

A valid and reliable test of technical skill for vision impaired football

Oliver R. Runswick , Alexander Rawlinson , Naomi Datson & Peter M Allen

To cite this article: Oliver R. Runswick , Alexander Rawlinson , Naomi Datson & Peter M Allen (2021): A valid and reliable test of technical skill for vision impaired football, Science and Medicine in Football, DOI: [10.1080/24733938.2021.1885725](https://doi.org/10.1080/24733938.2021.1885725)

To link to this article: <https://doi.org/10.1080/24733938.2021.1885725>



Accepted author version posted online: 04 Feb 2021.



Submit your article to this journal [↗](#)



Article views: 12



View related articles [↗](#)



View Crossmark data [↗](#)

Publisher: Taylor & Francis & Informa UK Limited, trading as Taylor & Francis Group

Journal: *Science and Medicine in Football*

DOI: 10.1080/24733938.2021.1885725

Running Head: A technical skill test for VI football

A valid and reliable test of technical skill for vision impaired football

Oliver R. Runswick (oliver.runswick@kcl.ac.uk)^{1 2}

Alexander Rawlinson¹

Naomi Datson¹

Peter M Allen³

¹ Institute of Sport, University of Chichester, Chichester, UK

² Department of Psychology, Institute of Psychiatry Psychology & Neuroscience, King's
College London, London, UK

³ School of Psychology and Sports Science, Anglia Ruskin University, Cambridge, UK

Corresponding Author: Oliver Runswick

Department of Psychology

King's College London London, UK

Email: oliver.runswick@kcl.ac.uk

Declaration of Interest: This project has been carried out with the support of a Classification Research Grant from the International Paralympic Committee. OR and PA receive funding from the International Paralympic Committee and International Blind Sports Federation. PA receives research funding via a Collaborative Research Grant from the College of Optometrists.

The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/u7eb8/?view_only=16b2a5ec16d74573891034dc144d3af1

Abstract

The International Paralympic Committee requires international federations to develop and implement sport-specific classification guidelines based on scientific evidence. As a result of these requirements, new evidence-based criteria are required in football for athletes with vision impairment (VI). Performance tests are key to the development of a new classification system. Therefore, the aim of this study was to develop a valid and reliable test of technical performance for VI football. To assure content and face validity, the Vision Impaired Football Skills (VIFS) test was based on recommendations from experienced players and coaches. To test construct validity, we compared 24 sighted football players that were split into two groups based on highest-level of performance achieved but matched on experience. To test reliability all players completed the VIFS three times on two separate days. Results supported construct validity through detecting significant differences in performance times between the two groups ($p = .004$, $g = 1.28$ 95% CI = 0.41 - 2.15). The small bias between visits ($.54s \pm 2.93s$; 95% LoA = -5.21– 6.29) and intraclass correlations (.81, 95% CI = .56 - .92) showed between-day agreement and reliability. Within-day reliability was good when participants had completed a familiarisation trial. Results support the suitability for the use of the VIFS test for classification research. Future work should establish feasibility for players with a VI.

Introduction

In Paralympic sport, athletes are grouped into classes with the aim to minimise the impact of the eligible impairment on the outcome of competition (Tweedy & Vanlandewijck, 2011). In the past, this classification process has been based solely on the nature of the impairment. However, this approach does not account for how impairment can impact performance differently across sports. It is therefore now a requirement of the International Paralympic Committee (IPC; International Paralympic Committee, 2015) that classification is based on the understanding of the relationship between a specific eligible impairment and performance in a specific sport. The understanding of this performance-impairment relationship must be based on sport and impairment specific research evidence (International Paralympic Committee, 2016; Tweedy et al., 2014; Tweedy & Vanlandewijck, 2011).

Conducting research to establish the performance-impairment relationship in parasport (any sport for people who have a disability) is a multistage process (Tweedy et al., 2014). The first step is specifying the impairment types eligible for a specific sport and developing valid measures of the relevant impairment. There is then a need for a model of the determinants of performance in the specific sport and for valid and reliable methods to test those. Such tests do not currently exist for vision impaired (VI) football. An ability to measure the determinants of performance combined with an ability to measure the level of impairment allows for research into the performance-impairment relationship. Understanding this relationship allows for the development of the evidence-based minimum impairment criteria (MIC; Tweedy et al., 2014) and sport classes. The MIC represents the minimum level of impairment that impacts performance and dictates who is eligible to compete in the para

version of a sport. Those athletes who are eligible can be grouped into sport classes with others whose impairment have a similar impact on performance.

Several sports for athletes with physical and cognitive impairments have begun to implement evidence-based classification systems (e.g. Beckman et al., 2014; Pastor et al., 2019; Reina et al., 2018; Vanlandewijck et al., 2011). However, in the majority of VI sports, athletes are still classified in the same fashion. This system was originally based on World Health Organisations definition of blindness. Individuals are categorized as B1, B2 or B3; where B1 athletes are effectively blind with some individuals able to perceive light (LogMAR Visual Acuity worse than 2.6), B2 (LogMAR visual acuity 1.5 – 2.6; visual field radius < 5 degrees) and B3 (LogMAR visual acuity 1.0 – 1.4; visual field radius < 20 degrees) athletes have progressively more vision (Mann & Ravensbergen, 2018).

The current system for athletes with VI does not satisfy the aim of classification in Paralympic sport and there is a need for research to develop an understanding of performance and impairment in a range of VI sports (Ravensbergen et al., 2016). The IPC and the International Blind Sports Federation outlined models for conducting research to establish the performance-impairment relationship in VI sports with the aim of developing evidence-based classification systems (including minimum impairment criteria and sport classes; Mann & Ravensbergen, 2018). Research has begun working towards sport-specific models in VI swimming (Ravensbergen et al., 2018), VI judo (Krabben et al., 2018, 2019), VI skiing (Stalin, 2020), and, following a significant body of work, new systems have now been implemented in VI shooting (Allen et al., 2016, 2018, 2019; Latham, 2020; Myint et al., 2016). However, evidenced-based classification is yet to be developed for the VI version of the world's most popular sport, football.

Measuring individual performance, while straightforward in sports where outcomes such as race times are readily available, is a more complex endeavour in a team sport such as football. VI football is an adapted version of futsal. Futsal is itself a 5-a-side adaptation of football, played on a court smaller than a standard football pitch using a smaller and harder ball which bounces less. These adaptations make it more suitable for VI players than its 11-a-side counterpart. In VI football B1 and B2/B3 athletes compete in separate versions of the game. B1 athletes compete with sound in the ball, wearing blindfolds, and with guides and coaches able to vocally guide players whereas B2/B3 athletes compete with use of their vision. Both versions incorporate sighted goalkeepers. Uniquely in VI sport, currently only B1 athletes compete in the Paralympic games. According to the International Blind Sports Federation and IPC guidelines it is crucial to establish the MIC in the unadapted form of the sport. Therefore, valid and reliable measures of the aspects of futsal performance that could be affected by vision impairment are required (Mann & Ravensbergen, 2018).

