australian football players, but comparisons were also done between RAT and a COD speed test (CODST) (see Figure 19), and also between groups of two different performance levels. Similar to Gabbett et al. (2008) an 0.878 ICC was found for RAT. This is also one of the only studies to also determine objectivity- the interrater reliability was also high (ICC 0.904).

Many tests of agility are in use by strength and conditioning coaches. Semenick (1990) described a simple test that measures the “ability to change directions rapidly while maintaining balance without loss of speed”. The T-test (see Figure 20) was however only described in this article, no statistical analysis was run. Fry et al (1991) also used the T-test to evaluate an off-season training program for female volleyball players without the determination of reliability. Pauole et al. (2000) has since published a reliability and validity study of this test as well as the hexagon test (see Figure 21) using college men (n=153) and women (n=152). The results showed the T-test to be reliable (0.98) across three trials. Likewise so was the hexagon (0.95). The hexagon was used as the criterion measure, although this may not be appropriate, the correlation was statistically significant p<.05, r=.48 but the partial correlation was low at .31 for both males and females, indicating the T-test is not a valid measure of agility based using the hexagon as the criterion.
The hexagon had been used previously (Roetert, Garrett, Brown, & Camaione, 1992), to evaluate agility in nationally ranked junior tennis players without a determination of reliability or validity. Pauole and colleagues (2000) correctly conclude that the hexagon may not be the best choice, but also there is no “gold standard”.

Statistical analysis also found leg speed (36.6m dash), leg power (vertical jump), and agility (hexagon) only predicted between 50 and 62% of the variability in the T-test. It must then be considered that other factors such as dynamic balance or other yet to be measured characteristics (Pauole et al, 2000). The results also contain a norm table for future comparisons and classification. More recently Beekhuizen et al. (2009) attempted to establish reliability of the hexagon test 1 hour after baseline and 48 hours after baseline. Intraclass correlation coefficients “indicated excellent intrarater reliability for both same-day (0.938, 3,1) and between-day (0.924, 3,1)” (Beekhuizen et al., 2009, p. 2170). Twenty-six participants performed 3 trials with a 10 second rest interval during each of the three testing sessions. Based on the mean testing time of 13.83 seconds, this would indicate a near one to one work to rest ratio.

Most recently, Sassi et al. (2009) set out to determine the reliability of a modified version of the T-test (MAT). Eighty-six physical education students served as participants and performed a modification in which the distance total distance was decreased to 20m (see Figure 22). The participants also completed a standard T-test. With respect to the MAT, the inter-session reliability was high and similar for both men (0.95) and women (0.92). It should be noted that the test of reliability was done within a few minutes and not the minimum of 48 hours which has been the norm in previous investigations. The participants were also given a significantly greater amount of rest-
minutes between trial and 5 between tests. The work to rest ratio would be 1:18. The mean times between the standard T-test and MAT were also significantly different for the male participants, but not for the female participants.

Methodological Limitations

Psychometric Properties

Overall, the test construction process has been largely ignored when it comes to agility. As a construct, agility has not seen the same scrutiny in the literature that other dimensions of fitness and performance have seen. With no gold standard of measurement, each successive test has just been accepted as valid on its face. To complicate evaluation of the construct are the various definitions. Without a universally accepted definition, individual authors have taken the liberty of using a definition that meets the test they are using in their investigation. If we used the definition proposed by Sheppard and Young (2006), many of the tests would be invalid based on not measuring what it purports to measure. Reliability has typically not been an issue, but the results of certain studies may be questioned based the amount of rest between trials. Time between trials, in most cases, seems to be arbitrary and not prescribed by the test protocol. Additionally, objectivity has been completely ignored in the majority of the literature. Investigators are not asking outside researchers to try to replicate the results of their study to insure the directions for executing the test are followed and results are subsequently evaluated properly prior to publication or before using it to evaluate a training program. One positive trend is the use and development of tests of agility for an appropriately selected sport. However, the majority of literature and nearly all of the practitioners are still using a small group of tests of agility that at worst are invalid because they do not
require reaction to a stimulus or are at least inappropriate for either their sport and/or developmental level based on the skills being used or the distances being covered.

3-cone Test

The 3-cone test (3CR) (Figure 31) to be evaluated in this study has only been used in one published study. Hoffman et al (2007) used the 3CR to examine bilateral power differences in football players. No significant differences were seen when the players ran the test to the right (3CR) or to the left (3-cone left (3CL)) (Figure 32). The authors report a test re-test coefficient of (R>0.90). However, there is no indication of the method used to calculate the value or the protocol used to collect the data. It appears the data used were from six trials all collected on the same day. The data to be collected in this study will not only be used to confirm this finding, but also to determine the reliability of a modified version of the test 3-cone modified (3CM) (Figure 33). The creation of this new test will be guided by the test construction process.

