The Development of a Field-Based Kicking Assessment to Evaluate Australian Football Kicking Proficiency

The development of a field-based kicking assessment to evaluate Australian Football kicking proficiency

Abstract

In Australian Football (AF), the ability to proficiently kick the ball is a critical skill and has been shown to be advantageous to a team's successful performance; however, a valid and reliable match referenced kicking assessment remains absent. Therefore, the aim of this study was to develop a valid and reliable AF kicking proficiency assessment comparative to match play kicking performance. Youth male Australian Football players (*n* = 251) from different stages within the AF talent pathway were recruited. The developed AFFB-DKA (Australian Football Field-Based – Dynamic Kicking Assessment) considered particular constraints of match play kicking demands such as kick type, distance; delivery; and locomotion of the player receiving the ball. In total, fourteen kicks were completed during the test. Validity (i.e., content, logical and construct) and reliability (i.e., test re-test) were assessed. Findings indicate the kicking test can distinguish across and between age (i.e., U14; U16; U18) and skill groups (i.e., club; sub-elite; elite). The timeframe between U14 and U16 was identified as a potential key period where kicking skill acquisition may be most impressionable; however, further research is recommended to support this. The developed AFFB-DKA is the first Australian Football specific kicking assessment to consider and apply match play kicking constraints to make a more representative, valid and reliable assessment.

The development of a field-based kicking assessment to assess Australian Football kicking proficiency

In Australian Football (AF), effective kicking is a critical technical skill to maintain ball possession and more importantly, scoring. Skilled performance is characterized by delivering the ball with more consistent and accurate technical skill executions (Bennett et al., 2017; Cripps, Joyce, Woods, & Hopper, 2017). A player's ability to successfully kick the ball to a teammate has been shown to be advantageous to their team's success (Robertson, Back, & Bartlett, 2015). Despite the importance of kicking in AF, for talent identification purposes, no assessment of technical kicking performance has been developed with sufficient content, logical and construct validity. These forms of validity are important to ensure the test measures what it claims to, is representative of match-play and can distinguish between known skill levels thereby effectively assessing the skill.

Recent AF performance analysis investigations have suggested greater ball possession and kicking skill proficiency relative to the opposition have been shown to have greater influences on match outcomes (Robertson et al., 2015). When an AF team was successful in winning a match, there was a reduced physical output (i.e., distance travelled by the team) and an increase in skill involvements (i.e., kicking and handballing) suggesting more technically skilful AF teams do not have to rely upon superior physical traits to perform at a high standard (Sullivan et al., 2014). A higher level of skill proficiency allows for a more effective delivery to either score, obtain and/or maintain possession of the ball. Effective kicking, particularly long kicks beyond 40m (Sullivan et al., 2014), is a crucial component to team success in AF. Stewart, Mitchell, and Stavros (2007) investigated which skill was the most important to winning games and attempted to quantify their contribution. They discovered when a player accurately delivered the ball over 40m it increased their team's average winning margin by 0.99 of a point, whilst an inaccurate kick (of any distance) to the opposition reduced their team's winning margin by 0.62 of a point. Despite evidence demonstrating the importance of kicking performance to AF team success, only two studies have investigated kicking proficiency as part of a multidimensional approach for talent identification (Tribolet, Bennett, Watsford, & Fransen, 2018; Woods, Raynor, Bruce, McDonald, & Robertson, 2016) and only two studies have investigated kicking proficiency in an attempt to predict playing status (Cripps et al., 2017; Woods, Raynor, Bruce, & McDonald, 2015).

Although game-based data does identify important elements of match performance (e.g., kicking effectiveness), it is not standardised. In an attempt to assess kicking performance of AF players, the Australian Football League (AFL) included a kicking test at the AFL National Combine (i.e., large scale standardized testing for talent identified 17-18 year-olds). The test involves a player running towards a feeder who receives the ball, turns and executes a kick to one of six randomly assigned stationary targets (Cripps, Hopper, & Joyce, 2015). Kicking performance is subjectively assessed on a scale from 0 to 5 (5 being the highest score) for each kick. Researchers have attempted to assess the validity and reliability of the test with conflicting results. Cripps and colleagues (2015) investigated the inter-rater reliability and validity of the AFL kicking and handballing tests. They assessed 121 semi elite under 16 male players and whilst they found the inter-rater reliability to be high, there was a poor correlation between the test score and coach's perception ($r = 0.13$) indicating a lack of validity. The authors found the test could differentiate between dominant and non-dominant kicking leg accuracy; however, they did not assess the ability of the test to discriminate between levels of ability. In another study Woods et al. (2015) compared biological maturation, anthropometric, physical and technical skill measures between talent and non-talent identified junior Australian footballers. They assessed 50 under 18 male players (25 state representatives and 25 non state representatives) and found when kicking accuracy and ball speed were combined playing status (i.e., sub-elite, elite) was able to be predicted. In addition to these results, a limitation of the current kicking test, is the assessment is conducted in isolation and does not assess the range of in-game kicking constraints (e.g., distance of the kick, movement patterns of players) typically performed within the performance environment. As a result, kicking ability is not assessed under match referenced conditions and consequently players may perform alternative actions and performances (Araujo, Davids, & Hristovski, 2006). Therefore, a significant gap remains within both applied research and industry for the development of a valid and reliable standardised match referenced kicking test that could be used for talent identification and monitoring of skill development.