In a study into the needs of a classification system for VI football, Runswick et al., (2020) adopted the Delphi process to establish consensus amongst experts in VI football on the needs of a sport-specific classification system. The expert panel identified the aspects of performance that would most likely be negatively affected by the presence of a vision impairment and how important these factors are in winning games. These findings offered clear guidance that valid measures of performance should target technical and perceptual-cognitive skills (e.g. anticipation and decision-making). Despite the wealth of football-specific sport science literature, and the existence of comparable tests in basketball (Conte et al., 2019) and netball (Mungovan, 2018), there is almost a complete paucity of representative and valid assessments that incorporate multiple aspects of technical skill and can be undertaken by an individual player.

The efforts made to date in developing measures of technical skill in football (for review see Ali, 2011) have specific issues that make them unfit for use in classification research in VI football. For example, tests often focus on single or limited skills such as passing (e.g. Loughborough Soccer Passing Test, Ali et al., 2007), shooting (e.g. Loughborough Soccer Shooting Test, Ali et al., 2007) or heading and dribbling (Rösch et al., 2000). Others test aspects of football that are not representative of game play, such as ball juggling (Rösch et al., 2000) or wall volleys (Vanderford et al., 2004), or lack control due to reliance on full match play (Rampinini et al., 2007).

One exception does exist, the Futsal Special Performance Test (Farhani et al., 2019). The test incorporates multiple skills and does have potential benefits for the development of classification in VI football. However, while the Futsal Special Performance Test was shown to be valid and reliable, reliability was only assessed using correlations across two trials conducted on the same day, an issue that needs to be addressed in order to conduct classification studies in VI football. Furthermore, a key requirement is that the test is accessible to players with (e.g. for the development of classes) and without (e.g. for simulation studies) VI and can be practically administered at a variety of locations for individual players. The Futsal Special Performance Test is not currently accessible to VI players due to its use of cones for dribbling and it also lacks practical usability due to its involvement of four skilled and sighted players to act as passers in addition to the player being tested.

Therefore, a test is required that incorporates the elements identified by the expert panel in Runswick et al's (2020) Delphi study, is accessible to VI athletes, and is practical to use in a variety of testing locations. The test also needs to display face (logical) validity, content validity, and construct validity. Face validity refers to the degree that a test would be subjectively viewed as measuring technical performance, and content validity the degree to

which a test measures all facets of technical performance (Currell & Jeukendrup, 2008). These are hard to measure objectively but can be established by working with input from individuals who are experienced in the sport in question. Construct validity refers to the degree to which a test can measure a hypothetical construct. In this case technical performance is considered such a construct. Therefore, construct validity can be established objectively through the comparison of two groups who perform at different levels. A test with good construct validity should easily discriminate between the skill-level groups (Currell & Jeukendrup, 2008; Thomas & Nelson, 2001). This approach has been adopted to validate performance tests in futsal (Farhani et al., 2019) and basketball (Conte et al., 2019), and has been widely used in research applying the expert performance approach in sport (Williams & Ericsson, 2005).

The aim of this study was to develop a valid and reliable futsal specific test of technical performance that can be used in research to develop understanding of the performance-impairment relationship in VI football. To meet this aim and to ensure face and content validity we based the design of the test on the expert opinions gleaned in Runswick et al's (2020) Delphi study and further consultation with subject matter experts. We then tested the construct validity of the test by comparing two groups who had performed at different levels but were matched on experience (Currell & Jeukendrup, 2008; Thomas & Nelson, 2001), and tested within- and between-day reliability through the completion of the test a total of six times across two days.

Method

Test Development

To ensure both content and face validity, the Vision Impaired Football Skills Test (VIFS Test) was developed in partnership with experienced coaches, players, and sport scientists.

The current test included all the technical skills that the expert panel in Runswick et al's (2020) Delphi study agreed would be negatively affected by VI. These included ball control, dribbling, passing, spatial awareness and movement around the court. The Delphi study also identified that anticipation and decision-making are likely to be negatively affected by VI. While other screen-based tests that focus on these aspects already exist (e.g. Roca et al., 2014), it was decided to include elements of anticipation and decision-making in this test through the use of a defender and the inclusion of choice on how best to complete the course as fast as possible based on the individual strengths of each participant.

An experienced B1 coach who was not part of the previously mentioned Delphi study and a previous head of sports science for an international football team were consulted with to adapt the Futsal Special Performance Test (Farhani et al. 2019). Key changes included the removal of cones, the removal of shooting that did not reach agreement in Runswick et al (2020), and to enhance practical usability, the four skilled players were replaced with use of common gym benches based on the Loughborough Soccer Passing Test (LSPT; Ali et al., 2007; Supplement Figure 1). Time penalties from the Loughborough Soccer Passing Test were also included in the design to maintain a game realistic speed and accuracy trade-off while producing a single time-based performance score. Finally, a defender was added to incorporate an element of anticipation and decision making and to enhance content and face validity. A fully detailed step by step overview of the test set-up, course (Supplement Figure 2), and penalties is available in the supplementary material.

Participants

Having consulted with players and coaches from VI football during test development, sighted players were then required for the remainder of this study. Without an understanding of the performance-impairment relationship in this sport, the use of sighted players was the only

way to control for effects of vision. Therefore, 24 football or futsal players (age 21.16 ± 6.57 years) with normal or corrected to normal vision volunteered to participate in the study. The *competitive* group ($n=12$) consisted of players who had played in UK tier 11 (county leagues) or above and had 12.9 ± 5.1 years competitive experience and included one female professional player. The *social* group ($n = 12$) had equivalent experience (14.4 ± 7.3 years; $t = .56$, $p = .58$) but had only played below UK tier 11 (e.g. local Sunday leagues) and included one female club level player. Participants signed written informed consent, ethical approval was granted from the University of Chichester ethics committee (1920_07) and all experimental procedure conformed to the ethical standards of the Declaration of Helsinki.

Procedure

Participants completed the VIFS test a total of six times across two visits (three trials per visit). Participants were instructed to wear suitable clothing and footwear to participate in futsal on an indoor court. Upon arrival during visit 1, participants read information sheets and were offered the opportunity to ask questions before signing informed consent forms. Following this, participants completed a standardised warmup which consisted of jogging, change of direction and change of speed drills. Participants were also offered the opportunity to complete any other exercises they would normally complete to be ready for play. Following the warmup, participants were immediately walked through the course by the tester and completed each action at a walking pace (full details on the course are available in the supplementary material). After the walk-through, participants were asked to verbally talk the tester through the course to ensure familiarity with all action points and rules for penalties and rewards. The participant then completed the course in the fastest possible total time three times with a two-minute rest period after each trial. An identical procedure was executed on a second visit a minimum of 48 hours and maximum of one week later.

Dependent Measures

Raw Time. The raw time for each trial was calculated as the time from the ball touching the ground at the start to the ball touching the final bench, the defender's body, or the white line as it left the court (whichever occurred soonest; see supplement). Practicality was a primary aim for this test, so a manual time to the nearest second (stopwatch) was chosen and accuracy confirmed using video footage.

Penalty Time. The penalty time for each trial was calculated as the sum of all penalties and deductions incurred on that trial.

Total Time. The total time for each trial was calculated as the sum of raw time and penalty time.

Data Analysis

Total Time (the combination of raw and penalty time) was used for further analysis. To investigate a variety of possible methods for implementing the VIFS test to best capture optimal performance and consistency, we used individual trial times to calculate overall visit times for each of the dependent measures in three different ways.

