Reliability and Validity of a New Test of Agility and Skill for Female Amateur Soccer Players

by

Mehmet Kutlu¹, Hakan Yapıcı², Abdullah Yılmaz³

The aim of this study was to evaluate the Agility and Skill Test, which had been recently developed to assess agility and skill in female athletes. Following a 10 min warm-up, two trials to test the reliability and validity of the test were conducted one week apart. Measurements were collected to compare soccer players’ physical performance in a 20 m sprint, a T-Drill test, the Illinois Agility Run Test, change-of-direction and acceleration, as well as agility and skill. All tests were completed following the same order. Thirty-four amateur female soccer players were recruited (age = 20.8 ± 1.9 years; body height = 166 ± 6.9 cm; body mass = 55.5 ± 5.8 kg). To determine the reliability and usefulness of these tests, paired sample t-tests, intra-class correlation coefficients, typical error, coefficient of variation, and differences between the typical error and smallest worthwhile change statistics were computed. Test results showed no significant differences between the two sessions (p > 0.01). There were higher intra-class correlations between the test and retest values (r = 0.94–0.99) for all tests. Typical error values were below the smallest worthwhile change, indicating ‘good’ usefulness for these tests. A near perfect Pearson correlation between the Agility and Skill Test (r = 0.98) was found, and there were moderate-to-large levels of correlation between the Agility and Skill Test and other measures (r = 0.37 to r = 0.56). The results of this study suggest that the Agility and Skill Test is a reliable and valid test for female soccer players and has significant value for assessing the integrative agility and skill capability of soccer players.

Key words: agility and skill test, ability, change direction, female soccer fitness test.

Introduction

Sports performance and success in the complex and hybrid sport of soccer depends on multiple factors, including technical skills, tactics, mental aspects, and physical capacities (i.e., strength, power, sprinting ability, agility) (Lockie et al., 2013; Stolen et al., 2005). Performance assessment and the ability to execute skilled movement patterns efficiently and effectively are the most important aspects for talent identification in soccer (Ali, 2011). In team sports such as soccer, agility is an important quality for evading opponents when attacking or for placing pressure on opponents when defending (Young and Willey, 2010). Soccer also involves perceptual-motor skills that operate simultaneously in a rapidly changing environment (Bullock et al., 2012; Sheppard and Young, 2006).

The physical capacities of soccer players may be tested by various physical fitness tests. Fitness tests for soccer can be divided into laboratory tests and soccer-specific field tests. Even though laboratory tests are considered reliable and useful in measuring general fitness levels, their validity at measuring soccer-specific capacities has been questioned (Aandstad and Simon, 2013). Soccer fitness tests traditionally consist of submaximal running, straight sprints, and agility sprints (Aandstad and Simon, 2013).

1 - School of Physical Education and Sport, University of Istanbul Gelisim ISTANBUL, Turkey.
2 - School of Physical Education and Sport, Kirikkale University, Turkey.
3 - Department of Statistics, Art & Science Faculty, the University of Kirikkale, Kirikkale, Turkey.
However, some soccer-specific tests also include movement or actions with the ball (Stolen et al., 2005), which increases the validity of such a test. Yet such tests might still not be ideal for measuring physical fitness, since ball handling skills likely influence the result (Da Silva et al., 2011; Gabbett et al., 2010; Psotta et al., 2005; Sporis et al., 2010). Reilly and Holmes (1983) obtained preliminary data on the validity and reliability of a battery of tests involving soccer players that used simple protocols that included a shooting test. Yet, in this type of assessment, shots at the goal were taken from a static position, and it can therefore be argued that the test was actually assessing ‘technique’ rather than ‘skill’ (Ali, 2011).