The development of a new dynamic field-based kicking test would contain a more integrated approach of match-play components (i.e., technical, tactical, physiological, psychological) and a higher requirement from the performance demands (e.g., pressure) to be more representative of match play (Bonney, Berry, Ball, & Larkin, 2019). These factors are important to consider as they will influence how accurately a player will perform their skills (Ali, 2011). Players could then be assessed on their skill execution with an outcome focus (e.g., kick effectiveness) rather than a performance focus (e.g., mechanics of the kicking actions) to provide a greater indication of match day performance.

Numerous studies have developed methods to assess physiological qualities of AF players (for examples see Pearson, Naughton, & Torode, 2006; Pyne, Gardner, Sheehan, & Hopkins, 2005; Veale, Pearce, & Carlson, 2010); however, there remains no valid and reliable AF kicking test to assess and monitor kicking performance and development, with the potential aim to establish skill benchmarks for player identification purposes. Therefore, the aim of the current study is to develop a match referenced Australian Football Field-Based–Dynamic Kicking Assessment (AFFB – DKA) and establish the validity and reliability of the test to assess AF kicking performance.

Method

Participants

Male Australian Football players (n = 251) from different stages within the AF talent pathway were recruited. The players were recruited from a local club (club; U14: *n* = 28; U16: *n* = 69; U18: *n* = 47); a state junior representative group (sub-elite; U16: *n* = 22; U18: *n* = 26) and a national junior representative group (elite; U16: *n* = 4; U18: *n* = 55).

Test Development

When developing new assessments, validity is an important consideration as it ensures the test measures what it claims to measure (Thomas, Nelson, & Silverman, 2011). To establish the content validity of the test, in-game kicking constraints were considered. Notational analysis of six under 18 sub-elite level matches (Mantle, 2017) and 19 AFL matches (Back, 2015) was used to quantify the in-game kicking constraints. The analysis revealed specific in-game kicking demands such as the distance of the kick (i.e., more kicks are performed <40m than >40m); how the kick is delivered (i.e., more kicks are delivered from a stationary position than from a running movement); locomotion of the player receiving the ball (i.e., more kicks are delivered to a leading player than a stationary player); and the type of kick being performed (e.g., switching play to the opposite side of the ground; kicking for goal; or delivering the ball over a teammate's shoulder whilst they are running away towards goal). To ensure the kicking test was more representative of the game constraints, these kicking demands were used as a guide and implemented where possible, to develop a more dynamic kicking assessment.

A suggested method to assess how player kicking behaviours performed during the performance environment compare to player kicking behaviours performed during the test was to assess action fidelity (Travassos, Duarte, Vilar, Davids, & Araujo, 2012). To provide a level of fidelity the in-game kicking constraints were compared between the AFFB-DKA and sub-elite under 18 (Mantle, 2017) and senior elite level match play (Back, 2015). The AFFB-DKA required players to deliver the ball with 85% of kicks <40m (U18 56% (Mantle, 2017); senior elite 63% (Back, 2015)) and 15% >40m (U18 44%; senior elite 37%); deliver the ball with 57% from a stationary position (U18 61%; senior elite 57%) and 43% from a running position (U18 39%; senior elite 43%); and players receiving the ball will be in a stationary position in 15% of events (U18 34%; senior elite 35%) and running for the ball in 85% of events (U18 66%; senior elite 65%). The test also required players to kick the ball to the opposite side of the ground (switch play) and kick for goal to replicate match play kicking patterns.

Further constraints were applied to the test to ensure particular patterns of behaviour would occur (Davids, Renshaw, & Glazier, 2005). For example, the test required players to only perform drop punt kicks, as this is the preferred kick in AF due to its accuracy and high speed of execution (Rendell, Masters, Farrow, & Morris, 2011). Other constraints included participants delivering the ball with only their dominant leg, no handballs and no opposition players, to enable the focus of the test to be on kicking proficiency. All sections of the ground were utilized to more accurately replicate the movement patterns of competitive play (Loader, Montgomery, Williams, Lorenzen, & Kemp, 2012).

To assess the content validity of the test, a pilot study was conducted to obtain feedback on the design, functionality and assessment process from an expert panel including current elite AFL coaches (*n* = 2); a current elite under 18 representative squad coach (*n* = 1); recently retired elite AFL players (*n* = 4); current sub-elite under 18 AF players (*n* = 26); skill acquisition experts (*n* = 3); a biomechanist (*n* = 1); and a senior sport scientist working within an AFL club (*n* = 1). Feedback was obtained from the panel and slight modifications to the procedure of the test were applied. For example, it was suggested (and applied) to advise participants to only kick the ball when they believed they could successfully deliver the ball to their teammate (i.e., on stations were a 40m kick was to be conducted, players could run closer to the receiver as long as the kick was still perceived as being long). In addition, whilst it is acknowledged elements of the test are not directly representative of the performance environment (e.g., no opposition), feedback from the panel revealed it was an appropriate balance between controlling constraints (e.g., starting positions) and having an open, dynamic test (e.g., the ability of the kicker to adjust their kick to meet the demands of the receiver). Furthermore, it was not the intent of this paper to create a test that is a perfect correlation to the performance setting, merely to design a test that is a progression along the continuum and to fill the void between static tests and match play (Bonney et al., 2019).

