Journal of Physical Education and Sport ® (JPES), 18(2), Art 109, pp. 745 - 752, 2018
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Standardization of selected control exercises for soccer technical evaluation

PAVOL PERÁČEK[1], PETER ŠAGÁT[2], TIMOTEJ VAJDÍK[1], MARTIN MIKULIČ[1]
[1] Sports Game Department, Faculty of Physical Education and Sport, Comenius University, Bratislava, SLOVAKIA
[2] Department of Physical Education, Health and Recreation, Prince Sultan University, Ryiadh, SAUDI ARABIA

Published online: June 30, 2018
(Accepted for publication May 14, 2018)
DOI:10.7752/jpes.2018.02109

Abstract:
Soccer needs the appropriate tools to perform diagnostics (for performance analysis). But some tests give us incomplete data. We have decided to examine the test battery used by DFB (Deutscher Fußball-Bund) clubs for the long-term diagnosis of the technical aspects of gaming activities of youth soccer players in Germany. We have searched for the tests that are reliable and provide us with enough information about the technical aspects of gaming skills. The object of the survey were four motoric tests: Dribbling Tests, Ball Control Test, Agility Test, and Juggling Test that we applied to the elite U16 and U17 age categories (n = 66). Based on the ICC method (interclass correlation), we found that the lowest correlation value of the four observed tests had a Juggling test r = 0.5384 (U16) and r = 0.4379 (U17). Based on the values found, we claim that this test is not appropriate or does not provide relevant information about players of elite youth soccer teams U 16 and U 17. This test is inadequate and inappropriate for judging the game skills required for game performance, also because it is isolated from the player's movement structure in the game and its outcome does not significantly affect game performance. Other observed tests - Dribbling test, Ball control test, Agility test achieved a high degree of correlation r > as 0.6 (0.97 **). The measurement error in these tests was < 0.06 ***. The measurement error in the Juggling test was at the level of 0.71403 and therefore highest in the observed tests. From a practical point of view, the trainers have a tool for evaluating the technical aspects of gaming skills that provide detailed information about this area regarding the movement patterns of the players. The information obtained can serve as a feedback to players, and provide information to coaches for the next strategy of developing a player's technique and their movement in the training process.
Key words: soccer, diagnostic, youth team, technical abilities, reliability, ICC

Introduction

Diagnosis of training is considered to be an important prerequisite for increasing the efficiency of training and sports performance in most sectors of current top sport. The problem of diagnostics, especially when using laboratory tests, is quite often lacking the specificity of the player's movement from the match, as an eccentric, muscular activity in the change of direction (Currell & Jeukendrup, 2008). Field tests are conducted in a natural environment for soccer. Their importance lies in the specificity of load or movement (Balsom, 1994). Advantages of field tests include mostly lower price, greater availability and the possibility of realization with large groups of players (Bunc, 2006). The trend of modern stress diagnostics is a combination of both ways of evaluation of the player's current performance. When selecting a test, its properties must also be taken into account. It is important that the tests used are reliable and valid (Maxwell, 1992).

Another requirement is the standardization of conditions. Standardization of the test procedure is also important. It is ensured by preserving the same method of preparation of test subjects, in particular warm-up, instructions, motivation, and encouragement in the course of the test itself. Players should be sufficiently recovered before testing to prevent test results from affecting by fatigue (Psotta, 2006). Tests should be performed at the same time of day and the players should be well hydrated before and with the same eating strategy (Bangsbo & Mohr, 2011).

Often, we also encounter a validation statement where a certain number is given, but no one is talking about what type of validity it is and what it was actually detected. We can assume that it is so-called the aggregate validity we are talking about if we do not use one method but the whole set of methods, but it is far more likely that the information is simply not complete. When considering the suitability of using a particular method, we should also look at its properties (specific type of validity, reliability, etc.). However, if we do not know what the property was verified against, we do not know anything. When we determine the properties of the

Corresponding Author: PAVOL PERÁČEK, E-mail: pavol.peracek@uniba.sk

method, we should also know the research conditions from which the author has acquired it (MacKall & Pynes, 1991).