Recognizing the limitation of available skill and technical tests, Ali et al. (2007) developed the Loughborough Soccer Passing Test (LSPT) and Loughborough Soccer Shooting Test (LSST). Nonetheless, these skill-based tests did not specifically aim to measure agility. Lockie et al. (2011) recently developed a new test for field-based sports - the Change-of-Direction and Acceleration Test (CODAT), which monitors player’s performance. However, this test did not include an assessment of ball-shooting skills. Taking these test limitations into consideration, Kutlu et al. (2012) developed a combined soccer-specific agility and skill test, including sprinting, change-of-direction, acceleration, deceleration, forward and backward sprinting, and shooting at the goal. This agility and ball-shooting test was deemed valid and reliable for male soccer players; however, it remained uncertain whether it was reliable and valuable for female amateur soccer players. Most reports examining fitness performance and tests have been based on male players, with only a few involving female counterparts (Ali et al., 2008; Jason and Terence, 2014; Kaminski et al., 2008). Therefore, the purpose of this study was to evaluate the content validity and test/retest reliability of agility and skill-specific field tests for female amateur soccer players that included both measure to change direction and ball shooting skills to a target with decision making.

Material and Methods

Participants

Thirty-four amateur female soccer players (age 20.8 ± 1.9 years; body height 166 ± 6.9 cm; body mass 55.5 ± 5.8 kg) that played in the university soccer team (training experience 2.7 ± 1.1 years) volunteered to participate in this study. The study was approved by the University’s Ethics Committee, and conformed to the policy statement with respect to the Declaration of Helsinki and ethical guidelines. None of the athletes had existing medical conditions that would have prevented participation in the study. Before the start of the study, each participant received a clear explanation of the study, including the risks and benefits of participation, and written informed consent was obtained.

Preliminary Procedures

Before the initial testing session, each subject’s age, body height and body mass were recorded. Prior to testing, two one-hour familiarization sessions were conducted, one day apart. Athletes completed two agility and skill-specific field testing sessions with trials spaced one week between each session (i.e., test/retest) (Lockie et al., 2013; Sheppard et al., 2006; Young and Willay, 2010). A 10 min standardized warm-up consisting of jogging, change of direction, sprinting and passing exercises was performed prior to preliminary and experimental test procedures. During the familiarization session, participants were provided with a brief explanation of how to perform the tests and then allowed four attempts to ensure full familiarization with the procedures.

Testing Procedures (Performance Assessment)

Three successful trials were completed for each test and the mean scores were computed for each session. No evaluation was provided to athletes as to the most effective movement technique and they were instructed to complete the tests as quickly as possible. Athletes were also instructed that they must cut around markers and not skip over them. Trials were stopped and re-attempted after a brief rest period if any subject cut over the top of a marker. A minimum of 3 min rest was provided between test and retests trials, and approximately from 5 to 6 min of rest were allowed between different tests to reduce the likelihood of fatigue. Participants wore sport shorts, t-shirts, and soccer shoes during tests. If the athletes made mistakes or hesitated during the tests, the trial was disregarded and another attempt was allowed after a recovery period. Time for each distance was recorded to the
nearest 0.01 s for a 20 m sprint test (Jarvis et al., 2009). The T-drill classic test and the Illinois agility run (IAR) test were conducted in accordance with established methods (Wilkinson et al., 2009). The Change-of-Direction and Acceleration Test (CODAT) was conducted following the previously described procedure (Lockie et al., 2013). Then, the Agility and Skill (AS) test (measuring both ball skills and goal success) was performed in accordance with the methodology described by Kutlu et al. (2012). All tests were conducted in the same order as mentioned above. A diagram of the newly developed Agility and Skill Test (AS) for soccer is shown in Figure 1.

In the gymnasium, a carpet fragment was placed under the four balls to prevent them from sliding and to ensure an easy setting. At the end of the AS test, goal success scores were determined based on the recorded time, as follows: if the subject could score four goals, 1 s was subtracted from the recorded completion time; 0.75 s were subtracted for three goals, 0.50 s were subtracted for two goals, and 0.25 s were subtracted for one goal. If the subject did not manage to strike the goal on any of the four trials, the raw complete time score for the subject was recorded. Without goal success, AS test results were accepted only as recorded time (as a second AS test), ignoring whether the shots resulted in a goal or not (see Kutlu et al., 2012, for further scoring parameters).