The available research indicates that the principles of test construction are rarely followed or are not completed prior to the use of tests for evaluation of a specific training program. One of the primary foci of this study is to go through the process of test construction in order to create a new test of agility based on a previously created test. While this study will not go through all of the steps; each will be discussed and future investigations will be briefly described for those that will not be part of this study.

Test Construction

In order to develop a set of agility tests the construct first had to be clearly defined. After 75 years and the work of Chelladurai (1976; 1977) and Sheppard and Young (2006) the next step(s) is(are) much clearer. The definitions and framework they
have put forth allows for the process of constructing a new set of tests to begin. One the

tests being examined will be “new”- the 3CM. The design will be based on the

procedures described below in the following steps.

**Evaluation of the Skill/Component**

Based on the new definition of agility, the first step of test construction relies on

physical education teachers and sport coaches knowledge of the activity when identifying

the qualities of movements and patterns that might be affected by agility training.

Lessons and practices are filled with warm-up and skill specific agility movements. It is

these movements that should be used in the construction of the new set of tests. This

phase of test construction has historically not been a difficult task. Careful consideration

has led to tests that measure the quality of the movement pattern for which they were

designed. The skill(s) chosen should be representative of the complete skill set needed in

a game situation and in proportion how it is used in a game. For example, a forward on a

soccer team is most often going to be sprinting and cutting both with and without the ball,

therefore, a test that has components including back-pedaling or side shuffling would not

be appropriate because while they may need to side shuffle during a match, it is not the

primary skill required to play forward. This has become the biggest dilemma in agility

testing; most of the tests are being used universally when in fact they are limited to

specific activities, populations, sports, or developmental levels.

There is an opportunity to conduct additional research which will result in valid

and reliable tests for each position in every sport. This is an extremely large task and

therefore this study focuses specifically on a commonly used test to potentially evaluate

American football players.
Specificity of Item Selection in Relation to the Evaluation

Few sport-specific tests have been investigated for reliability and validity. To date no position specific tests have been constructed. Strength and conditioning professionals, instead, often provide coaches with results from the commonly used tests described above; however, this violates the principle of sport specificity in most cases. Specificity is a common term used in the strength and conditioning field to describe the type of activities that should be selected when designing a program. The activities chosen should match as closely as possible the actual movements athletes perform. The same test may be used for point guards in basketball and setters in volleyball. Like many early physical educators, coaches have a lack of understanding regarding test construction and statistical analysis. The result is often tests that do not hold up against scrutiny. While these attempts at sport specificity are not necessarily wrong, they do not always meet the criteria (valid, reliable, and objective) set forth for proper construction. Current tests and drills are possibilities for modification to insure that the items chosen are highly specific to the skills.

The 3CR is one of two tests used to evaluate potential football players at the National Football League Pre-Draft Combine (NFLPDC). Like many of the tests of agility it lacks specificity and is administered to all of the players when in actuality it is probably limited to only a position or two. For this reason and the fact that only a single study has reported reliability in the literature (Hoffmann et al., 2007) it was chosen instead of creating a completely new test. However, in addition to evaluating a current test, this study will create a new test that is a modification of the 3CR.

Description of administration
The administration of many tests have been described, some in great detail others only briefly. This is based both where and when the test was published in the literature. In some cases the test has just been described in a practitioner journal, while others are detailed as part of research studies. Several variables need to be considered when creating a new test according to Chelladurai's model of agility (see Figure 1) (1976) The first variable is spatial arrangement. This includes distances between cones and angles of paths between cones. The 3CR (see Figure 31) has set distances between cones of 4.6m (5 yards) in the shape of an upside down L (Hoffmann et al., 2007). The 3CL (see Figure 32) will also use the same distance but flipped to the left. The 3CM (see Figure 33) will use a T shaped orientation that combines the 3CR and 3CL. Secondly, the 3CM will have one temporal component. Following the second turn in the test the participant will be cued (see figures 34a and b) to continue the remainder of the test to either the right or left. The cue was chosen based on its simplicity for the tester to reliably execute as well as its ease of visualization for the runner. A central visual cue initiated by the tester was chosen because as Magill (1993) noted, “When all sensory systems are available to us, we tend to use and trust vision the most” (p. 103). In addition, for most motor skills, the most critical sensory receptors for providing feedback are those related to providing proprioceptive and visual information to the central nervous system.” (Magill, 1993, p. 105). A visual cue initiated from cone C at the point in which the runner is given the cue would be at 63 degrees to the right or left. A person’s field of vision covers up to 200 degrees, but this cue is much more of a challenge than placing the tester in the central vision which is 2-5 degrees (Magill, 1993). Visual search is especially important like this
where the runner has little time to make the decision following the perception of the cue (Magill, 1993).