Reliability - In general, it tells us about the "accuracy" of the test, which also reflects the size of the test or measurement error. High reliability can be demonstrated, for example, by repeating the test for the same person under almost identical conditions, to obtain virtually the same results. In fact, even in strictly standardized tests, repeated measurements differ from each other. In another sense, reliability is "the validity of the test to itself" (Měkota & Blahuš, 1983). It is a range in which the test results are still and accurate, and the representation of the overall population in the study is known to be reliable and when the study results can even be reproducible on the basis of a similar methodology, then we say that the research tool is considered reliable (Joppe , 2000). Charles (1995) states that the consistency with which subjects are being tested is very important, and the individual score may be quite the same if the determined method is the same when repeating the test at two different times. Kirk and Miller (1986) recognize three types of reliability in quantitative research, which concern in particular: the degree of repeat measurement that remains the same, the stability of measurement in time and the similarity of measurement in other time segments.

Reliability of the test is expressed by the coefficient of reliability (r), which expresses the ratio of the actual variance to the observed variance. The reliability coefficient expresses the level of equilibrium of the actual level of the given physical ability (Safrit, 1976).

When calculating reliability, other aspects are used in addition to the ICC, but they can be confusing. It is also Pearson's correlation coefficient and LOA (Bland-Atman plots). Pearson's product of the current correlation coefficient was often used in the past to calculate reliability but has one great weakness, because it does not show a systematic measurement error (Rousson, Gasser & Seifert, 2002).

Material & methods

The observed team consisted of 66 elite soccer players U16 and U17. In the U16 category, the average age was 15.25 years, the body height was 177.6 cm and the body weight was 67.75 kg. Average age values in the U17 category were 15.78 years; body height 180.7 cm and body weight 71.06 kg. The basic method was measurement - testing. We tested the players after the end of season with 12-day setback. Tests were performed on an artificial grass surface. The test battery was assembled from four tests. In particular, Agility test, Dribbling test, Ball control and Juggling test. The players performed the tests in the same order and the order of the players in the individual tests was the same, as well as the time frame of the day was the same. To capture the data in the given test we used Fitro Gates photocells (except juggling). To analyze the obtained data, we used the IBM SPSS Statistic 21 mathematical and statistical program for the Windows operating system. In addition to ICC coefficient (intracalass correlation), we also calculated SEM (standard error of the mean) and interval of reliability (Confidence interval).

Research hypothesis

The aim of the study was to determine the reliability of selected control exercises to assess the technical skills of young soccerers.

H1: We expect the Juggling test to be the lowest value of reliability
H2: We expect the other tests to achieve a value of > 0.6
H3: We expect the measurement error to be the lowest in the Juggling test

Protocoling obtained data

The distribution and interpretation of amount of the correlation coefficient of the Pearson correlation are different from the authors. According to Cohen (1988), the correlation below 0.1 is trivial, 0.1 - 0.3 is small, 0.3 - 0.5 is moderate and above 0.5 it is large. Correlation 0.7 - 0.9 is often referred to be as very large and 0.9 - 1 as almost perfect. Hand and Taylor (1987) interpret ICC results as follows: r >0.9 reliability is high, r = 0.90 - 0.80 reliability is average, r < 0.80 reliability is uncertain for physiology, and r < 0.7 is unacceptable for non-social sciences.

Results

Ball control

Table 1. Measurement results in the test Ball control.

Order of measurement	1.(U16)	2.(U16)	1.(U17)	2.(U17)
Average	9.6	9.55	9.3	9.26
Median	9.65	9.53	9.37	9.31
Max.value	10.84	10.76	10.44	10.33
Min.value	7.75	7.79	7.89	7.81
Scatter	0.48	0.46	0.48	0.46
Standard deviation	0.69	0.68	0.69	0.68

We used the ICC statistical method to calculate reliability. This method serves in particular to calculate two or more data. In the Ball control test, ICC was (U16) = 0.9887 and ICC (U17) = 0.9892, which means that the reliability is very high.