Procedure

An AF oval was set up as outlined in Figure 1. Cones were placed on the field, where the green triangles are located, for players to commence their lead from (i.e., running to receive the ball). Two Defender Man mannequins (Knight Sport) were placed at the goal kicking end of the field 25m out, on a 45-degree angle, for stations 7 and 14 to kick the ball over. In total, 14 kicks were performed in the test, seven going in one direction and then the same seven kicks going back in the mirror image. For example, stations one and eight require the player to perform a stationary kick to a player on the lead at approximately a 45°angle. The duration of the test was approximately 20 minutes, with 15 players completing the test in this time. The length of the test was representative of one quarter; however, the number of kicks performed in this timeframe was not. Similar duration lengths have previously been reported (Robertson, Burnett, & Cochrane, 2014).

******Figure 1 near here******

To increase the stabilization of performance data, the protocol for testing included a requirement for familiarization. Research has shown when trials have been included in a reliability analysis the coefficient of variation between the first two trials is 1.3-fold greater than between following trials (Hopkins, Schabort, & Hawley, 2001). A small replica test was set-up on the side line to initially walk the players through the activity. Once completed players commenced a practice session of approximately ten minutes to allow the players time to familiarize themselves with the test, as recommended by Currell and Jeukendrup (2008). This time allocation allowed each of the players to experience each of the seven unique kicks they were required to complete during the test. A five-minute break then occurred before the test commenced.

Fifteen players, the number of players involved in each test, were randomly allocated numbered bibs and a GPS unit (Catapult, Minimax S5) to wear. Each bib number corresponded with the station number the player commenced on the test. The test commenced with player 1 kicking the ball out from the goal square (i.e., Station 1). Each participant followed their kick, moving to the next station and waiting for the ball to come back around to them. This procedure continued until all fourteen kicks were completed. Players were instructed to complete the test at 'game speed' to simulate match requirements. The receiver was instructed to lead for the ball once the kicker had the ball in their hands. Each kick required the kicker to deliver the ball at different angles and speeds (e.g., to a teammate leading in a particular direction at a particular speed).

Three cameras were positioned on the field to capture the test performance. One camera was positioned behind the goals facing the middle of the ground (i.e., to capture footage within the 50m zone), whilst two cameras were on opposite sides of the field, ten meters in from the boundary line at approximately the half way mark (i.e., to capture test footage in that specific side of the field). The test finished once all participants had completed the kick at each corresponding station.

Logical validity is seen as the base level validity (Lather, 1986) and highlights how the study is deemed to be worthwhile; which assists in gaining acceptance by the population being tested. When considering logical validity, the test should closely resemble the natural environment in which the skill is conducted (Larkin, Mesagno, Berry, & Spittle, 2014). To assess for logical validity, approximately ten minutes after the test players and coaches were asked to complete a five-point Likert scale questionnaire which had 12 questions pertaining to the assessment effectiveness and how comparative the kicking patterns, kicking distances, test intensity and test suitability were in comparison to match play (i.e., the kicking patterns used in this test are similar to that performed in match play at your level).

Reliability is an important consideration as without reliability assessment it is unknown whether a change in a particular performance is a result of development or due to an unreliable test (Larkin et al., 2014). The stability of the test performances was determined by test re-test reliability approximately one week apart (Lubans, Smith, Harries, Barnett, & Faigenbaum, 2014). On both test occasions, standardised procedures and instructions were followed. The stability of individual responses was determined by one sub-section of sub-elite U18 participants (n =15) completing the protocol on two occasions, seven days apart, as long retest intervals can result in large variations due to factors such as participant behaviour or circadian variations (Robertson et al., 2014). Inter-rater reliability was examined with two trained independent assessors analysing 30% (105 players, 1575 kicks) of the kicks using the scoring procedure outline in Table 1. The kappa (k) correlation was interpreted as follow: poor (<0.20), fair (≤0.20-0.40), moderate (≤0.40-0.60), good (≤0.60-0.80) and very good (≤0.80-1.00) (Altman, 1991). The result was k = 0.92 and classified as very good.

Data Analysis

Video footage from the three cameras were stacked (i.e., having the three camera angles showing on the one screen side-by-side) and coded using SportsCode 10.3.25. Each view provided different contextual pieces of information. For example, the camera from behind the goal displayed footage of player 1 and 8 kicking the ball whilst the cameras on the side of the oval displayed footage of players 2 and 9 receiving the ball. Each kick was scored according to how accurate the kick was executed. The scoring sheet (see Table 1) was developed in conjunction with coaches and sport scientists working within the AFL to ensure it was representative of match play kicking actions. Scoring was tailored for the kicking player when they were kicking to a player who was not leading. For example, five points were awarded to the kicking player when the receiving player was stationary and marked the ball out in front without moving. Kicks 2-6 and 8-13 were scored from 0-5 (i.e., 5 being the highest point value for the most accurate kick) whilst kicks 7 and 14 were scored either 0, 1 or 6 as they were kicks directed towards goal and this is the scoring system used in the AFL (*Laws of Australian Football*, 2015). The maximum score a player could achieve on the test was 72.