By providing a simple cue that is in the runner’s central vision, we are attempting to minimize processing time and maximize the ability to change direction while still at near full speed. As the runner makes the return turn towards cone B the image of the tester will begin to grow larger on the runner’s retina, while it is unknown at what distance the cue will trigger the appropriate locomotor response, Lee (1987) argued that the time it takes for the image to hit the retina will determine the ability of the person to make a change in their gait. One challenge for the runners will be the fact that the tester is beyond cone B so the target for changing direction will be 2.17m (2 yards) closer than their visual focus. In the experience of the investigator this buffer is similar to the distances athletes have to make decisions regarding changing direction, and it between the distances used by Sheppard in the RAT (3 m) (see Figure 17) and the CODST (1.5m) (see Figure 18) (2006). This new component of the test will challenge the runners” ability to change the motor pattern to both brake and arrange their footwork to make a cut that is nearly 90 degrees. Lee (1984) found a similar pattern of adjustment in long jumpers as they elevated their knees to be sure to they hit the take-off board as close to then as possible. The cerebellum is responsible for the modulation of this change in direction. It receives information from the motor cortex, visual input, proprioceptors in the muscles and tendons, and vestibular feedback in an attempt to make coordinated movements. All of these external factors contribute to the choices of what cue to use and how far away the cue is being given. The final pieces of the protocol are related directly to the administration of the test.
Each of these pieces is interrelated and must be considered based on others. First the number of trials must be determined. Each test will require a minimum of 3 trials, this number is based on the mode (n=9) of the 24 studies who have similar methods to the test being evaluated in this study. An additional trial has been added for the modified test block to keep participants from anticipating a direction while getting two trials per testing session. The next step is determining the rest interval between trials. Unfortunately, this key component is often missing from the descriptions of several of the tests previously described. In those cases where it is described, the rest interval seems considerably shorter than what might be expected to allow for the anaerobic sources of energy to recover adequately. According to Hoffman et al (2007) the average time to complete the 3CR and 3CL averaged 8.0 ± 0.5 seconds. Based on this result the 3C tests, like football and weightlifting rely 100% on the muscles high energy phosphates (ATP/PCr) (Fox, Bowers, & Foss, 1989). Fox and Matthews (1974) recommend for training times less than ten seconds a 1:3 work-to-rest ratio with a maximum of 50 repetitions. As we are not training the subjects, a 1:6 work-to-rest interval was chosen to allow for a more complete recovery. Fox and Matthews also recommend having the participants walk or flex their muscles during the rest period (1974). A similar determination based on long-term physiological recovery is needed to determine the time between testing sessions or the test re-test reliability interval.

Using the same physiological rationale of full recovery in less than five minutes and only using 10 total trials; an acceptable reliability re-test interval of 24-hours is appropriate. While testing on subsequent days may be acceptable from a physiologically perspective, subject and facility availability necessitate a 48-hr separation.
Once the protocol has been completed the next step is to determine how it will be scored, and will there be any penalties for not following the prescribed protocol (i.e. false start, hitting a cone). Each trial will be timed from the participant’s first movement until he crosses the start/finish line. Any contact with a cone will result in a re-test, as well as if the participant fails to follow the correct path around the cones at any point.

The main goal of preparing the protocol is to improve the chances of having a high degree of reliability, objectivity and validity. If the test is administered even slightly differently each time a person is tested it will be very difficult to evaluate the results to determine improvement or do comparisons between individuals. Additionally, if the directions are so complicated that the performer cannot understand what is being asked of him/her this will impact validity (Johnson & Nelson, 1986).

**Test-retest Reliability**

The literature provides relatively few examples of reliability being examined in previously identified tests of agility (Kirby, 1971; Gates & Sheffield, 1940; Cozens, 1928; Sassi et al., 2009; Seils, 1951; Pauole et al., 2000). While none of these meet the current definition of the construct, ICC for these tests ranged from 0.88-0.97. In addition one study has examined reliability for 3-cone test (3CR). Hoffman et al. (2007) reported reliability coefficients R>0.90 in a study of sixty-two college football players. Based on this specific finding and the range shown in other tests of similar duration, the tests being examined during this study will be expected to have similar ICC’s.

**Calculate Objectivity**

Several agility tests purport to have objectivity scores of 0.96 or higher in a measurement text, (Johnson & Nelson, 1986), however, this claim cannot be
substantiated from the original development of the test. Many years later some the tests were evaluated for objectivity but the data were never published in any refereed journal. Often objectivity is overlooked which can lead to a drift in the tests original protocol or the rating of performance based on different scoring criteria (Johnson & Nelson, 1986). This step will be included in a future follow-up large-scale study.

*Establish Validity*

Any new test should be validated based on comparing its results with another previously validated test. However, the construct of agility has only recently been delineated so there is no gold standard for comparison, and with the degree of specificity needed for this test this method of comparison may not be appropriate. Most tests rely on face validity (the test looks to be measuring skills used game-like situations), which while logical does not allow for a statistical evaluation. Even though the 3CR is used extensively to evaluate the agility of college football players (primarily because it is also the test used during the NFLPDC), no systematic examination of the test has been completed. Based Sheppard and Young's (2006) definition of agility which requires a response to a stimulus; the 3CR and 3CL are invalid on their face because they require no reactive component. The 3CM does require a reactive response and has the potential to be validated. No test can be considered valid if it is not first found reliable, and based on this fact, the goal of this study to determine validity; but rather to instead focus on the development of a reliable test that in the future could be validated via content, criterion, or construct measures.