- (i) *All Trials* – the average of all trials from a visit
- (ii) *Fastest Two* – the average time for the two fastest trials from a visit
- (iii) *Fastest* – the single fastest time for a visit.

To ensure both visits were always accounted for in group comparisons, we also calculated an *All Trials Mean* that consisted of all six trials (three from each of the two visits), a *Fastest Two Mean* that consisted of four trials (the fastest two from each of the two visits), and a *Fastest Mean* that consisted of two trials (the fastest from each of the two visits). Independent

sample t-tests were used to compare competitive and social players across the three measures of total time (All Trials, Fastest Two, and Fastest).

A three-way mixed ANOVA was conducted to detect the effect of group, trial, and visit on the Total Time for All Trials and assess the needs for familiarisation. A Bonferroni adjustment was employed when multiple comparisons were being made to avoid Type I errors (McLaughlin & Sainani, 2014). Violations of sphericity were corrected for by adjusting the degrees of freedom using the Greenhouse Geisser correction when epsilon was less than 0.75 and the Huynh-Feldt correction when greater than 0.75 (Girden, 1992).

To assess within-day and between-day relationships, Pearson's correlation and two-way random effects, absolute agreement, single measurement intra class correlation coefficients (ICC) were used between all trials and each of the three measures for visit time (All Trials, Fastest Two and Fastest). To assess agreement Bland-Altman analyses were conducted. Bland-Altman (or difference plots) are a graphical method for comparing two different measurements and evaluating agreement through calculating and displaying both individual data points and a bias (average discrepancy between two trials or visits). Limits of agreement (LoA) display a range within which 95% of repeated measures would lie when compared to the first measurement.

Effect sizes were calculated for all analyses; to account for sample size Hedges' g effect sizes were used for group comparisons, partial eta squared (η_p^2) for ANOVA analyses, correlation coefficient (r) for Pearson's correlations and ICCs were accompanied by a 95% confidence interval. The alpha level (p) for statistical significance was set at .05.

Results

All data for this paper including raw, penalty, and total time are available via the Open Science Framework ([Link Here](#)). Results presented below focus on Total Time as a combination of Raw Time and Penalty Time.

Group Differences

The competitive group displayed significantly faster times across all calculations of visit time. This included All Trials Mean (competitive = 28.40 ± 2.66 s; social = 31.94 ± 3.36 s; $t = 2.87$, $p = .009$, $g = 1.13$, 95% CI = 0.28 - 1.98; Figure 1), Fastest Two Mean (competitive = 27.06 ± 2.61 s; social = 30.58 ± 3.60 ; $t = 2.74$, $p = .012$, $g = 1.08$ 95% CI = 0.23 - 1.93), and Fastest Mean (competitive = 24.92 ± 2.52 s; social = 28.63 ± 3.05 ; $t = 3.24$, $p = .004$, $g = 1.28$ 95% CI = 0.41 - 2.15).

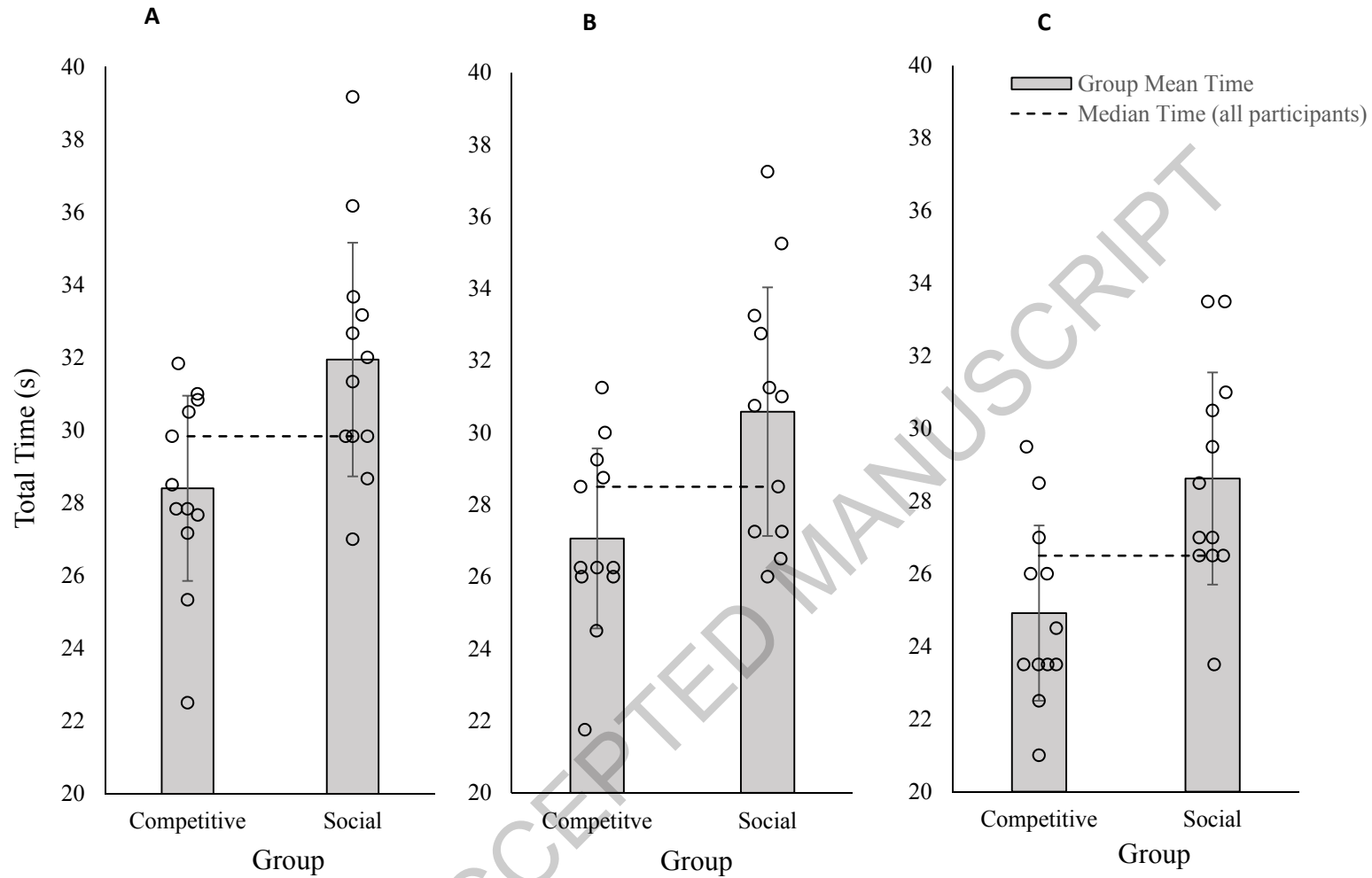


Figure 1. Total Time for individuals and group across (A) All Trials Mean (B) Fastest Two Mean and (C) Fastest Mean.