******Table 1 near here******

Logical validity gathered from the Likert scale questionnaire is presented in the form of descriptive statistics and was assessed by a cross-section of elite under 18 players (n = 15), sub-elite under 18 players (n = 15), elite under 18 representative squad coaches (n = 3), and skill acquisition experts (n = 2). The mean and standard deviation for each topic section were calculated from the 5-point Likert scale (i.e., 1, strongly disagree; 2, disagree; 3, neutral; 4, agree; 5, strongly agree) (Boone & Boone, 2012). Mean results were classified as strongly disagree (1-1.9), disagree (2-2.9), agree (3-3.9), strongly agree (4-4.9). Likert scale questions were provided to two senior sport scientists and one elite U18 representative coach for feedback. To ensure reliability sub-elite under 18 players (n = 10), and an elite under 18 representative coach (n = 1), were given the same questionnaire on two separate occasions, one week apart. Their results were assessed using Cronbach's alpha with a score of .906, indicating excellent reliability (Helmerhorst, Brage, Warren, Besson, & Ekelund, 2012).

To determine the construct validity a One-Way ANOVA was used to compare between groups (e.g., U14, U16 and U18) (independent variable) and kicking proficiency score (dependent variable) and within group comparisons assessing skill (e.g., U18 club, U18 sub-elite and U18 elite) (independent variable) and the kicking proficiency score (dependent variable). A multinomial logistic regression (Thomas et al., 2011) was used to determine if kicking proficiency percentage could accurately classify age (i.e., U14, U16, U18) or skill groups (i.e., club, sub-elite, elite) of players.

The Catapult Sprint 5.1.7 software was used to download the GPS data. Individual player meters travelled per minute ($m\text{-}min^{-1}$) and maximum velocity (max. velocity) were recorded as measures of central tendency. Significance for data sets were set at p = <.05. Effect sizes (ES) for ANOVAs were reported as partial eta squared (η_p^2) (Olejnik & Algina, 2003) and post hoc effect sizes were calculated using Cohen's *d* with 95% confidence intervals (CI) (Cohen, 1988). All other calculations were made using the statistical package SPSS Statistics (SPSS Version 25.0).

The consistency of the test was measured by test-retest reliability using the intra-class correlation coefficient (ICC) with 95% confidence limits, the coefficient of variation (CV) and the standard error of measurement (SEM) (Hopkins, Marshall, Batterham, & Hanin, 2009). For intra-rater reliability, the ICC was measured by one tester analysing the same video recorded testing session, involving 15 participants, on two separate occasions seven days apart. The ICC classifications used were 0.81-1.00 Almost perfect, 0.61-0.80 Substantial, .41-.60 Moderate, 0.21-0.40 Fair and 0.00-0.20 Slight (Landis & Koch, 1977).

Results

Logical validity was supported through both players (mean ± SD; 3.60 ± 0.66) and coaches (3.88 ± 0.78) agreeing the test can assess kicking proficiency and identify potential player weakness. Further, players (3.58 ± 0.79) and coaches (3.10 ± 0.71) agreed the test used patterns and distances similar to match play although the coaches (2.67 ± 0.95) did not feel the drill simulated the time constraints and intensity experienced during match play. Both players (3.96 ± 0.76) and coaches (3.63 ± 0.95) agreed the kicking test was suitable for the age and ability of the groups tested.

A one-way between subjects ANOVA was conducted to compare the effect of age on kicking proficiency at the under 14 (mean ± SD; 41.90 ± 5.66), under 16 (54.93 ± 8.57) and under 18 (61.51 ± 8.82) age groups. There was a significant effect of age on kicking proficiency at the $p<.05$ level for the three age groups [$F(2, 237) = 48.23$, $p = 0.000$, $\eta_p{}^2 = .289$]. Post hoc comparisons using the Tukey HSD and the Cohen's *d* tests indicated the mean score for the under 14 group was significantly different with a large ES compared to the under 16 group (p = .000, Cohen's *d* (95% CI) = 1.2 (.76 – 1.70)) and a very large ES compared to the under 18 group (p = .000, Cohen's *d* (95% CI) = 1.9 (1.48 – 2.41)). The under 16 group was also significantly different from the under 18 group with a medium to large ES (p = .000, Cohen's *d* (95% CI) = .76 (.45 – 1.05)). A multinomial regression analysis was conducted, using kicking proficiency percentage as a predictor of age group. This analysis identified the AFFB-DKA could correctly identify 60.7% of U14 players, 51.2% of U16 players and 76.6% of U18 players with an overall accuracy of 65.8%. While this study was not longitudinal in nature, analysis of the kicking performance across age groups show an increasing trend.

When comparing kicking proficiency within skill groups, a significant effect was not found at the $p<.05$ level for the U16 age group for the three conditions of club, sub-elite and elite [$F(2,86) = 1.29$, $p = 0.289$, $\eta_p{}^2 = .029$]. However, there was a significant effect of skill on kicking proficiency at the $p<.05$ level for the three under 18 conditions of club, sub-elite and elite [$F(2,157) = 19.71$, $p = 0.000$, $\eta_p{}^2 = .201$]. Post hoc comparisons using the Tukey HSD and Cohen's *d* tests indicated the mean score for the club group was not significantly different to the sub-elite group (p = 0.48, Cohen's *d* (95% CI) = .46 (-.03 – 0.94)) but was significantly different to the elite group with a large ES (p = .000, Cohen's *d* (95% CI) = 1.12(0.70 – 1.53)). The sub-elite group was also significantly different with a moderate to large ES to the elite group (p = .001, Cohen's *d* (95% CI) = .76 (0.25 – 1.21)). A multinomial regression analysis was conducted, using kicking proficiency percentage as a predictor of skill group. This analysis identified the AFFB-DKA could correctly identify 86.1% of novice players, 16.2% of sub-elite players, 57.6% of elite players with an overall accuracy of 68.3% of players. Taken together, these results support the construct validity of the test and suggest as age (U14 to U18) and skill (club to elite) increase so too does kicking proficiency. On average, players increased kicking proficiency at the club skill level by 9.29% from under 14 to under 16 and by 2.77% from under 16 to under 18 (Table 2). At the sub-elite level, players increased kicking proficiency from under 16 to under 18 by 1.22% and by 6.18% at the elite skill level.