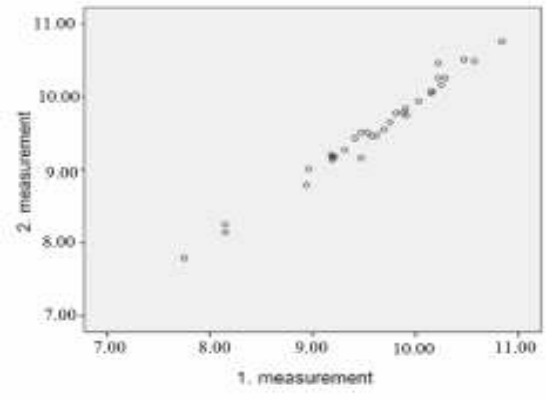

Fig.1. Correlation of U16 category in Ball control test

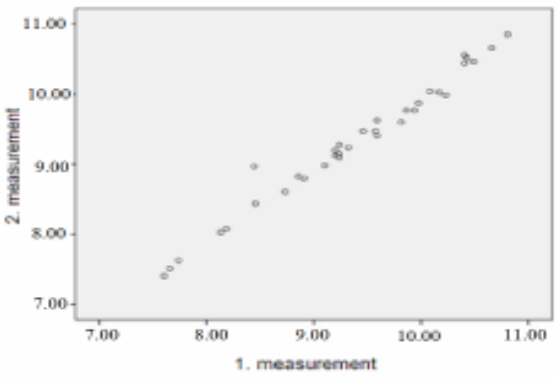

Fig. 2. Correlation of U17 in Ball control test

Juggling test

Table 2. Measurement results in the Juggling test.

Order of measurement	1(U16)	2(U16)	1(U17)	2(U17)
Average	**5.37**	**6.84**	**7**	**7.64**
Median	**6**	**8**	**8**	**8**
Max.value	**8**	**8**	**8**	**8**
Min.value	**1**	**4**	**1**	**4**
Scatter	**7.11**	**2.38**	**3.18**	**0.64**
Standard deviation	**2.7**	**1.56**	**1.80**	**0.81**

The reliability of the test stated was in category U16 ICC = 0.5384 and for U17 was ICC = 0.4379, which is unacceptable reliability for science other than social, according to Hand and Taylor (1987).

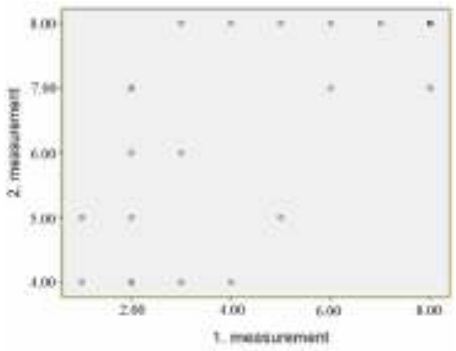

Fig. 3. Correlation of the U16 category in the Juggling test

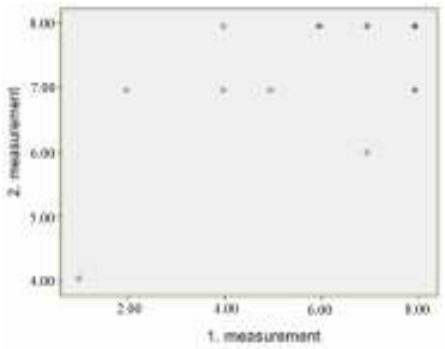

Fig. 4. Correlation of the U17 category in the Juggling test

Dribbling test

Table 3. Measurement results in the Dribbling test.

Order of measurement	1(U16)	2(U16)	1(U17)	2(U17)
Average	11.82	11.73	11.76	11.74
Median	11.81	11.75	11.76	11.71
Max.value	14.07	13.92	13.06	12.98
Min.value	10.47	10.59	11.13	11.06
Scatter	0.66	0.56	0.16	0.18
Standard deviation	0.83	0.76	0.41	0.43

Reliability in the ball guiding test reached a coefficient for the U16 category ICC = 0.976 and for the U17 category ICC = 0.9841, which is considered very high by Hand and Taylor (1987).