Familiarisation

There was a significant main effect of group on Total Time (competitive = 28.40 ± 2.66 s; social = 31.94 ± 3.36 s; $F = 8.20$, $p = .009$, $\eta_p^2 = .27$). There was no effect of visit on Total Time ($F = 2.672$, $p = .12$, $\eta_p^2 = .11$) and no Visit by Group interaction ($F = 3.33$, $p = .08$, $\eta_p^2 = .13$). There was a main effect of Trial on Total Time ($F = 8.13$, $p = .001$, $\eta_p^2 = .27$; Figure 2). Post hoc analysis revealed that participant's first trial was significantly slower than their second ($p = .04$) and third ($p = .004$) trials during each visit. There was no difference between Trial 2 and Trial 3 ($p = .32$). There was a significant Trial by Group interaction ($F = 4.93$, $p = .01$, $\eta_p^2 = .18$) suggesting that the effect of trial was greater in the competitive group, and a Trial by Visit interaction ($F = 3.32$, $p = .05$, $\eta_p^2 = .13$) suggesting the effect of trial was slightly greater in Visit 1 than Visit 2. There was no three-way interaction between Group, Trial and Visit ($F = .65$, $p = .53$, $\eta_p^2 = .03$).

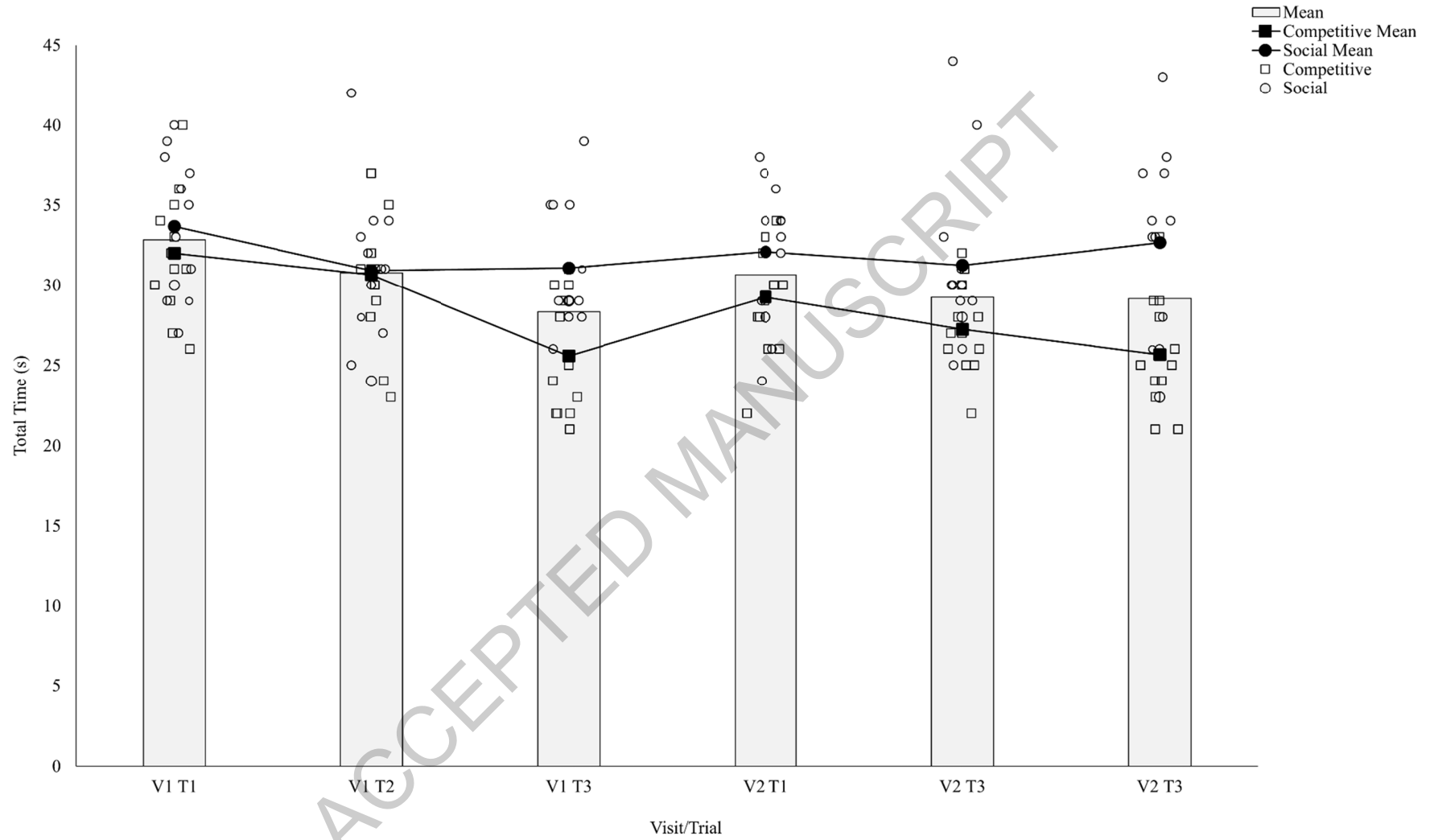
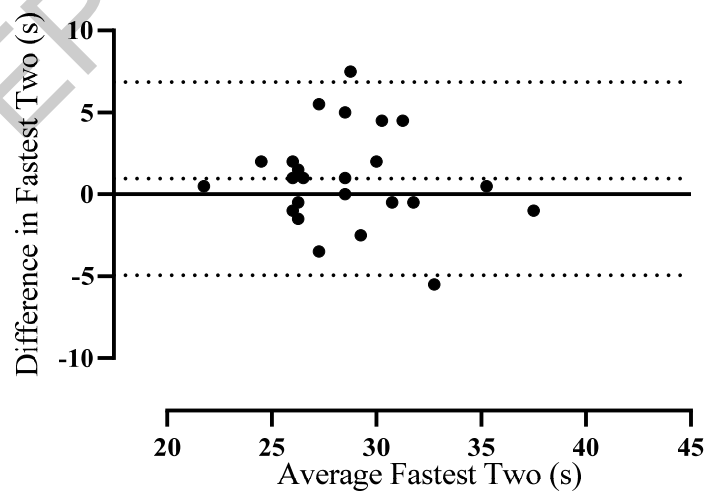
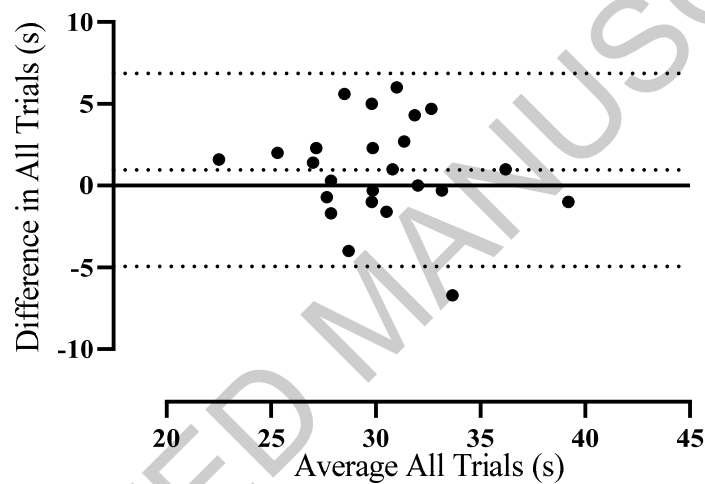


Figure 2. Mean and Individual Total Times for all participants across the three trials in each of the two visits.