During test and retest sessions, ambient temperature and relative humidity were recorded (test = 21.2 ± 0.7°C, 35 ± 2%; re-test = 22.0 ± 0.9°C, 33 ± 1.8%; p > 0.01). Prior to testing, all players warmed up for approximately 10 min with slow-to-moderate running interspersed with short sprints of progressive intensity. CODAT, Illinois Agility Run, and T-Drill classic tests were used to validate the new AS test as they evaluated acceleration, deceleration and balance control aspects of agility. Additionally, these tests are comparatively easy to learn and can be used to evaluate athletes in field sports such as soccer. Reported validity and reproducibility of these tests suggest their suitability as comparison instruments for this study (Lockie et al., 2013).

Sprint time was measured with the use of timing gates (Tumer Electronic, Timing System; Ankara, Turkey). The players used a standing start 30 cm behind the start line in order to trigger the first gate (Sporis et al., 2009), placing their preferred foot in the forward position. If the athlete made a mistake or hesitated during the tests, the trial was disregarded and another attempt was allowed. Time for each distance was recorded to the nearest 0.01 s (Gabbett, 2010; Sporis et al., 2009).

Statistical Analysis

Data analyses were performed using the Statistics Package for Social Sciences (SPSS, V20.0; IBM Corporation, NY, USA). The normality of data was confirmed using the Shapiro–Wilk procedure. Intra-class correlation coefficients (ICC) were calculated for test/retest results. An ICC equal to or above 0.70 was considered acceptable (Hopkins, 2004). Test, means, standard deviations, and coefficients of variance (CV) were calculated. For validity, Pearson correlation coefficients (r) were used. Absolute reliability was also assessed by paired-sample t-tests, to determine the differences between the test/retest sessions (p > 0.01). A coefficient of variance (CV) of less than 5% was also set as the reliability criterion (Lockie et al., 2011). The usefulness of the tests was determined by comparing typical error (TE) to the smallest worthwhile change (SWC) in the timing for each test (Hopkins, 2004). If the TE was below the SWC, the test was rated “good”; if the TE was similar to SWC, the test was rated “acceptable”; and if the TE was higher than the SWC, the test was rated “marginal”. To determine the validity of AS tests, Pearson correlations (p ≤ 0.05) were computed among the tests to determine the interrelationships (Hopkins, 2004). If a correlation coefficient (r) was less than 0.30, the correlation was considered small; 0.31 to 0.49 was considered moderate; 0.5 to 0.69 was considered large; 0.7 to 0.89 was considered very large; and 0.9 and higher was considered nearly perfect for predicting relationships (Hopkins, 2004). The results are presented as means ± SD. The level of significance was accepted at p < 0.05.

Results

Reliability and Usefulness

Table 1 shows the mean times ± SD, 99 percent confidence intervals (99% CI), p values for paired t-tests, intra-class correlations (ICCs), typical error (TE), coefficients of variations (CV), and smallest worthwhile changes (SWC) for
rating the usefulness \((0.5 \times SD = SWC_{0.5})\) of the six tests over the two testing sessions. Paired sample t-tests revealed no significant \((p > 0.01)\) differences between the testing sessions for all speed, agility and AS tests regarding their relative reliability. The ICCs for each of the performance tests were more than 0.70, with a range of 0.94 to 0.98. The CV was equal to 5% for the AS goal success test, and less than five percent for all other tests, a good acceptable indicator of reliability (Lockie et al., 2013; Shan et al., 2012). The Typical Error was either similar to or below the \(SWC_{0.5}\) for each performance test. The Typical Error (TE) was below the \(SWC_{0.5}\) for all tests.

**Validity**

To understand the validity quotient of the new agility test, Pearson correlations between agility tests were computed. If correlations exist, it is presumed that there is a relative similarity between the agility tests. Pearson correlations and \(p\)-values for each of the speed, agility, and AS tests are displayed in Table 2. The strongest significant positive relationships were found between the two AS tests (with ball and goal success; \(r = 0.98; p < 0.01\)), and moderate-to-large level significant correlations were computed for each of the other performance tests (0.39 to 0.66; \(p < 0.01\) to \(p < 0.05\)).