*Test Item Revision*
Like any evaluation process the next step after establishing both reliability and validity of an item is to make any revisions based on the results (Johnson & Nelson, 1986). Revisions are based on multiple factors including but not limited to (a) statistical analysis of reliability and validity, (b) ease of administration, and scoring, and (c) advances in technology with relationship to measurement.

**Develop Normative Data Sets**

Many of the tests developed for physical education have published norm tables. Norms have been developed for the following tests of agility: Burpee (college and high school), AHHPERD Shuttle Run (9-18yo), Quadrant Jump (college), SEMO Agility Test (college), Right-Bommerang Run (middle school and college students), LSU Agility Obstacle Course (college), and Cozens’ Dodging Run Test (college). The range of students used to develop norms for each of these tests was described as “limited” or “small group” to greater than 200. Future large-scale studies will be needed to determine agility norms using the tests examined in this study.

**Summary of the Review of Literature**

Measurement in physical education has a long and interesting history. As one of the primary responsibilities of a physical educator, sport or performance coach, assessment is routinely completed to demonstrate improvement in a variety of skills and fitness for each of their students or athletes. This has led to development of a set of standard assessments by which most are evaluated. The majority of these “standards” are used based on the fact that they have been shown to be valid and reliable for a variety of populations and developmental ages. Unfortunately agility is not a one of the components of fitness that has a gold-standard. In part this is due to a lack of consensus
in the literature as to: 1) what agility is, 2) how should it measure it be best measured, 3) what constitutes a valid and reliable test of agility, 4) what role should technology play in the measurement of agility, and 5) how much specificity is required when assessing different sports or across developmental levels.

While it is not the intention of this study to completely answer all of these questions; this study will critically evaluate one of the most commonly used tests (3CR), and two modifications (3CL and 3CM) of the test that require a reactive component. As a first step and due to the complete absence of any reliability data on the original 3CR test this study is novel, will contribute to the literature, and lead to a significant line of future research.
References: Review of Literature


Alden, F.D., Horton, M.O., & Caldwell, G.M. (1932). A motor ability test for university
women for the classification of entering students into homogenous groups.

*Research Quarterly*, 85-120.


Macmillian.


childhood motor skill proficiency predict adolescent fitness? *Medicine Science
and Sports & Exercise*, 40(12), 2137-2144.


evaluation in physical education & exercise science* (8th ed.). Boston: McGraw
Hill.


Appendix C

Test Re-Test Reliability of Three Versions of the 3-Cone Test
Additional Methods
Table 1

*Definitions of Agility Since 1945*

<table>
<thead>
<tr>
<th>Author</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Connor &amp; Cureton (1945)</td>
<td>“as the capacity for quick, efficient control of the body, is important to most sport skills, and important to the safety of the individual in emergencies where quick reaction is necessary”</td>
</tr>
<tr>
<td>Cureton (1947)</td>
<td>“fast reaction controlled movement…ability to handle the body quickly and precisely”</td>
</tr>
<tr>
<td>Clarke (1950)</td>
<td>“rapid movement of the entire body, in different directions and in response to unexpected circumstances”</td>
</tr>
<tr>
<td>McCloy (1954)</td>
<td>“ability to change direction rapidly”</td>
</tr>
<tr>
<td>Barrow (1969,1973)</td>
<td>“ability of the body or parts of the body to change directions rapidly and accurately”</td>
</tr>
<tr>
<td></td>
<td>“the ability of a performer to change direction quickly and accurately”</td>
</tr>
<tr>
<td>Kirby (1971)</td>
<td>“rapid and accurate movement of the total body in response to stimulus”</td>
</tr>
<tr>
<td>Chelladurai (1976,1977)</td>
<td>“ability to quickly change body positions and make directional changes in body movement”</td>
</tr>
<tr>
<td>Jacobson (1977)</td>
<td>“moving the body through space by an efficient combination of coordination and strength”</td>
</tr>
<tr>
<td>McClenaghan (1978)</td>
<td>“ability to change body direction and position rapidly”</td>
</tr>
</tbody>
</table>
Altug (1986) “ability to perform a series of powerful movements in opposing directions, in rapid succession”

Johnson & Nelson (1986) “rapid and precise change of body position and direction”

NSCA (1994) “ability to stop, start, and change direction of body movements of less than 10s in duration”

Hastad & Lacy (1994) “rapidly and precisely alter the position and direction of the body”

Bloomfield (1995) “incorporates elements of movement speed, as well as the ability to coordinate changes in direction and modification of the normal locomotion process”

NSCA (2000) “ability to explosively brake, change direction, and accelerate again”

Young (2001) “ability to change direction and start and stop quickly”

Craig (2004) “ability to change direction of movement quickly”

Gambetta (2004) “ability to recognize, react, accelerate, change direction, and stop quickly”