Reliability and Agreement

Between-Visit

Strong relationships (Evans, 1996) and ICCs were found between visits for All Trials ($r = .69, p < .001, ICC = .80, 95\% CI = .55 - .91$), Fastest Two ($r = .69, p < .001, ICC = .81, 95\% CI = .57 - .92$) and Fastest ($r = .68, p < .001, ICC = .81, 95\% CI = .56 - .92$). Figure 3 shows Bland-Altman plots for All Trials (Bias \pm SD, 95% Limits of Agreement; $0.95 \pm 3.01, -4.95 - 6.86$), Fastest Two ($0.92 \pm 2.97, -4.91 - 6.74$) and Fastest ($0.54 \pm 2.93, -5.21 - 6.29$) time.



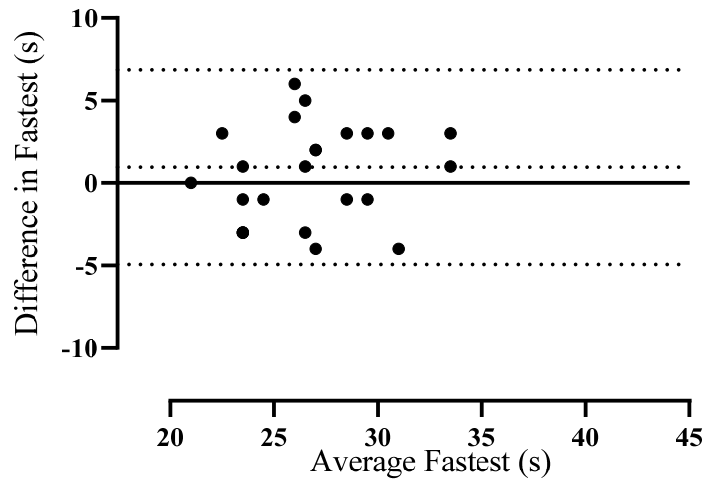


Figure 3. Bland-Altman plots showing Bias and 95% limits of agreement for the three possible measures of time for each visit for All Trials, Fastest Two and Fastest.

Between Trial

Table 1 displays the data for relationships, reliability, and agreement between individual trials. In support of the familiarisation findings, comparisons that include a trial one from either visit show the weakest relationships and largest bias.

Table 1. Trial-Trial reliability for Total Time (s). Pearson's (r) correlations, 95% confidence intervals (CI), and significance values (p). Bland-Altman analysis shows Bias \pm SD and 95% limits of agreement (LoA) between individual trials. Significant relationships displayed in bold.

		Visit 1			Visit 2		
		Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	
Visit 1	Trial 2	r (95% CI)	.54 (.15 - .84)				
		P	.003				
		Bias \pm SD	2.04 \pm 4.20				
		95% LoA	-6.18 – 10.27				
	Trial 3	r (95% CI)	.49 (.09 - .78)	.37 (-.04 - .76)			
		p	.008	.04			
		Bias \pm SD	4.50 \pm 4.51	2.46 \pm 5.16			
		95% LoA	-4.34 – 13.34	-7.65 – 12.57			
	Trial 1	r (95% CI)	.33 (-.08 - .75)	.42 (.02 – .89)	.30 (-.13 - .82)		
p		.05	.02	.08			
Bias \pm SD		2.17 \pm 4.81	0.13 \pm 4.70	-2.33 \pm 5.25			
	95% LoA	-7.27 – 11.60	-9.09 – 9.34	-12.62 – 7.95			
Visit 2	Trial 2	r (95% CI)	.31 (-.10 - .64)	.53 (.15 - .86)	.41 (.01 - .81)	.63 (.25 - .86)	
		p	.07	.004	.02	< .001	
		Bias \pm SD	3.58 \pm 5.26	1.54 \pm 4.48	-0.92 \pm 5.12	1.42 \pm 3.87	
		95% LoA	-6.72 – 13.89	-7.24 – 10.33	-10.95 – 9.11	-6.16 – 8.96	
	Trial 3	r (95% CI)	.37 (-.03 - .55)	.50 (.09 - .67)	.68 (.28 - .79)	.26 (-.12 - .48)	.52 (.12 - .72)
		p	.04	.01	< .001	.11	.04
		Bias \pm SD	3.67 \pm 5.86	1.625 \pm 5.36	-.83 \pm 4.44	1.50 \pm 6.30	.08 \pm 5.33
		95% LoA	-7.81 – 15.15	-8.87 – 12.12	-9.54 -7.87	-10.86 – 13.86	-10.37 – 10.53

Discussion

The aim of the study was to produce a valid and reliable test of technical performance that can be used for research into the performance-impairment relationship in VI football. Content and face validity were ensured by developing the test in partnership with recommendations from experts in football and VI football (Runswick et al., 2020). Construct validity was established by comparing competitive and social players and reliability through within- and between day relationships.

Simulation studies used to establish the MIC use repeated measures across several different levels of impairment. Similarly, studies to establish classes involve testing large samples of players with a variety of impairments. Therefore, it was also important to establish the reliability of the VIFS Test across visits, an aspect that was lacking in the reliability testing for the Futsal Special Performance Test (Farhani et al., 2019). Results from between day (visit) analysis showed low bias and Pearson's and intraclass correlations suggested between-day reliability. Reliability was displayed across all three measures of on-the-day performance with the *All Trials* and *Fastest Trial* showing only slightly stronger relationship than the *Fastest Two*. While any of these measures could be used for developing a single performance score for a visit or level of impairment in a simulation study, the fastest time represents the simplest method with the least bias. However, this does not account for the need to familiarise participants on the day.

Within day (trial) analysis displayed a variety of weak to strong correlations and some larger bias, with the majority occurring when a Trial 1 from either visit was used. Results from the repeated measures ANOVA suggested that while there were no significant differences between visits, there was a main effect of trial. Post hoc testing showed that the first trial of a visit was significantly slower than the subsequent trials. This, in combination with poor

within-day reliability for Trial 1s, suggests that a familiarisation trial should be used as the first trial of each visit. Furthermore, wide LoAs between individual trials suggest the test is best used with multiple trials to create a single visit time.

Group comparisons showed significantly faster Total Time in the competitive group with large effect sizes for the All Trials Mean, Fastest Two Mean and Fastest Mean measures of individual performance. The largest effect for between group difference was found when taking the fastest time from each visit. Effect sizes displayed for group differences by VIFS Test are smaller than those displayed by the Futsal Special Performance Test (Farhani et al., 2019). However, here we deliberately matched groups on experience and closely matched levels of performance for the groups. Therefore, a large effect size for group differences suggests supports the construct validity of the test (i.e. the ability to distinguish time between the groups).

Previous football skill tests have lacked strong face, content, and construct validity, often due to focusing on single aspects of performance or using unrepresentative tasks (Ali, 2011). However, the VIFS Test represents an important development in these areas and will allow for the development of simulation studies to investigate the MIC and mass testing of athletes with impairment to develop evidence for classes. This evidence will form the foundation of future work to provide an evidence-based classification system for VI football (Mann & Ravensbergen, 2018; Tweedy et al., 2014; Tweedy & Vanlandewijck, 2011). Simulation studies involve systematically impairing the vision of sighted players (e.g. Allen et al., 2016) and testing players with impairment to build evidence for sport classes, require clear differentiation between levels of performance. It is therefore important that the VIFS Test has been able to reliably detect differences in technical proficiency while testing an individual's performance in multiple aspects of the game likely to be affected by VI (Runswick et al., 2020).