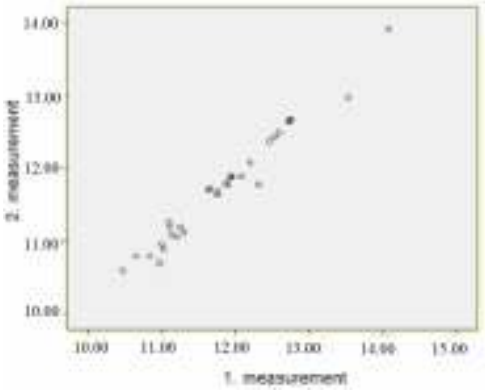

Fig. 5. Correlation of category U16 in the Dribbling test

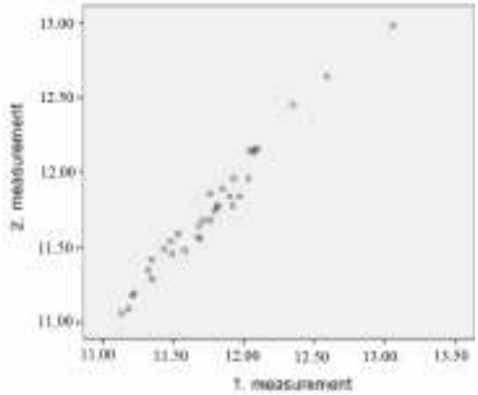

Fig. 6. Correlation of category U17 in the Dribbling test

Agility test

Table 4. Measurement results in the Agility test.

Order of measurement	1(U16)	2(U16)	1(U17)	2(U17)
Average	9.06	9.05	9.22	9.20
Median	9.08	9.12	9.22	9.24
Max.value	9.55	9.61	9.76	9.78
Min.value	8.48	8.39	8.00	8.11
Scatter	0.08	0.09	0.12	0.11
Standard deviation	0.28	0.31	0.35	0.33

In the U16 category, the reliability in the agility test achieved coefficient ICC = 0.9754. In U 17, the coefficient was ICC = 0.9781, what was considered by Hand and Taylor (1987) to be high.

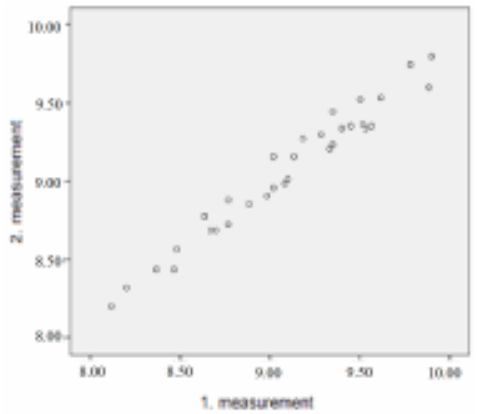

Fig. 7. Correlations of category U16 in the Agility test

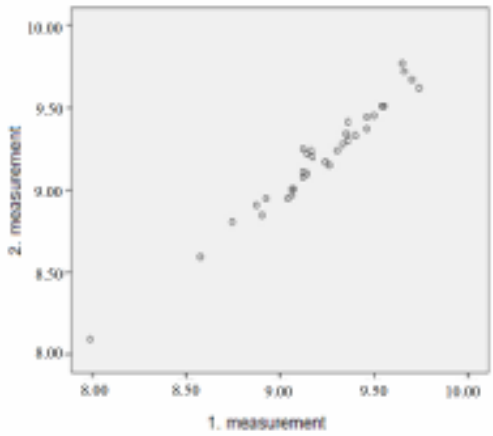

Fig. 8. Correlation of category U17 in the Agility test

Discussion

Diagnosis in soccer is as old as soccer itself. Most often, we do diagnose the individual gaming and game performances of the team. We also meet diagnostics of performance through general and specific motor performance tests. Soccer associations will determine test batteries at first to assess the performance of young players and based on their results, we can compare the results, for example, with players of other teams in the same age category, and to operatively adapt the training process with specific adjustments as required by players or teams. A specific example is the rating of soccer technique. By the eye, for example, in the game, it is possible to evaluate which player is technically better than the other player. On the other hand, from the point of view of diagnostics, the field of soccer player technique is relatively difficult to make in the test battery.

When Höner et al. (2009) tested a sport population in 4-year duration (twice a year), in the end, they were comparing the variations in this time period between the tests. The highest reliability was found during the four-year juggling test. In our testing, on the other hand, this player's skill was the lowest value compared to other skills (tests). On the other hand, the lowest reliability value of the authors found in the test was dribbling when the values of the test did not exceed the r = 0.5 limit over a four-year period.