Hoffman (2006) “ability to change direction rapidly”

Sheppard (2006) “a rapid whole-body movement with change of velocity or direction in response to a stimulus”

Miller (2006) “ability to rapidly change the position and direction of the body or body parts”

Baumgartener (2007) “ability to change the direction of the body or body parts rapidly”
NSCA (2008) “the skills and abilities needed to explosively change movement velocities or modes”
Table 2

*Terms Used in the Definitions of Agility Since 1947*

<table>
<thead>
<tr>
<th>Author</th>
<th>Precision/Accuracy</th>
<th>Change of direction</th>
<th>Body (parts)</th>
<th>Quickness/Rapidity</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Conner &amp; Cureton (1945)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cureton (1947)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Clarke (1950)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>McCloy (1954)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barrow (1969,1973)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Kirby (1971)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chelladurai (1976,1977)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jacobson (1977)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McClenaghan (1978)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draper (1985)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altug (1986)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson &amp; Nelson (1986)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>NSCA (1994)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hastad &amp; Lacy (1994)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Bloomfield (1995)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>NSCA (2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young (2001)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craig (2004)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambetta (2004)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoffman (2006)</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Sheppard (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Miller (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baumgartener (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSCA (2008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3

*Classifications of Agility* (Chelladurai, 1976)

<table>
<thead>
<tr>
<th>Type</th>
<th>Presence of temporal uncertainty (when?)</th>
<th>Presence of spatial variation (where?)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>No</td>
<td>No</td>
<td>Field events</td>
</tr>
<tr>
<td>Temporal</td>
<td>Yes</td>
<td>No</td>
<td>Sprint events</td>
</tr>
<tr>
<td>Spatial</td>
<td>No</td>
<td>Yes</td>
<td>Receiving a serve</td>
</tr>
<tr>
<td>Universal</td>
<td>Yes</td>
<td>Yes</td>
<td>Goal keeping</td>
</tr>
</tbody>
</table>
**Table 4**

*Balance Latin Square for Determining Test Block Order*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 6

*Balance Latin Square for Determining Trials of the 3-cone Modified*

<table>
<thead>
<tr>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>L</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>R</td>
<td>R</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>R</td>
<td>R</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>
Figure 1. Model of Agility (Chelladurai, 1976)
Figure 2. Conceptual Framework of Agility (Sheppard and Young, 2006)
Figure 3. Cozens’ Dodging Run Test (Cozens, 1929)

Procedures: “The performer starts from behind the line to the right of the first hurdle (or cones), which is the starting line. On the signal “Go,” the performer runs to the left of the second hurdle and follows the course as shown in the figure. Two complete round trips constitute a run, and two runs are made.”

Scoring: “The better of the two runs recorded to the nearest tenth of a second is the score.”
**Procedures:** “Lower the body to a squat-rest position, leaning forward, and placing hands on the floor in front of the feet. Thrust the legs backward to the front leaning-position. Return almost to the squat-rest position, and then to the standing position. Repeat the movements as rapidly as possible until the command to stop is given.”

**Scoring:** “The test is scored in terms of the number of performances completed in ten seconds. A complete performance is scored as 1. Scores for a partial performance are as follows: ¼ for first touching the hands to the floor, ½ for thrusting the legs backward, and ¾ for returning to a squat-rest position with the hands still on the floor.”
Figure 5. Zig-Zag Run (Young K. E., 1937)

**Procedures:** The player started at one end of the starting line (labeled 10), running around through the hurdles and back to the other end of the line.

**Scoring:** Time to run two laps around the course.
**Figure 6.** Potato Race, 10 yards (Young K. E., 1937)

**Procedures:** “The runner starts from the first line, runs and picks up one block, brings it back to the first line and leaves it there, then she gets the second block and touches the first line with it and returns it to its original place. She then comes back to the first line and picks up the first block and returns it to the second line and then run back across the starting line.”

**Scoring:** “The best score is the smallest and the time is taken from the word “go” until the runner crosses the starting line at the finish of the run.”
Figure 7. Howe Exercise (Young K. E., 1937)

Procedures: “The participant lies on the floor on her back with her feet behind the starting line, hands over her head on the floor. On the command “go” she springs to her feet, crosses the two lines and assumes the same position on the other side of the second line. She then taps her hands on the floor twice above her head, springs up and returns to the starting position. The participant does this back and forth until three round trips have been made.”

Scoring: “The time is taken from the starting command until the girl has completed the second tap at the end of the third round.”
Figure 8. AAHPERD Shuttle Run (AAHPERD, 1976)

**Procedures:** The student stands at one of the lines with the two blocks at the other line. On the signal to start, the student runs to the blocks, takes one and returns to the starting line and places the block behind that line. He then returns to the second block which he carries across the starting line on his way back. Two students could run at the same time if two timers are available, or if one test administrator has a split-second timer, and of course, if there are two sets of blocks. Two trials are permitted. If the students start first at one line and then at the other, it will not be necessary to return the blocks after each race. Sneakers should be worn or the student may run barefooted.