A key aim of this work was to produce a practically useful test that can be applied by anyone whilst minimising the need for additional equipment. The test has proved valid and reliable based on manual timing to the nearest second on a stopwatch and without any specialist equipment requirements aside from a futsal and four gym benches. In addition, only two further personnel who need not have any futsal or football skill are required to run the test. Based on the findings from the present study we would suggest that, when utilising the test with sighted individuals (i.e. in simulation studies), the test is implemented with at least one familiarisation trial at full speed after the walk through from the tester and multiple trials are conducted to generate a single visit time.

Results of this work and applications of the VIFS test should be considered alongside limitations. According to International Blind Sports Federation and IPC guidelines the research to establish the MIC should be conducted using the unadapted form of the sport using simulation studies with sighted players (Mann & Ravensbergen, 2018). This is to allow for systematic intrasubject assessment of performance at various levels of VI from fully sighted to severely impaired. However, testing with VI players will be required to develop classes. In the current study, expert consultation was sought in an endeavour to make this test accessible for when it is needed for VI athletes or sighted athletes with simulated impairment and consequently the authors are confident that it is accessible to the levels of impairment required to test the MIC. Future work should aim to assess feasibility of the test with athletes who have an impairment.

This study displayed reliability over two visits when familiarisation was included. However, simulation work for MIC testing and work on developing classes may require more visits over a longer period or more trials within a visit. While we have endeavoured to enhance content validity and include some aspects of decision-making, the test is not an unpredictable game scenario and the participants will become familiar with the course. Future work that

implements the test should be careful to include aspects in designs that check for learning effects over the course of data collection.

In summary, this study has progressed the development of an evidence-based MIC through the development of a test assessing technical skill in VI football. The study aimed to develop a test that incorporated all the elements of performance identified as important in classification research (Runswick et al., 2020), is practical to deliver in a variety of settings, and offers accessibility for players with a VI or simulated VI. We have presented the VIFS Test as a valid and reliable method to achieve this goal and offered practical guidance on its implementation. The test can now be used for research to develop understanding of the performance-impairment relationship in VI football and evidence-based classification.

References

- Ali, A. (2011). Measuring soccer skill performance: a review. *Scandinavian Journal of Medicine & Science in Sports*, 21(2), 170–183. <https://doi.org/10.1111/j.1600-0838.2010.01256.x>
- Ali, Ajmol, Williams, C., Hulse, M., Strudwick, A., Reddin, J., Howarth, L., Eldred, J., Hirst, M., & McGregor, S. (2007). Reliability and validity of two tests of soccer skill. *Journal of Sports Sciences*, 25(13), 1461–1470. <https://doi.org/10.1080/02640410601150470>
- Allen, P. M., Latham, K., Mann, D. L., Ravensbergen, R. H. J. C., & Myint, J. (2016). The level of vision necessary for competitive performance in rifle shooting: Setting the standards for paralympic shooting with vision impairment. *Frontiers in Psychology*, 7(NOV). <https://doi.org/10.3389/fpsyg.2016.01731>
- Allen, P. M., Latham, K., Ravensbergen, R. H. J. C., Myint, J., & Mann, D. L. (2019). Rifle Shooting for Athletes With Vision Impairment: Does One Class Fit All? *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01727>

- Allen, P. M., Ravensbergen, R. H. J. C., Latham, K., Rose, A., Myint, J., & Mann, D. L. (2018). Contrast sensitivity is a significant predictor of performance in rifle shooting for athletes with vision impairment. *Frontiers in Psychology, 9*(JUN). <https://doi.org/10.3389/fpsyg.2018.00950>
- Beckman, E. M., Newcombe, P., Vanlandewijck, Y., Connick, M. J., & Tweedy, S. M. (2014). Novel strength test battery to permit evidence-based paralympic classification. *Medicine (United States), 93*(4). <https://doi.org/10.1097/MD.0000000000000031>
- Bland J. M., & Altman D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*(2):135-160. [doi:10.1177/096228029900800204](https://doi.org/10.1177/096228029900800204).
- Conte, D., Smith, M. R., Santolamazza, F., Favero, T. G., Tessitore, A., & Coutts. A. (2019). Reliability, usefulness and construct validity of the Combined Basketball Skill Test (CBST), *Journal of Sports Sciences, 37*(11), 1205-1211, [doi:10.1080/02640414.2018.1551046](https://doi.org/10.1080/02640414.2018.1551046)
- Currell, K., Jeukendrup, A.E. (2008). Validity, Reliability and Sensitivity of Measures of Sporting Performance. *Sports Medicine, 38*, 297–316. <https://doi.org/10.2165/00007256-200838040-00003>
- Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. In *Straightforward statistics for the behavioral sciences*. Thomson Brooks/Cole Publishing Co.
- Farhani, F., Rajabi, H., Negaresh, R., Ali, A., Amani Shalamzari, S., & Baker, J. S. (2019). Reliability and Validity of a Novel Futsal Special Performance Test Designed to Measure Skills and Anaerobic Performance. *International Journal of Sports Physiology and Performance, 14*(8), 1096–1102. <https://doi.org/10.1123/ijsp.2018-0850>

Girden, E. (1992). *ANOVA*. <https://doi.org/10.4135/9781412983419>

International Paralympic Committee. (2015). *IPC Athlete Classification Code: Rules, Policies and Procedures for Athlete Classification*.

International Paralympic Committee. (2016). *International standard for athlete evaluation*.

Krabben, K. J., Ravensbergen, R. H. J. C., Nakamoto, H., & Mann, D. L. (2019). The development of evidence-based classification of vision impairment in Judo: A Delphi study. *Frontiers in Psychology, 10*(FEB). <https://doi.org/10.3389/fpsyg.2019.00098>

Krabben, K. J., van der Kamp, J., & Mann, D. L. (2018). Fight without sight: The contribution of vision to judo performance. *Psychology of Sport and Exercise, 37*, 157–163. <https://doi.org/10.1016/j.psychsport.2017.08.004>

Latham, L., Mann, D.L., Dolan, R., Myint, J., Timmis, M., Ryu, D, Frisson, S., Allen, P.M. (2020) Do visual fields need to be considered in eligibility criteria within visually impaired shooting? *Journal of Sport Sciences* (Currently under review)

Mann, D. L., & Ravensbergen, H. J. C. (2018). International Paralympic Committee (IPC) and International Blind Sports Federation (IBSA) Joint Position Stand on the Sport-Specific Classification of Athletes with Vision Impairment. *Sports Medicine, 48*(9), 2011–2023. <https://doi.org/10.1007/s40279-018-0949-6>

McLaughlin, M. J., & Sainani, K. L. (2014). Bonferroni, holm, and hochberg corrections: Fun names, serious changes to P values. *PM and R*. <https://doi.org/10.1016/j.pmrj.2014.04.006>

Mungovan, S. F., Peralta, P. J., Gass, G. C., & Scanlan, A. T. (2018). The test-retest reliability and criterion validity of a high-intensity, netball-specific circuit test: The Net-Test. *Journal of Science and Medicine in Sport, 21*(12), 1268-1273.