******Table 2 near here******

A one-way between subjects ANOVA was conducted to compare the effect of age on meters ran per minute at the under 14 (mean ± SD; 41.90 ± 3.19), under 16 (54.69 ± 6.75) and under 18 (57.09 ± 5.08) age groups. There was a significant effect of age on meters ran per minute (m-min^{-1}) at the $p<.05$ level for the three age groups [$F(2, 211) = 70.75$, $p = .000$, $\eta_p^2 = .541$]. Post hoc comparisons using the Tukey HSD and Cohen's *d* tests indicated a very large difference between the under 14 group and the under 16 group (p = .000, Cohen's *d* (95% CI) = 2.17 (1.58 – 2.72)) and the under 14 group and the under 18 group (p = .000, Cohen's *d* (95% CI) = 3.31 (2.60 – 3.96)). The under 16 group was not found to be significantly different to the under 18 group (p = .057). A similar result was found when comparing the effect of age on average maximum velocity at the under 14 (mean ± SD; 5.52 ± 0.29), under 16 (6.81 ± 0.56) and under 18 (6.92 ± 0.60) age groups. There was a significant effect of age on average maximum velocity at the $p<.05$ level for the three age groups [$F(2,211) = 69.125$, $p = .000$, $\eta_p^2 = .527$]. Post hoc comparisons using the Tukey HSD and Cohen's *d* tests indicated a very large ES between the under 14 and under 16 groups (p = .000, Cohen's *d* (95% CI) = 2.55 (1.98 – 3.08)), under 14 and under 18 groups (p = .000, Cohen's *d* (95% CI) = 2.68 (2.05 – 3.26)) group. The under 16 group was not found to be significantly different to the under 18 group (p = .487).

There was a significant effect of skill on the amount of meters ran per minute [F(2, 108) = 19.603, p = 000, $\eta_p{}^2$ = .416] at the p<.05 level for the three conditions of club, sub-elite and elite. Post hoc comparisons for the under 16 group indicated a large difference between club and sub-elite players (p = .000, Cohen's *d* (95% CI) = 1.30 (0.62 – 1.94)) with no significant differences found between club and elite players (p = 0.89) or sub-elite and elite players (p = .941) (Figure 2). When comparing the under 18 group results indicate a large negative ES between club and sub-elite players (i.e., club players performed at a higher intensity) (p = .000, Cohen's *d* (95% CI) = -1.29 (-1.92 – -0.61)), no significant difference between club and elite players (p = .070) and a large ES between sub-elite and elite players (p = .000, Cohen's *d* (95% CI) = 2.25 (1.38 – 3.02)).

A significant effect of skill was found on the average maximum velocity at the p<.05 level for the three conditions of club, sub-elite and elite [F(2,108) = 13.080, p = 0.000, $\eta_p{}^2$ = .322]. Post hoc comparisons for the under 16 group indicated a moderate ES between club and sub-elite players (p = .044, Cohen's *d* (95% CI) = 0.72 (0.13 – 1.31)) with no significant differences found between club and elite players (p = .199) and sub-elite and elite players (p = .951). When comparing the under 18 group results indicate a large negative ES between club and sub-elite players (p = .001, Cohen's *d* (95% CI) = -1.24 (-1.87 – -0.57)), no significant difference between club and elite players (p = .659) and a large ES between sub-elite and elite players (p = .000, Cohen's *d* (95% CI) = 1.52 (0.75 – 2.22)).

Taken together, these results suggest as age increases so too does the running intensity of the players; however, trying to differentiate between skill groups, using running intensity parameters, is less insightful.

******Figure 2 near here******

For the reliability of the test, the ICC ± 95% CL, CV and the SEM indicated good reliability between the test re-test assessment (ICC = 0.90 ± 0.78-0.96, CV = 9.84, SEM = 1.89), with intra-rater correlation being almost perfect (ICC = 0.96 ± 0.85-0.99).

Discussion

This study effectively assessed the validity and reliability of a novel Australian Football kicking proficiency test. The AFFB-DKA is the first Australian Football specific skill test to consider and apply match play kicking constraints to make a more representative valid assessment. Content and logical validity was supported through players and coaches rating the test as representative of match play patterns and distances and can assess kicking proficiency whilst identifying player kicking weaknesses. Construct validity was supported through the test successfully identifying an increase in kicking proficiency as age and skill level increase. The test could also identify an increase in physical output as age increased; however, results were less apparent when analysing from a skill level perspective. Reliability was supported through test re-test.

Obtaining and applying notational analysis is the foundation stone of skill-based assessments (Bonney et al., 2019). Notational analysis enables the skill assessment to include more match specific constraints (e.g., time pressure), to help provide a more accurate assessment of skill (Bonney et al., 2019). Careful consideration was given in the design process of the AFFB-DKA to ensure, where possible, match specific kicks transpired and kicking constraints were applied. To assess the content validity of the AFFB-DKA, an expert panel of coaches, players and sport scientists reviewed and provided feedback to the development of the test. Larkin et al. (2014) recommend this approach suggesting the involvement of an expert panel can provide more scientific rigor to the content validity of a study. The notational analysis used for the design and implementation of the AFFB-DKA, to the specific age groups, were supported by the expert panel confirming content validity.