**Instructions:** On the signal to “Go” you must run as fast as you can to the next line and pick up a block. You should return the block over the second line where you place it on the floor. Do not throw it. You return for the second block and this time, you may run across the starting line as fast as you can without placing the block on the floor.

**Scoring:** The score is the elapsed time recorded is seconds and tenths of seconds for the best of two trials
Figure 9. Modified ZigZag-run (Johnson R. D., 1960)

Procedures: “The subject was instructed to stand behind the middle of the starting line, and on the command Go, to run either to the right of left of the first chair, to zigzag around the three remaining chairs, to touch the X, and to return in the same manner and to touch the starting line with his foot.”

Scoring: “The score is the time to the nearest tenth of a second that was required for running the course.”
Figure 10. Zigzag-run Dribbling a Basketball (Johnson L. W., 1934)

Procedures: “The player started at one end of the starting line (labeled 10) and dribbled around through the hurdles and back to the other end of the line.”

Scoring: “The player’s score was the number of zones he passed in 30-seconds.” Zone 1 was the left side of the first hurdle, zone 2 was the right side of the second hurdle…and zone 10 was the starting point.”
Figure 11. Edgren Side step (NSCA)

**Procedures:** “The athlete stands astride the centerline. On the “Go” command, the athlete side steps to the right until the right foot has touched or crossed the outside line. The athlete then side steps to the left until the left foot has touched or crossed the left outside line. The athlete side steps back and forth to the outside lines as rapidly as possible for 10 seconds.”

**Scoring:** “Each traversal of a 3-ft subsection (from centerline to first increment, from first increment to outside line, from outside line back to first increment, etc.) counts as 1 point.”

**Penalty:** 1-point penalty for each time the feet cross or if the outside foot does not cross the line.
**Figure 12.** Edgren Coordination test (1932)

**Procedures:** “This test was devised as a means of judging the ability of an individual to shift his body from left to right similarly to the way a basketball player is forced to do when guarding an opponent. The participant must work with feet spread in a good base and must then shift the body across an eight-foot lane. The participant worked inside the lane and needed only to touch the line with the outside foot on each shift.”

**Scoring:** “He was scored on the time it took to make ten shifts from left to right and from right to left. One shift across counted as one time.”
**Figure 13.** Up and Back Agility Test (Draper 1985)

**Procedures:** “Involves running a total of 30m, 15m there and 15m back. The participants were required to stand on the 15m mark and on the command „Go” they would run to the zero line, then run back and past the 15m mark.” “One foot is required to touch the zero line during the change of direction.”

**Scoring:** “Time taken to cover the 30m was recorded.” Times were recording using a video timer.
Figure 14. 505 Test (Draper 1985)

Procedures: Completed during the turning phase of the UAB. (solid line)

Scoring: Time was recorded using the video timer “from when the participant originally passed the 5m mark, until the returned to it.”
Figure 15. Illinois Agility Test (Getchel, 1979)
Procedures: “Unlike the traditional 505 test (where players assume a starting position 10m from the timing gates and therefore 15 m from the turning point), players started 5 m from the timing gates.”

Scoring: Time was recorded using the gate to the nearest 0.01s, with the fastest value used for the score.
**Figure 17.** „L” run (Gabbett et al., 2008)

**Procedures:** “3 cones were placed 5m apart in the shape of an „L“. Players were instructed to run as quickly as possible along the „L” run.”

**Scoring:** Times were measured to the nearest 0.01s with the fastest value used for the score.
Figure 17. Reactive agility test (Sheppard and Young, 2006)

Procedures: “The tester (investigator) stood opposite and facing the participants. The tester stood behind a set of timing lights. Each test trial involved the tester initiating movement, and thereby beginning the timing. The athlete reacted to the movements of the tester by moving forward, to the left or right in response to, and in the same direction as, the left or right movement of the tester. The tester displayed one of four possible scenarios for the athlete to react to: 1) step forward with right foot and change of direction to the left, 2) step forward with the left foot and change of direction to the right, 3) step forward with the right foot, then left, and change of direction to the right, and 4) step forward with the left foot, then right, and change of direction to the left.”
**Scoring**: Time starts as the tester steps off a pressure sensitive mat, and end when the athlete passes through the gate. The recorded score was the mean of all trials (8), which was the average of all trials to the left (4) and right (4).
Figure 18. Change of direction speed test (CODST)

Procedures:

Scoring: Timing was begun as the athlete broke through the first gate, and stopped as athlete broke the beam of the lateral gate to either the right or left.
**Figure 19.** T-Test Semenick (1990)

**Procedures:** The runner starts at A in a two-point stance, sprints to B touches the cone, side shuffle to C touches the cone, side shuffles to D touches cone, side shuffles back to B touches cone, and back pedals finishing past A.