- Myint, J., Latham, K., Mann, D., Gomersall, P., Wilkins, A. J., & Allen, P. M. (2016). The relationship between visual function and performance in rifle shooting for athletes with vision impairment. *BMJ Open Sport & Exercise Medicine*, 2(1), e000080.
<https://doi.org/10.1136/bmjsem-2015-000080>
- Pastor, D., Campayo-Piernas, M., Pastor, J. T., & Reina, R. (2019). A mathematical model for decision-making in the classification of para-footballers with different severity of coordination impairments. *Journal of Sports Sciences*, 37(12), 1403–1410.
<https://doi.org/10.1080/02640414.2018.1560617>
- Rampinini, E., Bishop, D., Marcora, S. M., Ferrari Bravo, D., Sassi, R., & Impellizzeri, F. M. (2007). Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *International Journal of Sports Medicine*, 28(3), 228–235. <https://doi.org/10.1055/s-2006-924340>
- Ravensbergen, H. J. C., Genee, A. D., & Mann, D. L. (2018). Expert consensus to guide the classification of Paralympic Swimmers with vision impairment: A Delphi study. *Frontiers in Psychology*, 9(OCT). <https://doi.org/10.3389/fpsyg.2018.01756>
- Ravensbergen, H. J. C., Mann, D. L., & Kamper, S. J. (2016). Expert consensus statement to guide the evidence-based classification of Paralympic athletes with vision impairment: A Delphi study. *British Journal of Sports Medicine*, 50(7), 386–391.
<https://doi.org/10.1136/bjsports-2015-095434>
- Reina, R., Iturricastillo, A., Sabido, R., Campayo-Piernas, M., & Yanci, J. (2018). Vertical and horizontal jump capacity in international cerebral palsy football players. *International Journal of Sports Physiology and Performance*, 13(5), 597–603.
<https://doi.org/10.1123/ijsp.2017-0321>
- Roca, A., Williams, A. M., & Ford, P. R. (2014). Capturing and testing perceptual-cognitive

- expertise: A comparison of stationary and movement response methods. *Behavior Research Methods*, 46(1), 173–177. <https://doi.org/10.3758/s13428-013-0359-5>
- Rösch, D., Hodgson, R., Peterson, L., Graf-Baumann, T., Junge, A., Chomiak, J., & Dvorak, J. (2000). Assessment and evaluation of football performance. *American Journal of Sports Medicine*, 28(5 SUPPL.). https://doi.org/10.1177/28.suppl_5.s-29
- Runswick, O. R., Ravensbergen, H. J. C., Allen, P. M., & Mann, D. L. (Minor Revisions Submitted). Expert opinion on classification for footballers with vision impairment: Toward evidence-based minimum impairment criteria. *Journal of Sports Sciences*.
- Stalin, A. (2020). Development of sports-specific classification for Paralympic skiers with visual impairment. *Unpublished Doctoral Thesis. UWSpace*.
<http://hdl.handle.net/10012/15933>
- Thomas JR, Nelson JK. (2001). *Research methods in physical activity*. Champaign (IL): Human Kinetics.
- Tweedy, S. M., Beckman, E. M., & Connick, M. J. (2014). Paralympic Classification: Conceptual Basis, Current Methods, and Research Update. *PM and R*, 6(8 SUPPL.).
<https://doi.org/10.1016/j.pmrj.2014.04.013>
- Tweedy, S. M., & Vanlandewijck, Y. C. (2011). International Paralympic Committee position stand-background and scientific principles of classification in Paralympic sport. In *British Journal of Sports Medicine* (Vol. 45, Issue 4, pp. 259–269).
<https://doi.org/10.1136/bjism.2009.065060>
- Vanderford, M. L., Meyers, M. C., Skelly, W. A., Stewart, C. C., & Hamilton, K. L. (2004). Physiological and sport-specific skill response of Olympic youth soccer athletes. *Journal of Strength and Conditioning Research*, 18(2), 334–342.

<https://doi.org/10.1519/R-11922.1>

Vanlandewijck, Y. C., Verellen, J., Beckman, E., Connick, M., & Tweedy, S. M. (2011).

Trunk strength effect on track wheelchair start: Implications for classification.

Medicine and Science in Sports and Exercise, 43(12), 2344–

2351. <https://doi.org/10.1249/MSS.0b013e318223af14>

Williams AM, Ericsson KA. (2005). Perceptual-cognitive expertise in sport: some considerations when applying the expert performance approach. *Human Movement Science*, 24(3), 283-307 <https://doi.org/10.1016/j.humov.2005.06.002>. PMID: 16095739.

ACCEPTED MANUSCRIPT

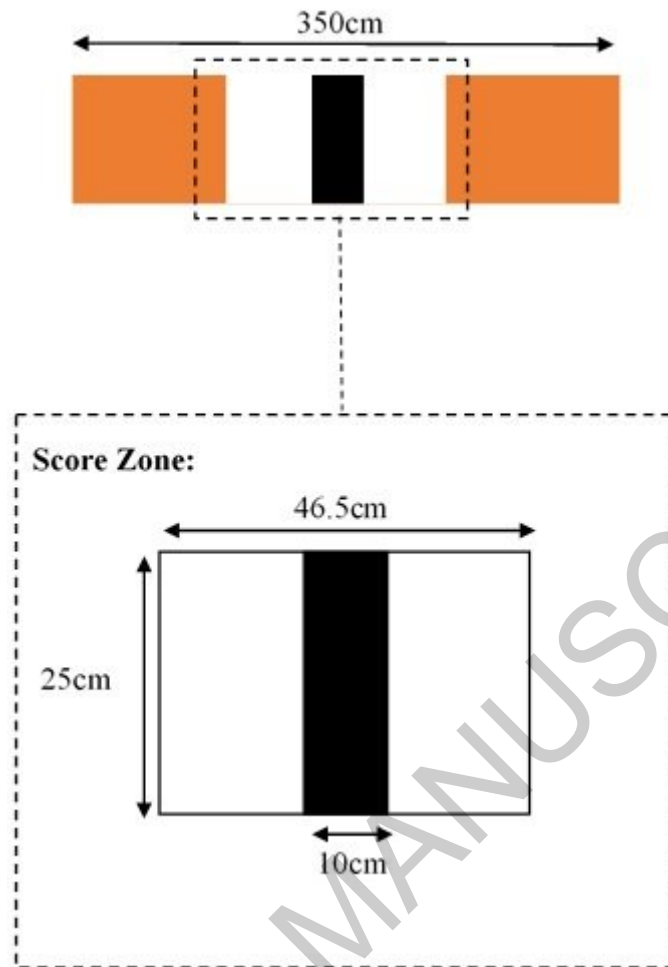
Supplementary Material

The Vision Impaired Football Skills Test (VIFS Test)

Set-Up

Standard gymnasium benches were placed so that the flat face lay perpendicular to the floor and the futsal could be bounced off. Benches included targets adapted from the Loughborough Soccer Passing Test (Ali et al., 2007) to consist of a laminated white card ($46.5 \times 25.0\text{cm}$) taped to the middle of each bench and a black mark ($10.0 \times 25.0\text{cm}$) printed vertically in the middle of the target areas to maximise contrast (Supplement Figure 1). Passing penalties and rewards matched those used in the Loughborough Soccer Passing Test (Ali et al., 2007) and were applied to participants' final time as follows. Striking the black section of a bench was rewarded with a 1 second deduction from the overall time, striking the white section incurred no penalties or rewards, and striking any other section of the bench resulted in 3 seconds being added to the final time. Missing the correct bench with the final required pass resulted in a 5 second penalty, whilst missing any other action point or bench voided the trial.