---

**Figure 1**

*A diagram and schematic representation of the newly developed agility and skill (AS) test for soccer.*
Table 1
Test/retest results for all performance tests: descriptive data, p values for paired t-tests between sessions, ICC between sessions, CV, mean and confidence intervals (99% CI), and other reliability statistics (TE and SWC rating) for amateur female soccer players (n = 34).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Test</th>
<th>Re-test</th>
<th>p</th>
<th>ICC</th>
<th>Mean &amp; 99% CI</th>
<th>TE</th>
<th>CV</th>
<th>Rating</th>
<th>SWC 0.5 Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-Drill</td>
<td>12.17 ± 0.56</td>
<td>12.21 ± 0.53</td>
<td>0.30</td>
<td>0.95</td>
<td>12.19 (11.9–12.5)</td>
<td>0.09</td>
<td>0.04</td>
<td>Good</td>
<td>0.27</td>
</tr>
<tr>
<td>AS</td>
<td>12.65 ± 0.57</td>
<td>12.72 ± 0.57</td>
<td>0.03</td>
<td>0.95</td>
<td>12.68 (12.4–13.0)</td>
<td>0.10</td>
<td>0.04</td>
<td>Good</td>
<td>0.28</td>
</tr>
<tr>
<td>AS (goal)</td>
<td>12.02 ± 0.62</td>
<td>12.04 ± 0.62</td>
<td>0.34</td>
<td>0.98</td>
<td>12.03 (11.7–12.3)</td>
<td>0.11</td>
<td>0.05</td>
<td>Good</td>
<td>0.31</td>
</tr>
<tr>
<td>20 m Sprint</td>
<td>3.36 ± 0.15</td>
<td>3.37 ± 0.14</td>
<td>0.56</td>
<td>0.94</td>
<td>3.36 (3.3–3.4)</td>
<td>0.02</td>
<td>0.04</td>
<td>Good</td>
<td>0.07</td>
</tr>
<tr>
<td>CODAT</td>
<td>6.36 ± 0.28</td>
<td>6.39 ± 0.27</td>
<td>0.07</td>
<td>0.98</td>
<td>6.38 (6.3–6.5)</td>
<td>0.05</td>
<td>0.04</td>
<td>OK</td>
<td>0.13</td>
</tr>
<tr>
<td>IAR</td>
<td>19.07 ± 0.70</td>
<td>19.12 ± 0.74</td>
<td>0.04</td>
<td>0.98</td>
<td>19.19 (18.8–19.4)</td>
<td>0.12</td>
<td>0.04</td>
<td>Good</td>
<td>0.36</td>
</tr>
</tbody>
</table>

*SWC: Smallest Worthwhile Changes for Rating the Usefulness of the Measured Tests (Hopkins, 2004)

Table 2
Pearson correlations and p values for the 20 m sprint, AS, CODAT, IAR and T-Drill agility tests or amateur female soccer players (n = 34).

<table>
<thead>
<tr>
<th>Tests</th>
<th>AS (with a ball)</th>
<th>T-Drill (classic)</th>
<th>AS (goal success)</th>
<th>20 m Sprint</th>
<th>CODAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS (with a ball)</td>
<td>r 0.56**</td>
<td>p 0.00</td>
<td>r 0.98**</td>
<td>p 0.00</td>
<td>r 0.49**</td>
</tr>
<tr>
<td>AS (goal success)</td>
<td>r 0.56**</td>
<td>p 0.00</td>
<td>r 0.52**</td>
<td>p 0.01</td>
<td>r 0.37*</td>
</tr>
<tr>
<td>20 m Sprint</td>
<td>r 0.46**</td>
<td>p 0.00</td>
<td>r 0.52**</td>
<td>p 0.02</td>
<td>r 0.43*</td>
</tr>
<tr>
<td>CODAT</td>
<td>r 0.52**</td>
<td>p 0.00</td>
<td>r 0.66**</td>
<td>p 0.03</td>
<td>r 0.54**</td>
</tr>
<tr>
<td>IAR</td>
<td>r 0.46**</td>
<td>p 0.00</td>
<td>r 0.52**</td>
<td>p 0.00</td>
<td>r 0.43*</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).
AS Test: Agility and Skill Test (Newly Developed)
CODAT: Change-of-Direction and Acceleration Test
IAR: Illinois Agility Run Test
Discussions