**Scoring:** Two trials and record the fastest time.
Figure 20. Hexagon Test (Roetert, 1992)

Procedures: Participants faced forward and began in the hexagon and began on the command “Ready -Go” by jumping feet together over one side of the hexagon and immediately back into the hexagon. This pattern was followed for three revolutions.

Scoring: Time in seconds was recorded. No details on start/stop.
Figure 21. Modified T-Test Sassi et al. (2009)

Procedures: The runner starts at A in a two-point stance, sprints to B touches the cone, side shuffle to C touches the cone, side shuffles to D touches cone, side shuffles back to B touches cone, and back pedals finishing past A.

Scoring: Two trials and record the fastest time.
Figure 31. 3-cone right (3CR) (Hoffman, 2007)

Procedures: Players start in a three-point stance in front of three cones that are set up in a triangle or L shape, with each cone five yards apart. They then sprint five yards to one cone (B), sprint back to the starting cone (A), and head back to the second cone (B) where they run around it and cut right to the third cone (C). The players then run a circle around the third cone (C) from the inside to the outside and run around the second cone (B) before returning to the first cone (A).
**Figure 32. 3-cone left (3CL) (Hoffman, 2007)**

**Procedures:** Players start in a three-point stance in front of three cones that are set up in a triangle or L shape, with each cone five yards apart. They then sprint five yards to one cone (B), sprint back to the starting cone (A), and head back to the second cone (B) where they run around it and cut left to the third cone (C). The players then run a circle around the third cone (C) from the inside to the outside and run around the second cone (B) before returning to the first cone (A).
Figure 33. 3-cone modified (3CM)

Procedures: Players start in a two-point stance in front of three cones that are set up in a triangle or L shape, with each cone five yards apart. They then sprint five yards to one cone (B), sprint back to the starting cone (A), and head back to the second cone (B) at the 2.5–yard line the tester will point either right or left requiring the player to react and take a path toward C or D. The players then run a circle around the third cone (C or D) from inside to the outside and run around the second cone (B) before returning to the first cone (A).
In step three, as the runner is returning toward cone B for the second time the tester will start with his hands in directly in front of his chest as shown in the figure. As the runner passes the midpoint (equivalent to 2.3m) of this portion of the test the tester will cue to either the right (as depicted in the figure) or left.

Figure 34a Starting position of the cue for 3CM.

Figure 34b Position of cue for 3CM when directing the runner to the tester’s right.
Appendix D

Test Re-Test Reliability of Three Versions of the 3-Cone Test
Informed Consent
CONSENT AND INFORMATION FORM

Principal Investigator: Sean M. Bulger
Department: West Virginia University (WVU) Coaching and Teaching Studies in the College of Physical Activity and Sport Sciences

Tracking Number:

Study Title: Test Re-Test Reliability of Three Versions of the 3-Cone Test

May 10, 2010

Co-Investigator(s): Jason G. Langley, Kristen Dieffenbach, Guy Hornsby, Robert Wiegand

Contact Persons
In the event you experience any side effects or injury related to this research, you should contact Mr. Jason G. Langley at (304)293-0852. (After hours contact: Mr. Jason G. Langley at (262) 909-6661). If you have any questions, concerns, or complaints about this research, you can contact Mr. Jason G. Langley or Dr. Sean M. Bulger at (304) 293-0845.

For information regarding your rights as a research subject, to discuss problems, concerns, or suggestions related to the research, to obtain information or offer input about the research, contact the Office of Research Compliance at (304) 293-7073.

Introduction
You, ______________________, have been asked to participate in this research study, which has been explained to you by Jason G. Langley ABD. This research is being conducted to fulfill the requirements for a doctoral dissertation in Kinesiology in the Department of Coaching and Teaching Studies at West Virginia University, under the supervision of Sean M. Bulger Ed.D.

Purposes of the Study
The purpose of this study is to learn more about the reliability of a common measurement of agility and two alternative measurements. Each subject will perform multiple trials of the 3 cone drill commonly used to evaluate agility at the NFL Pre-Draft Combine, as well as two modifications of the same test. You will be expected to give maximal effort on each of the trials during the test re-test period. WVU expects to enroll approximately 50 subjects the general study body. The study will be conducted at Stansbury Hall 83 Beechurst Ave, Morgantown, WV 26505.

Description of Procedures
This study involves a 10-15 minute warm-up targeted at the lower body, a set of sub-maximal agility trials with adequate rest between, followed by a total of 10 maximal agility trials for the three tests, and completing with a cool-down. It will take approximately 45 minutes for you to complete. Prior to any activity you will be asked to fill out a questionnaire regarding your athletic and injury history. This will take
approximately five minutes. You do not have to answer all the questions. You will have the opportunity to see the questionnaire before signing this consent form.

**Risks and Discomforts**
There are no known or expected risks from participating in this study, except for the mild frustration associated with answering the questions.

**Alternatives**
You do not have to participate in this study.

**Benefits**
You may not receive any direct benefit from this study. The knowledge gained from this study may eventually benefit others.