Action points and ball trap lines shown in Supplement Figure 2 were marked on the blue floor with white masking tape. A smart phone video camera was carried by the defender (shown as the green mark in Supplement Figure 2) and was used to record each trial in its entirety. The tester gave verbal instructions (specified below) to create a more accurate representation of a game situations and to aid future use for those with vision impairments.

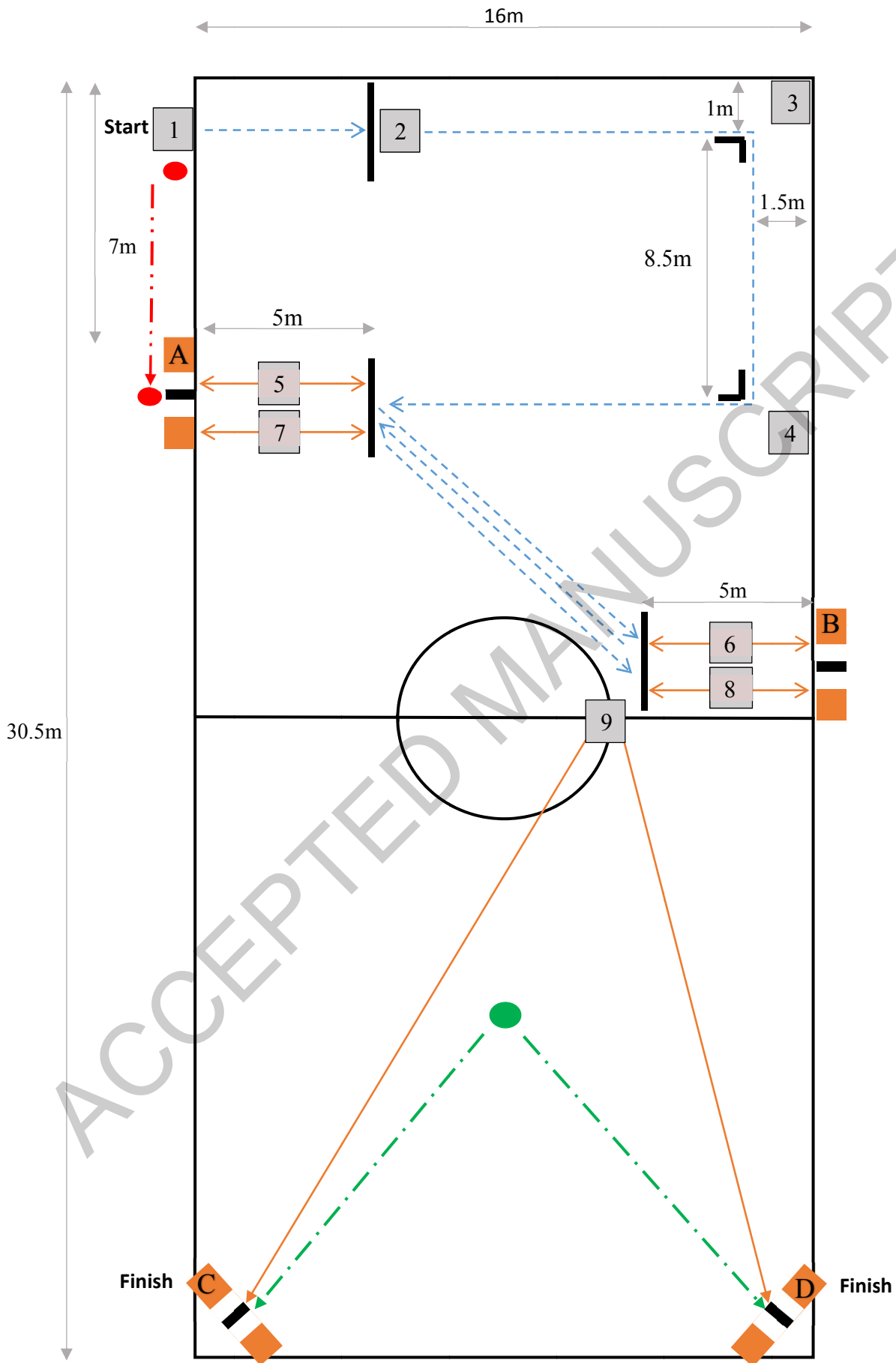


Supplement Figure 1. Bench design adapted from Loughborough Soccer Passing Test (Ali et al., 2007)

The Course

Supplement Figure 2 shows a schematic of the set up and action points on the VIFS Test. Participants started each test at action point one, facing onto the court. The start was indicated by a call of “3, 2, 1, go!” at which point the tester dropped a futsal from the participants waist height so that it hit the floor no more than 50cm from the participant’s feet. The stopwatch started at the moment the futsal hit the floor. The participant was required to control the ball and begin dribbling it to action point two where the tester called ‘trap’ and the participant was required to trap the ball. The trap was required to prevent participants from using a kick and chase technique and to ensure the need to keep control of the ball (to replace the FSPT

cones). Once the ball was at a complete stop, they resumed dribbling to action point three where the tester called 'turn' and participants turned 90 degrees right, and on to action point four where again the tester called 'turn' and participants turned 90 degrees right. The participant then dribbled to action point 5 where the tester called "pass!", whilst stood directly behind the Loughborough bench and participants were required to pass the ball against bench A. The participant was able to pass the ball on the line, or at any point beyond it to incorporate an element of decision-making between time and accuracy. Once the ball struck the bench the participant controlled the ball and turned to dribble the ball to action point 8, where they were required to make a pass against bench B, in the same fashion as done at bench A. From here they turned and returned with the ball to action point 5, repeating the two passes against benches A and B. At the moment the participant touched the ball following their second pass against bench B, the defender stood at action point 9 ran either towards bench C or bench D. This decision was pre-determined to ensure a balance across visits for pass directions. This was only communicated to the defender and so participants were not aware of either the pre-determined direction or that there would be a balance. The participant again had a time vs accuracy decision to make and choose their preferred quickest route (pass and/or dribble the ball) to get the ball to the bench the defender was not covering (C or D). The stopwatch was stopped when the ball struck the bench, left the court, or was intercepted by the defender (whichever event happened first).



Supplement Figure 2. Schematic showing the set-up and execution of the Vision Impaired Football Skills Test

Key

1	Participant Action Points	
1	Ball drop by tester on marker tape, control after first bounce and begin	6 Pass to bench B after marker tape & gather rebounding ball
2	Trap ball to a complete stop after marker tape	7 Pass to A after marker tape & gather rebounding ball
3	Right turn about marker tape	8 Pass to bench B after marker tape & gather rebounding ball
4	Right turn about marker tape	9 Pass to undefended bench (C or D)
5	Pass to bench A after marker tape & gather rebounding ball	

	Dribble
	Pass to bench & gather rebound
	Pass
	Action marker tape
	Tester standing point
	Tester movement
	Defender starting point
	Defender movement
	Loughborough Bench

ACCEPTED MANUSCRIPT