Few studies have examined the physical characteristics and skills of female soccer players, even for such basic skills as kicking; however, no study has developed a combined agility and skill test for this group (Kutlu et al., 2012; Little and Williams, 2005). Sports-specific field tests not only reflect real game distances covered, but also the number of different movement categories. Agility sprints and straight sprints are among the major types of these tests (Bullock et al., 2012). Yet, such tests might not be ideal for measuring physical fitness of soccer players since these measures do not evaluate skills requiring decision-making processes, such as ball-shooting to a target. In actual soccer match situations, players must both perform with agility and have ball skills in order to be successful. Kutlu et al. (2012) previously developed an Agility and Skill Test (AS) for this purpose for male athletes. The present study aimed to analyze the utility of the AS test for female amateur soccer players. With this new test, directional-change agility can be measured together with the number of successful shots on the goal as a dimension of soccer skill.

The new AS tests were compared with CODAT, IAR and T-Drill classic agility tests to ensure validity. The 20 m linear sprint test was also used to determine the validity of the test, as previous research had shown acceptable correlation when comparing it to change-of-direction speed (Buchheit et al., 2011; Jarvis et al., 2009). Results of the present study showed that the AS test had a high level of reliability between the test and retest (AS with the ball, r = 0.95; AS with success, r = 0.98) for female soccer players, which was similar to other agility tests measured in this study and to that of adult male soccer players (Kutlu et al., 2012). Coefficients of variations (≤5%) we calculated for all tests, were also consistent with previous research (Green et al., 2010; Lockie et al., 2013; Sporis et al., 2011). In combination, these results imply substantial evidence for reliability of the new agility and skill test. However, it must be noted that there was no significant difference between the mean AS times for test-to-retest and with other sprint and agility tests (paired t-test p > 0.01) (Table 1). In line with our findings, some previous research has found good reliability for measures of linear and change-of-direction agility performance in other tests, including IAR, CODAT and T-Drill tests (Hachana et al., 2013; Hopkins, 2004; Kowacks et al., 1999; Lockie et al., 2013; Oliver and Meyers, 2009). Indeed, the acceptable level of relative reliability, and thus the degree to which the test is useful, is also in the ‘acceptable-to-good’ rating for the AS test and other measured tests for female soccer players. This was confirmed by the degree of difference between the TE and SWC (0.5) (Table 1). For all tests in this study, TE was below or similar to SWC, which is consistent with previous studies (Gabbert et al., 2008; Lockie et al., 2013; Thomas and Nelson, 2001).

Our results indicate that the AS test is a valid tool to assess agility and skill speed in female amateur soccer athletes. Mean times of the AS tests demonstrated its significant moderate-to-high relationships with all other sprint and agility tests measured in this study (Table 2). These results confirm that the AS tests assess physical abilities (i.e., quickness, change-of-direction speed) similar to other comparable measures. As observed in the present research, low, moderate, high and very high-level relationships exist between straight sprinting agility, power jumping agility, and sprint with directional-change agility (Gabbert, 2008; Lockie et al., 2013). In contrast, Sporis et al. (2011) found that there was no significant correlation between the agility test (zigzag with the ball) and speed and quickness (r = 0.093–0.247). Little and Williams (2005) observed particularly low correlations (r = 0.35) between a 10 m sprint test and a zigzag test that used three turns of 100º. Our study provides similar results; there were no or low significant correlations between agility and sprint tests (Table 2). Our results combined with findings from other studies (Buchheit et al., 2011; Gabbert et al., 2008) strongly suggest that the structure of a test of agility with the ball is much more complex compared to simple speed tests that previously have been applied.

