RELIABILITY AND DISCRIMINATIVE VALIDITY OF NOVEL METHOD IN EVALUATION OF WATER POLO PERFORMANCE FOR CADET-AGE PLAYERS

Dean Kontić[1,2], Đivo Ban[1] and Šime Veršić[3]

[1]University of Dubrovnik, Dubrovnik, Croatia
[2]Water Polo Australia, Sydney, Australia
[3]Faculty of Kinesiology, University of Split, Croatia

Original scientific paper

Abstract
This study aimed to evaluate a specific model of evaluation of cadet players' performance in water polo. The participants in this study were 61 young water polo players, members of Croatian and Montenegrin teams (15-16 years of age). Variables included six indicators of specific water polo performance. Three water-polo experts independently evaluated each subject on all performance-indicators using a scale from 1 to 5. Coefficient of variation (CV) and intra-class-coefficient (ICC) were calculated for reliability of the evaluation for performance-score. Discriminative analysis and t-test were used to determine discriminative validity of the applied system of evaluation of specific water polo performance. The reliability of the offensive and defensive performance-indicators showed appropriate ICC values from 0.71 for agility-offense up to 0.81 for agility-defence and CV values showed relatively small within-subject variations. Independent-sample t-test, and discriminative canonical analysis showed significant differences between qualitative groups of players (starters vs. nonstarters) for observed performance-indicators. Investigation confirmed appropriate reliability and applicability of proposed method aimed at evaluation of water polo performance in offensive and defensive tasks.

Key words: *water polo, cadet players, polyvalence, efficacy, agility.*

Introduction

Water polo is a team sport, and as is any other team sports (i.e. basketball, volleyball, football, and handball) the final achievement of the whole team depends on characteristic achievement of all team members. However, the evaluation of the performance for each member of the water polo team is lacking, since mostly relies on game statistics (Escalante, Saavedra, Mansilla, & Tella, 2011; Escalante et al., 2012). Although a valuable indicator of "overall team achievement", team statistics do not accurately depict the performance of an individual athlete. For example, some athletes do not score frequently, do not earn the exclusion of an opponent etc. At the same time, they have highly important duties that valuably (but indirectly) contribute to the team statistics and consequently to the final result of the game. Such duties are, for example, the ability to play in different playing positions, effective transitioning between defence and offence, pressuring the opponent, passing the ball etc. Even the authors of studies that used game statistics as a performance indicator recognised the problem and observed such data as a "team-performance-indicator" and not as a "single-player-performance-indicator" (Escalante et al., 2011; Escalante et al., 2012, 2013). Other type of performance-evaluation in team sport including water polo comprises clustering players into specific "performance-levels" or "quality-groups", such as national-level- vs. international-level-players, and/or team-players vs. national-team-players (Idrizovic, Uljevic, Ban, Spasic, & Rausavljevic, 2013; Uljevic, Esco, & Sekulic, 2014).

Indeed, such an approach is also logical since it allows relatively objective differentiation between those players who perform at higher level, against those less successful. However, this approach also has significant flow. Namely, in some situations, the differentiation does not present a true quality. For example, in some cases (i.e. countries, teams) the difference between qualitative levels observed are so profound and extreme that results between observed groups are hardly applicable in real sport settings, although being scientifically valid (Melchiorri et al., 2015). From this brief introduction, it is clear that a more objective and profound approach to defining a real-game performance in water polo is needed. In a recent study Sekulic et al. proposed a model of evaluation of water polo performance and evaluated it on a sample of junior players (Sekulic et al., 2016). In brief, following the suggestions from other team sports (i.e. basketball) (Trninic, Perica, & Dizdar, 1999; Trninic & Dizdar, 2000; Trninic, Dizdar, & Dezman, 2000) authors proposed and validated a method of evaluation of six performance indicators (offensive agility, offensive polyvalence, offensive efficacy, defensive agility, defensive polyvalence, and defensive efficacy) and independently evaluated these indicators in junior water polo players of advanced level. In general, results showed appropriate reliability of evaluation-system (Sekulic et al., 2016). However, authors indicated necessity of further evaluation of proposed system of performance-evaluation in females and other age categories of male water polo players (Kontic, Zenic, Uljevic, Sekulic, & Lesnik, 2017).

Therefore, this study aimed to evaluate a specific model of evaluation of players' performance in water polo. Specifically, in this study we studied male cadet age players (15-16 years of age) and hypothesized that suggested model will allow a reliable and valid insight into real game performance of young water polo players.