Logical validity was supported by players and coaches agreeing the AFFB–DKA can assess kicking proficiency and identify player weaknesses. They also agreed the test applied kicking patterns and distances representative of match play. Finally, they agreed the test was suitable to the age and ability level of the players assessed. The coaches did, however, note the time each player had to deliver the ball and the intensity at which they were working was different to that performed during match play. This observation was supported when comparing under 14 running intensity demands in the AFFB-DKA to match play. During the AFFB-DKA the amount of meters travelled per minute and average maximum velocity was approximately 42% and 80% of those reported during match play (Gastin, Tangalos, Torres, & Robertson, 2017). It should be noted however, the averages reported in the Gastin et al. (2017) study was from whole matches and may not represent times when the players were kicking the ball. For example, players may reach peak speed during match play when chasing an opponent whereas the intensities noted in this study are measurements of player speed during skill execution (i.e., kicking). A limitation of the current assessment was the AFFB-DKA occurred when the team's coaching staff were either talking to other players or observing silently from the side. This may have contributed to the lower intensity levels demonstrated in the test, as researchers have highlighted the importance of coach-athlete interaction in relation to physical performance (Selmi et al., 2017). Furthermore, the intention of this drill was to assess kicking proficiency without defenders. The absence of defenders (and therefore pressure to timely deliver the ball) may have afforded players more time to proficiently kick the ball at the expense of completing the task under constraints similar to match play. Finally, this research only assessed kicking proficiency on the dominant leg with further research being required to see if results change when assessing the non-dominant leg.

For the establishment of construct validity, the test should differentiate between known performance levels (Thomas et al., 2011). To date, no AF kicking test has reported the construct validity of the assessment prior to its implementation. Therefore, without assessment of construct validity prior to the implementation of a new kicking test, it may be unclear whether performance differences are a result of skill differences or an unreliable test (Larkin et al., 2014). The results of this study demonstrate the AFFB-DKA was successful in distinguishing between players across age groups (i.e., U14, U16 and U18) and across skill levels within the U18 age group (i.e., U18 club, U18 sub-elite and U18 elite) of the AFL talent pathway.

Finally, based on the results of the current study, these data may indicate there are key stages in the development of kicking performance. From a skill acquisition perspective, kicking skill development at a club level may be most evident between the age groups under 14 and under 16 (9.29% increase in comparison to the under 16 to under 18 age groups 2.77%); however, further research, using a longitudinal design, is required to consolidate these results. If shown to be accurate, this may highlight an important timeframe where kicking acquisition may be most susceptible to technique change. For example, player kicking motor pattern sequencing appears to be developing at the greatest rate between the age groups under 14 and under 16. However, it should also be noted the low levels of discriminatory ability of kicking performance at the under 16 age group. This may suggest kicking proficiency, at the under 16 age group, may not be a primary determinant of playing performance and provides a worthwhile area for future research.

Conclusion

This study contributes to AF talent identification research as it assessed the validity (i.e., content, logical and construct) and reliability (i.e., test re-test) of a new AF kicking test. Specifically, this study provided evidence the kicking assessment can distinguish across and between age and skill groups. As players progress along the AFL pathway so too does their kicking proficiency, with data indicating players between the age groups of under 14 to under 16 having the greatest percentage rate of improvement; however, further investigations are required to confirm this. Future investigations can use this test as a valid and reliable measure to assess or monitor kicking performance changes along the AFL pathway, providing a development continuum where talent is objectively assessed in a more inclusive rather than exclusive manner. Furthermore, this study indicates the stages and assessment methods that could potentially be used to develop a new valid and reliable performance skill assessment in other sport contexts.

References

Ali, A. (2011). Measuring soccer skill performance: a review. *Scand J Med Sci Sports, 21*(2), 170-183. doi:10.1111/j.1600-0838.2010.01256.x

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall.

Araujo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise, 7*(6), 653-676.

Back, N. (2015). *The influence of constraints on athlete kicking performance in training and matches at an elite Australian Rules football club.* (Bachelor of Applied Science Honours), Victoria University, Melbourne.

Bennett, K., Novak, A., Pluss, M., Stevens, C., Coutts, A., & Fransen, J. (2017). The use of small-sided games to assess skill proficiency in youth soccer players: a talent identification tool. *Science and Medicine in Football*. doi:https://doi.org/10.1080/24733938.2017.1413246

Bonney, N., Berry, J., Ball, K., & Larkin, P. (2019). Australian Football Skill-Based Assessments: A proposed model for future research. *Frontiers in Psychology*. doi:10.3389/fpsyg.2019.00429

Boone, H., & Boone, D. (2012). Analyzing Likert Data. *Journal of Extension, 50*(2).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: Erlbaum Associates.

Cripps, A., Joyce, C., Woods, C., & Hopper, L. (2017). Biological maturity and the anthropometric, physical and technical assessment of talent identified U16 Australian footballers. *International Journal of Sports Science and Coaching, 12*(3), 344-350.

Cripps, A. J., Hopper, L. S., & Joyce, C. (2015). Inter-Rater Reliability and Validity of the Australian Football League's Kicking and Handball Tests. *J Sports Sci Med, 14*(3), 675-680.

Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports Med, 38*(4), 297-316.