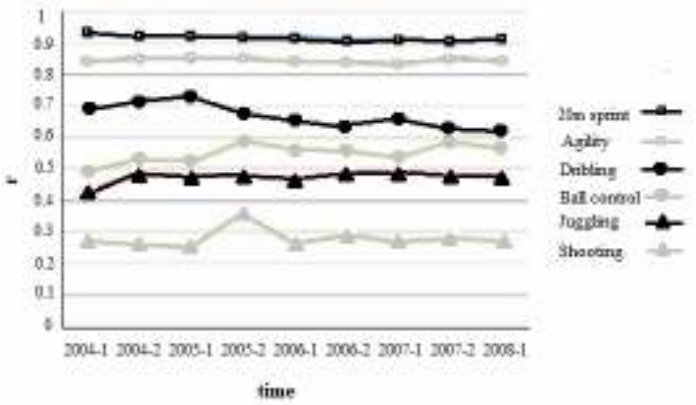

Fig. 9. r values during the four-year test period Höner et al. (2009)

AGILITY TEST

The issue of reliability of agility tests has been addressed in the past by a number of authors. E.g. authors Pauole, et al. (2000), Sporis et al. (2010) in a study to determine the reliability and factor validity of agility tests used in soccer rated slalom sprints for 4x5 m and sprint to 9-3-6-3-6-9 ms by 180 ° turns and confirmed high coefficient of reliability r = 0.992, respec. r = 0.976. In our case, we also found high reliability values r (U16) = 0.9754 and r (U17) = 0.9781 in the agility test. The following authors, on the basis of results and statistical significance, have determined which test is appropriate for the given player's function.

Also, Beekhuizen et al. (2009) measured the reliability of the agility test using a test-retest and also by ICC. The results showed that the agility test reliability reached 0.924 respec. 0.938, which is lower in comparison with our test.

According to Veale, Perace & Carlson (2010), test-retest reliability also achieved a strong correlation in the research of agility testing of australian soccer players when the resulting confidence value was r = 0.91. Measuring the reliability of agility tests was also dealt with by Kutlu et al. (2012), which included 6 tests in measurement, of which the overall coefficient of reliability r = 0.88 was later developed. Author Huijgen (2013) measured players with 2 days off. Her measurement contained a slalom sprint. In the slalom sprint, the value of r ICC = 0.71. Similarly, Lottermann et al. (2003), who assessed agility reliability (running through slalom gates). The authors calculated the reliability value at r = 0.85.

BALL GUIDANCE (DRIBBLING TEST)

Reilly, Holmes (1983) also measured reliability the ICC's ball dribbling, and on the basis of the test, achieved ICC = 0.95. Author Huijgen (2013) measured players with 2 days off. Her measurement included a shuttle dribble with a ball and slalom sprint with a ball. The resulting ICC values were as follows. To test the ball shuttle dribble test, the value of r ICC = 0.74 and value for the sprint slalom r ICC = 0.71.

Authors Kruger and Niedlich (1985) examined the reliability of the test in the form of the dribbling test and the Ball control test. The ball guidance test reliability achieved r = 0.99, resulting in a high correlation between testing. A test battery composed of 9 tests used by German clubs was evaluated by Lottermann et al. (2003) and Höner and Roth (2010), where they found that the reliability of the exercise is as follows: dribbling r = 0.49, which is a very low value and similar results were obtained by Höner et al. (2009) who measured the reliability of the dribbling test in the range of 2004-2008 and its value varied and changed during these 4 years at intervals r = 0.42 - 0.51.

Brahim, Bougatfa & Amri (2013) evaluated the players in the Ball-NMAT (New Multi-Change Direction Agility Test) test by Mujika et al. (2009) using the ICC, and which also evaluated CI (Stability Interval) and SEM (Standard Measurement Error). The results of these authors showed a high correlation r = 0.97, CI = 0.94 - 0.98 and SEM had a value of 0.09.