The results of the present study demonstrate that while the IAR, CODAT, and T-Drill agility tests are sufficient to assess change-of-direction speed and agility, there are other measurements of skill difference that benefit from the incorporation of the AS tests. The sprint distances covered and directional changes performed in AS tests are applicable to a range of field sports. AS tests also have practical use and
relevance in measuring agility and active ball-shooting skills in field testing of athletes, including soccer players. Athletes of different age and ability levels might also benefit from the use of AS tests, as they are relatively brief and easy to use as sport-specific tests.

One limitation of the present study is that the AS tests were conducted in an indoor gymnasium; thus, we posit that they should also be conducted on an outdoor soccer field to confirm their reliability.

Analyses of mean test results in this study also indicate that considering the differences between the duration of the T-Drill classic (12.19 ± 0.55 s), AS (with a ball) (12.68 ± 0.57 s), CODAT (6.38 ± 0.28 s) and IAR (19.09 ± 0.72 s) (Table 2) tests, the evaluated AS (with a ball) test for females requires additional ball-shooting skills incorporating decision-making. Thus, the AS (with a ball) test produces less successful performance taking into account recorded time results compared to the T-Drill classic test with the same metric (p > 0.01; independent t-test). The differences in scores between the T-Drill test and the AS-with-a-ball test can be explained by the additional requirement for decision-making and ball-kicking skills which are included in the AS test activities. Successful performance in team sports, also in soccer, requires well-developed perceptual and decision-making skills that are demonstrated by superior anticipatory motor performance (Gabbett et al., 2008; Sheppard and Young, 2006). Most of the athletes who were in the higher-performance percentiles were able to complete the decision-making processes in a shorter period of time during ball-shooting. This observation is in agreement with investigations that have addressed the differences between domain-specific anticipation and decision-making skills of novices and experts in sport (Gabbett et al., 2008; Kowacks et al., 1999). However, in the present study, decision-making time was not measured directly; therefore, assumptions about the hypothesized superior cognitive abilities of the higher-performance athletes cannot be made with certainty.

This research is the first study to demonstrate the reliability and validity of a newly developed soccer-specific agility and skill test in measuring speed, directional change and ball-shooting skill in female amateur soccer athletes. Our testing procedures and results are consistent with other agility and sprint tests and related research studies. The new AS test is not time-consuming, is inexpensive and simple to perform in limited space with sport-specific movements for both females and males. The new AS test may have particular value for soccer in identifying talent at different ages. Future research into this test should be considered with control groups including not only sprint tests and similar agility tests, but also tests assessing additional soccer-specific skills in both female and male soccer players.

This research study has demonstrated the reliability and validity of a newly developed soccer-specific agility and skill test that combines the measurement of speed, directional change, and ball-shooting skill in female amateur soccer athletes. Coaches can now evaluate skills that have not been considered by other tests. It may be particularly useful for talent identification among young soccer players. Additionally, the AS test can also be used by regional school centers and local clubs that wish to rank players. From a practical perspective, the AS test is relatively easy to incorporate within a battery of talent identification tests for it requires minimal preparation, few materials, and is relatively efficient with regard to time. Moreover, our results indicate that minimal familiarization is necessary to be able to distinguish between skilled male and female soccer players.

Acknowledgements

The authors express gratitude to all players and coaches involved in this study.

References


Reliability and Validity of a New Test of Agility and Skill for Female Amateur Soccer Players


Kutlu M, Yapıcı H, Yoncalı O, Celik S. Comparison of a new test for agility and skill in soccer with other agility tests. J Hum Kinet, 2012; 33: 143-150


Reily T, Holmes MA. Preliminary analysis of selected soccer skills. Phys Ed Rev, 1983; 6: 64-71

Sheppard JM, Young WB. Agility literature review: Classifications, training and testing. *J Sports Sci.*, 2006; 24(9): 919-932


Corresponding author:

Mehmet Kutlu PhD
School of Physical Education and Sport, University of Istanbul Gelisim, Istanbul, Turkey
Phone: +90 (0212) 422 70 00 / 267
Fax: +90 (0212) 422 74 01
E-mails: mkutlu@gelisim.edu.tr, mkutlu2012@gmail.com, mkutlu2002@yahoo.com