Davids, K., Renshaw, I., & Glazier, P. (2005). Movement models from sports reveal fundamental insights into coordination processes. *Exerc Sport Sci Rev, 33*(1), 36-42.

Gastin, P. B., Tangalos, C., Torres, L., & Robertson, S. (2017). Match running performance and skill execution improves with age but not the number of disposals in young Australian footballers. *J Sports Sci, 35*(24), 2397-2404. doi:10.1080/02640414.2016.1271137

Helmerhorst, H. J., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act, 9*, 103. doi:10.1186/1479-5868-9-103

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc, 41*(1), 3-13. doi:10.1249/MSS.0b013e31818cb278

Hopkins, W. G., Schabort, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Med, 31*(3), 211-234.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Larkin, P., Mesagno, C., Berry, J., & Spittle, M. (2014). Development of a valid and reliable video-based decision-making test for Australian football umpires. *J Sci Med Sport, 17*(5), 552-555. doi:10.1016/j.jsams.2013.08.001

Lather, P. (1986). Issues of Validity in Openly Ideological Research: Between a Rock and a Soft Place. *Interchange, 17*(4), 63-84.

Laws of Australian Football. (2015). Australian Football League.

Loader, J., Montgomery, P., Williams, M., Lorenzen, C., & Kemp, J. (2012). Classifying Training Drills Based on Movement Demands in Australian Football. *International Journal of Sports Science and Coaching, 7*(1).

Lubans, D. R., Smith, J. J., Harries, S. K., Barnett, L. M., & Faigenbaum, A. D. (2014). Development, test-retest reliability, and construct validity of the resistance training skills battery. *J Strength Cond Res, 28*(5), 1373-1380. doi:10.1519/JSC.0b013e31829b5527

Mantle, B. (2017). *Profiling kicking skill demands in elite Under-18 Australian Football.* (Honors), Victoria University, Melbourne.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol Methods, 8*(4), 434-447. doi:10.1037/1082-989X.8.4.434

Pearson, D. T., Naughton, G. A., & Torode, M. (2006). Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports. *J Sci Med Sport, 9*(4), 277-287. doi:10.1016/j.jsams.2006.05.020

Pyne, D. B., Gardner, A. S., Sheehan, K., & Hopkins, W. G. (2005). Fitness testing and career progression in AFL football. *J Sci Med Sport, 8*(3), 321-332.

Rendell, M. A., Masters, R. S., Farrow, D., & Morris, T. (2011). An implicit basis for the retention benefits of random practice. *J Mot Behav, 43*(1), 1-13. doi:10.1080/00222895.2010.530304

Robertson, S., Back, N., & Bartlett, J. D. (2015). Explaining match outcome in elite Australian Rules football using team performance indicators. *J Sports Sci*, 1-8. doi:10.1080/02640414.2015.1066026

Robertson, S. J., Burnett, A. F., & Cochrane, J. (2014). Tests examining skill outcomes in sport: a systematic review of measurement properties and feasibility. *Sports Med, 44*(4), 501-518. doi:10.1007/s40279-013-0131-0

Selmi, O., Khalifa, W., Ouerghi, N., Amara, F., Zouaoui, M., & Bouassida, A. (2017). Effect of Verbal Coach Encouragement on Small Sided Games Intensity and Perceived Enjoyment in Youth Soccer Players. *Journal of Athletic Enhancement, 6*(3).

Stewart, M., Mitchell, H., & Stavros, C. (2007). Moneyball applied: Econometrics and the identification and recruitment of elite Australian footballers. *International Journal of Sport Finance, 2*, 231-248.

Sullivan, C., Bilsborough, J. C., Cianciosi, M., Hocking, J., Cordy, J., & Coutts, A. J. (2014). Match score affects activity profile and skill performance in professional Australian Football players. *J Sci Med Sport, 17*(3), 326-331. doi:10.1016/j.jsams.2013.05.001

Thomas, J., Nelson, J., & Silverman, S. (2011). *Research Methods in Physical Activity, 6th edition*. Champaign: Human Kinetics.

Travassos, B., Duarte, R., Vilar, L., Davids, K., & Araujo, D. (2012). Practice task design in team sports: representativeness enhanced by increasing opportunities for action. *J Sports Sci, 30*(13), 1447-1454. doi:10.1080/02640414.2012.712716

Tribolet, R., Bennett, K. J. M., Watsford, M. L., & Fransen, J. (2018). A multidimensional approach to talent identification and selection in high-level youth Australian Football players. *J Sports Sci*, 1-7. doi:10.1080/02640414.2018.1468301

Veale, J. P., Pearce, A. J., & Carlson, J. S. (2010). Reliability and validity of a reactive agility test for Australian football. *Int J Sports Physiol Perform, 5*(2), 239-248.