Also by Hongyou et al. (2013) surveyed the reliability of game skill dribbling during a match. They found the ICC's skill level dribbling r = 0.96, and the SEM was at 0.23.

BALL CONTROL (BALL CONTROL)

Kruger and Niedlich (1985) investigated the reliability of the ball control test together with the ball guidance test. Reliability of ball control testing also reached a level of r = 0.99, as in the ball guidance test, resulting in a high correlation between testing. A test battery composed of 9 tests used by German clubs was evaluated by Lottermann et al. (2003) and Höner and Roth (2010), where they found that the reliability of ball control exercises was r = 0.56. Höner et al. (2009) report that during the four-year period measurement, the reliability values in the ball control test were at r = 0.49-0.6 intervals.

JUGGLING (JUGGLING TEST)

The reliability of the juggling test was evaluated in 2004-2008 by Höner et al. (2009) and found that the juggling test coefficient determined values r = 0.6 - 0.74. Also, a test battery composed of 9 tests used by German clubs included the juggling ball test and was evaluated by Lottermann et al. (2003) who found that the Juggling Test's reliability was r = 0.60.

Author Rosad (2014) evaluated the level of reliability of the juggling test. However, this test combines in a chain with other gaming activities and therefore the results may be slightly different. The first test was combined with an individual's play activity passing, where the r value reached r = 0.95. Later, juggling also associated with passing, but already in variable conditions with a possible choice of solution, and the resulting value r = 0.66.

In further research by Bekris, Gioldadis (2016), juggling has been evaluated with several options. In the 11 juggling tests in different ways, reliability was determined with ICC. The lowest ICC value was achieved by juggling testing in the form of alternating juggling with right and left foot, where the ICC value was 0.53. In contrast, the highest ICC value was achieved by authors in juggling alternately with the leg and chest, where the ICC value reached ICC = 0.91.

Table 5. Values of the reliability interval U16

U16	Lower Bound	Upper Bound
Dribbling test	11.4588	12.0106
Ball control test	9.3068	9.8063
Juggling test	6.2784	7.4091
Agility test	8.9374	9.1670

Table 6. Values of reliability interval U17

U17	Lower Bound	Upper Bound
Dribbling test	11.5970	11.8941
Ball control test	9.0469	9.5384
Juggling test	7.3637	7.9304
Agility test	9.0892	9.3213

Conclusions

The three selected standardized tests for assessing the technical skills of the four we have examined have been shown by the ICC to be very high reliability tests. Only the juggling test has reached a reliability value of r < 0.55 **, which is unacceptable to other than social sciences according to Hand & Taylor (1987). In the test reported was r (U16) = 0.5384 and r (U17) = 0.4379. This means that the results of the juggling test do not give us such relevant information that we can think that this skill can greatly influence game performance. Values (reliability) of the other observed tests exceeded values of 0.97 **. The measurement error in the juggling test was at 0.71403 and in the other tests this measurement error was in most cases below 0.06 ***. The trainer must therefore consider what tests he / she takes into the test battery so that the exercises are as prevalent as possible in the match and movement structure that is characteristic of the players in the match. Therefore, it is necessary to test all the tests that are used in practice (Germany) so that we can objectively assess the level of technical skills of the player.

Acknowledgment

This research is part of project VEGA 1/0529/16: Effectivity of preparation in sport games in Slovak clubs and national teams on game performance.

References

Balsom, P., (1994). Evaluation of physical performance. In: *Football (Soccer)*. Oxford: Blackwell Scientific Publications, pp. 102-123. ISBN 0-632-03328-2.

Bangsbo, J., & Mohr, M.(2011). *Fitness testing in football*. Espergærde:Bangsbosport.ISBN 978-87-994880-0-1.

Beekhuizen, K., et al. (2009). Test-retest reliability and minimal detectable change of the hexagon agility test**.** Journal Strength Cond Res. 23(7):2167-2171, 2009.

Bekris, E., Gioldadis, A. (2016). Juggling test battery. *The Swedish Journal of Scientific Research* ISSN: 2001-9211.Vol. 3. Issue 3. March 2016.

Brahim, M.B, Bougatfa, R., & Amri, M. (2013). *Reliability, Validity and Minimal Detectable Change of a New Multi-Change of Directionagility Test for Soccer Players*. APE Advances in Physical Education 03.04 (2013): pp. 190-96. Web.