**Financial Considerations**
There are no special fees for participating in this study, but any expense associated with current therapy or treatment of side effects will be billed to you or to your insurance company.

You will earn extra credit for participating in this study. Other options are available for earning the same extra credit.

**Confidentiality**
Any information about you that is obtained as a result of your participation in this research will be kept as confidential as legally possible. Your research records and test results, just like hospital records, may be subpoenaed by court order or may be inspected by the study sponsor or federal regulatory authorities (including the FDA if applicable) without your additional consent.

Audiotapes or videotapes will be kept locked up and will be destroyed as soon as possible after the research is finished.

In any publications that result from this research, neither your name nor any information from which you might be identified will be published without your consent.

We know that information about you and your health is private. We are dedicated to protecting the privacy of that information. Because of this promise, we must get your written authorization (permission) before we may use or disclose your protected health information or share it with others for research purposes.

You can decide to sign or not to sign this authorization section. However, if you choose not to sign this authorization, you will not be able to take part in the research study. Whatever choice you make about this research study will not have an effect on your access to medical care.

**Persons/Organizations providing the information:**
WVU students

Persons/Organizations receiving the information:
- The members and staff of any Institutional Review Board (IRB) that oversees this research study.
- West Virginia University Office of Research Compliance and Office of Sponsored Programs.
- The members of the research group.

The following information will be used: New information about you that is created or collected during the study such as: athletic and injury history, demographic data, and study forms.

The information is being disclosed for the following reasons: Review of your data for quality assurance purposes
- Publication of study results (without identifying you)
- Other research purposes such as improving the design of future agility testing.

You may cancel this authorization at any time by writing to the Principal Investigator:
Sean M. Bulger EdD
Coliseum 283
PO BOX 6116
Morgantown, WV 26506-6116

If you cancel this authorization, any information that was collected already for this study cannot be withdrawn. Once information is disclosed, according to this authorization, the recipient may redisclose it and then the information may no longer be protected by federal regulations.

You have a right to see and make copies of your records. You will not be able to see or copy your records related to the study until the sponsor has completed all work related to the study. At that time you may ask to see the study files related to your participation in the study and have the study coordinator correct any information about you that is wrong.

This authorization will not expire unless you cancel it.

SIGNATURE
I have read this section and all of my questions have been answered. By signing below, I acknowledge that I have read and accept all of the above.

____________________________________________ __________________
Signature of Subject or Authorized Representative  Date

____________________________________________
Print Name of Subject or Authorized Representative

Voluntary Participation
Participation in this study is voluntary. You are free to withdraw your consent to participate in this study at any time. Refusal to participate or withdrawal will not affect your future care, [or your employee status at West Virginia University or your class standing or grades, as appropriate] and will involve no penalty to you.

In the event new information becomes available that may affect your willingness to participate in this study, this information will be given to you so that you can make an informed decision about whether or not to continue your participation.

You have been given the opportunity to ask questions about the research, and you have received answers concerning areas you did not understand.

Upon signing this form, you will receive a copy.

I willingly consent to participate in this research.

__________________________________________  ____________________________________________  ______
Signature of Subject or Subjects Legal Representative Printed Name Date Time

The participant has had the opportunity to have questions addressed. The participant willingly agrees to be in the study.

__________________________________________  ____________________________________________  ______
Signature of Investigator or Co-Investigator Printed Name Date Time
Appendix E

Test Re-Test Reliability of Three Versions of the 3-Cone Test
Subject Demographic Questionnaire
Subject Demographic Questionnaire

ID

Date of Birth

Year at WVU Tech

Height ______ Weight ______ Which foot do you kick best with? L______ R____

Injury History

1. Have you ever had a lower extremity injury that will affect your participation in this study? Circle Yes or No. If yes, please explain: __________________________
   __________________________
   __________________________
   __________________________

2. Have you had surgery to the lower extremity that will affect your participation in this study? Circle Yes or No. If yes, please explain: __________________________
   __________________________
   __________________________
   __________________________

3. Have had any neurological disorders that will affect your participation in this study? Circle Yes or No. If yes, please explain: __________________________
   __________________________
   __________________________
   __________________________
4. Have you had or do you currently have any vestibular (balance) disorders that will affect your participation in this study? Circle Yes or No  If yes, please explain: ______________

______________________________________________________________

______________________________________________________________

______________________________________________________________

5. Have you had or do you currently have any visual disorders that will affect your participation in this study? Circle Yes or No  If yes, please explain: ______________

______________________________________________________________

______________________________________________________________

______________________________________________________________

Athletic/Physical activity History

1. Did you participate in competitive sport during high school? List which sport(s) and the highest level you played . ________________________________

______________________________________________________________

______________________________________________________________

2. Do you currently or have you ever participated in competitive sport at the collegiate level? List which sport(s) and at what level you played.____________________________

______________________________________________________________

______________________________________________________________

3. Do you currently participate in any of the following activities? Please circle all that apply.

Weight Training/ Cardiovascular Training/ Other

Please explain the frequency (days per week) of your participation. __________