Woods, C. T., Raynor, A. J., Bruce, L., McDonald, Z., & Robertson, S. (2016). The application of a multi-dimensional assessment approach to talent identification in Australian football. *J Sports Sci*, *34*(14), 1340-1345. doi:10.1080/02640414.2016.1142668

Woods, C. T., Raynor, J. A., Bruce, L., & McDonald, Z. (2015). The use of skill tests to predict status in junior Australian football. *J Sports Sci*, *33*(11), 1132-1140. doi:10.1080/02640414.2014.986501

TABLES

Table 1 *Description of the scoring system used to assess kicking ability within the AFFB-DKA*

Score	Description
5	Perfect Pass - to the lead or stationary position offered by teammate (advantage side) Leading Target: • Player receives ball at pace. Kicker forces the receiver to meet the ball • Kick is directed towards chest area; receiver makes no break in stride or running direction Stationary Target: • Receiving player does not move, marks out in front • If the receiver is shuffling away from the cone and the kicker matches this movement with the kick, a score of 5 is appropriate
4	Ball Out in Front Leading Target: • Receiving player has to change direction and/or jump to meet ball Stationary Target: • Receiving player moves comfortably to mark the ball (within 3-meters and/or jump)

3	Receiver in Trouble – opposition player has a chance to spoil Leading Target: • Receiving player has to slow and/or prop and/or adjust position significantly • Ball does not hit the ground • If receiver has to turn body to mark - default to Score 1 Stationary Target: • Receiving player moves hard to mark or forced to bend low (below knees at stretch), or moves backward and/or is forced to mark high overhead at a stretch
2	Missed Target - ball hits the ground to advantage. Leading Target: • Receiving player still has a realistic chance of gathering the ball in a game situation • Player receives ball at pace. Typically, one-bounce and run onto the ball • 75/25% chance of retaining possession • Ball must land within 5-meters • If ball lands at receiving player's feet then default to Score 1 (very hard gather) Stationary Target: • Default to Score 1 if ball lands on the ground

1	Missed Target - ball hits the ground to disadvantage Leading Target: • Ball must still land within 5-metres of receiving player if not, default to Score 0 • 25/75% chance or less of retaining possession Stationary Target: • Ball lands within 5-meters on the ground
0	Turnover • Ball does not land within 5-meters of the target area • In a game situation there would be a turnover (opposition would gain possession of the ball)

Note: Kicks 7 and 14 on the AFFB-DKA were scored as follows:

6 points – The ball goes through for a goal on the full. Player kicks ball over 'defender man'

1 point – The ball goes through for a point on the full. Player kicks ball over 'defender man'

0 points – The ball does not go over the goal line or the point line on the full and/or player does not kick the ball over the 'defender man'

Table 2 *Age and skill level mean (95% confidence intervals (CI)), standard deviation (SD) and standard error of measurement (SEM) for kicking proficiency percentage, meters travelled per minute and average maximum velocity*

		Club			Sub-elite			Elite		
		Mean (95% CI)	**SD**	**SEM**	**Mean (95% CI)**	**SD**	**SEM**	**Mean (95% CI)**	**SD**	**SEM**
Kicking proficiency	U14	45.19 (42.99-47.39)	5.66	1.79						
	U16	54.48 (52.43-56.53)	8.41	2.66	57.22 (52.45-61.99)	8.32	2.63	59.93 (49.68-69.77)*	5.47	1.73
	U18	57.25 (54.09-59.08)	8.02	2.66	58.44 (56.31-60.57)	6.41	1.63	66.11 (64.02-68.21)[ab]	7.75	2.43
Meters travelled per minute	U14	41.90 (40.62-43.19)	3.13	0.69						
	U16	52.32 (50.10-54.54)	6.77	2.14	60.06 (58.75-61.36)[c]	2.17	0.69	58.95 (56.89-61.01)*	1.12	0.35
	U18	58.24 (56.14-60.34)	4.37	1.38	53.35 (51.93-54.77)[a]	3.21	61.29	61.29 (59.00-63.57)[b]	3.99	1.26
Average maximum velocity	U14	5.52 (5.41-5.63)	0.28	0.09						
	U16	6.72 (6.37-6.67)	0.55	0.15	7.11 (6.82-7.40)[c]	0.48	0.15	7.20 (6.53-7.88)*	0.37	0.12
	U18	7.12 (6.90-7.34)	0.46	0.15	6.50 (6.27-6.73)[a]	0.53	0.17	7.27 (7.00-7.54)[b]	0.47	0.15

CI = Confidence Intervals; SD = standard deviation; SEM = standard error of measurement
* = Only 4 players were analysed in this age and skill group
[a] = Denotes a significant difference to the U18 club group; [b] = denotes a significant difference to the U18 sub-elite group
[c] = Denotes a significant difference to the U16 club group;

TITLES OF FIGURES

Figure 1. Schematic of stations 1-7 of the AFFB-DKA (stations 8-14 are the same in-reverse)

Figure 2. Kicking proficiency and running demands of the AFFB-DKA with SD error bars

Kicking Test Questionnaire

Name: ______________________________

Club: ______________________________

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
Player Assessment					
1. This test accurately assesses kicking proficiency	1	2	3	4	5
2. This test can distinguish between higher and lesser skilled players	1	2	3	4	5
3. This test assesses player ball control	1	2	3	4	5
4. This test can identify player kicking weakness	1	2	3	4	5
Game Simulation					
5. The kicking patterns used in this test are similar to how it would be performed in match play at your level	1	2	3	4	5
6. The distances used in this test simulate distances performed in match play at your level	1	2	3	4	5
7. The time the player had to kick the ball was similar to that performed in match play at your level	1	2	3	4	5
8. The intensity at which the players were working was similar to that performed in match play at your level	1	2	3	4	5
Test Suitability – Coach only to complete					
9. This test is appropriate for this age group	1	2	3	4	5
10. This test is appropriate for this ability level	1	2	3	4	5
11. I will use this test in the future to test my players	1	2	3	4	5

12. I will use this test in the future to help select my team	1	2	3	4	5