Bunc, V., (2006). Diagnostika trénovanosti hráčů fotbalu. In: R. Psotta et al. *Fotbal kondiční trénink*. Praha: Grada, s. 171-180. ISBN 80-247-0821-3.

Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. In: *Sports Medicine*. 38(4), pp. 297-316. ISSN 1179 203.

Hand, D., J., & Taylor, C., C. (1987). *Multivariate Analysis of Variance and Repeated Measures*. London: Chapman and Hall, 1987. ISBN 0 412 25810 2.

Höner, O., Laudenklos, P., Roth, K. & Schultz, F. (2009). Evaluation der technomotorischen Diagnostik im DFB -Talentförderprogramm. In M. Krüger, N. Neuber, M. Brach & K. Reinhart (Eds.), Bildungspotenziale im Sport (p. 144)

Höner, O. & Roth, K. (2010). *Testmanual für die sportmotorische Leistungsdiagnostik - Balljongliertest. Unveröffentlichtes Testmanual*, Eberhard-Karls-Universität Tübingen und Ruprecht-KarlsUniversität Heidelberg.

Hongyou, L.,et al. (2013). Inter-operator reliability of live football match statistics from OPTA Sportsdata, *International Journal of Performance Analysis in Sport*. 2013, 13, pp. 803-821.

Huijgen, B. CH. (2013). *A study on talent development and selection of youth soccer players*. ISBN: 978-90-367-6127-7. Groningen. 2013.

Charles, C. M. (1995). *Introduction to educational research*. druhé vydanie. San Diego, Longman. 1995.

Joppe, M. (2000). *The research process*. Retrieved february 25, 1998, http://www.ryerson.ca/~mjoppe/rp.htm

Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research*. Beverly Hills: Sage Publications.

Krűger, A. & Niedlich, D. (1985) ST-CF - *SPORTTEST: CREW-FUSSBALL-TEST* 1985; 100 Ballspiel-Fertigkeitstests, pp. 99-102. Schorndorf: Hofmann.

Kutlu, M., et al. (2012). *Comparison of a New Test for Agility and Skill in Soccer With Other Agility Tests*. School of Physical Education and Sport, Kırıkkale University, TURKEY. pp 143–150. ISSN (Online) 1899-7562

Lottermann, S., Laudenklos, P. & Friedrich, A. (2003). *Techniktraining - mehr als reine Ballarbeit*. Fussballtraining, 21 (4), 6-15.

Mackall, E., & Pynes, J. E. (1991). *The Predictive Validity of an Entry-Level Police Assessment Center - A Follow Up at Five Years*. International Personnel Management Association, Assessment Council.

Maxwell, J. A. (1992). *Understanding and validity in qualitative research*. Harvard Educational Review, 62(3), 279-300.

Pauole, L., et al. (2000) Reliability and Validity of the T-Test as a Measure of Agility, Leg Power, and Leg Speed in College-Aged Men and Women. *Journal of Strength and Conditioning Research*, 2000, 14(4), pp. 443–450. National Strength & Conditioning Association.

Pynes, J. E., & Bernardin, H. J. (1989). Predictive validity of an entry level police officer assessment center. Journal of Applied Psychology, 74 (5), 831-833.

Reilly, T, & Holmes, M. (1983) *A preliminary analysis of selected soccer skills*. Physical education review. 1983;6:64-71.

Rousson, V., Gasser, T., & Sefifert, .B. (2002). *Assessing intrarater, interrater, , and test-retest reliability of continuous measurements*. Stat. Med. 21:3431-3446-2002.

Safrit, M.J.E. (1976) *Reliability theory*. Washington. DC: American Alliance for Health, Physical Education, and Recreation, 1976.

Sporis, G., et al. (2010). Reliability and factorial validity of agility tests for soccer players. *J Strength Cond Res* 24(3): 679-686, 2010., p.194.

Veale, JP., Pearce, AJ. & Carlson, JS. (2010). Reliability and validity of a reactive agility test for Australian football*, International journal of sports physiology and performance*, vol. 5, no. 2, pp. 239-248.