Methods

The participants in this study were 61 young water polo players. At the time of the experiment they were 15–16 years old. The participants had been involved in water polo for more than 5 years, and were members of Croatian and Montenegrin teams. Variables included six indicators of specific water polo performance. Three variables were evaluated specifically for defensive and offensive performance as follows: (i) polyvalence; (ii) primary-position-efficacy, and (iii) agility. Polyvalence was defined as "the ability of the athlete to accomplish different playing-duties while playing at different playing positions over defense (polyvalence-defense) or offense (polyvalence-offense)".

Primary-position-efficacy was defined as "ability of the athlete to efficiently perform the position-specific tasks while playing at his primary playing-position during defense (efficacy-defense) or offense (efficacy-offense)". Agility was defined as "ability of the athlete to reasonably, quickly and efficiently change the playing position during defense (agility-defense) or offense (agility-offense)". For the purpose of the evaluation of performance-variables, three water-polo experts were interviewed. The interviewed experts were officials of the National Water Polo Federation and have been directly responsible for the selection of the National-Team members (i.e. National-team-coach and two assistants). Each expert (i.e. judge) independently evaluated each subject on all performance-indicators using a scale from 1 (poor; player is absolutely unreliable on a specific task and makes many mistakes) to 5 (excellent; player is absolutely reliable on a specific task and makes no mistakes at all). Therefore, regardless of their main duties and game-tasks, all of the participating athletes were evaluated on indicators of offensive and defensive-performance.

Following reliability analyses, the final score for each subject on each performance indicator was expressed as an average value of all three judges. Further, to establish the validity of the evaluation of defensive- and offensive-performance variables, players were additionally evaluated by their team-coaches with binomial criterion. Namely, team-coaches were asked to cluster their players into two qualitative groups (Starters vs. Non-starters) based on their own experience. The reliability of the evaluation for performance-scores was established by calculation of coefficient of variation (CV) and intra-class-coefficient (ICC) for the scores obtained from three evaluators (Hopkins, 2004; Huck, 2012). To establish the discriminative validity of the applied system of evaluation for specific water polo performance indicators, we have calculated: (i) discriminative canonical analysis and (ii) independent t-test between „starters" and „nonstarters" in all six performance indicators. Statistical significance of $p < 0.05$ was applied and Statistica ver. 13.0 (DELL, Tulsa, OK) was used for all calculations.

Results

Table 1 Descriptive statistics (Mean, Standard Deviation – SD) and reliability coefficients (CV – coefficient of variation, ICC – intra class coefficient) of the studied performance indicators.

	Mean	SD	CV	ICC
D_Poly (score)	2.80	1.44	0.05	0.79
D_Agil (score)	2.51	1.38	0.07	0.81
D_Effic (score)	2.03	0.91	0.05	0.77
O_Poly (score)	2.28	1.39	0.04	0.75
O_Agil (score)	2.00	1.31	0.04	0.71
O_Effic (score)	2.11	1.23	0.03	0.74

Legend: D_Poly – athlete's ability to accomplish different playing-duties in defense; D_Agil – athlete's ability to efficiently change the playing position during defense; D_Effic – athlete's ability to perform his primary position tasks in defense; O_Poly – athlete's ability to accomplish different playing-duties in offense; O_Agil – athlete's ability to efficiently change the playing position during offense; O_Effic – athlete's ability to perform his primary position tasks in offense.

Discriminant canonical analysis revealed signifciant differences ($p < 0.01$) between two observed groups in multivariate space. Indicators of defensive efficacy, offensive efficacy, and defensive agility mostly contribute to the differentiation between performance levels (Table 2).

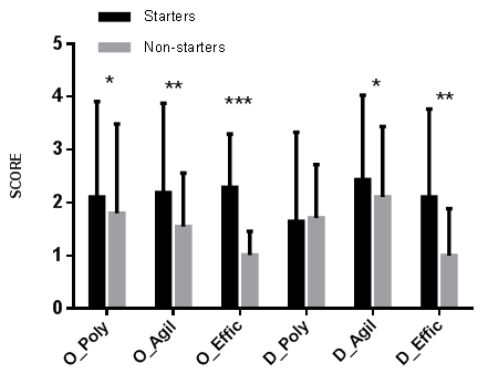

Figure 1. Differences between starters and non-starters in performance indicators with t-test for independent samples.

Legend: D_Poly – athlete's ability to accomplish different playing-duties in defense; D_Agil – athlete's ability to efficiently change the playing position during defense; D_Effic – athlete's ability to perform his primary position tasks in defense; O_Poly – athlete's ability to accomplish different playing-duties in offense; O_Agil – athlete's ability to efficiently change the playing position during offense; O_Effic – athlete's ability to perform his primary position tasks in offense: * denotes t-test significance of $p < 0.05$; ** denotes t-test significance of $p < 0.01$; *** denotes t-test significance of $p < 0.001$.

The reliability of the offensive and defensive performance-indicators showed appropriate ICC values from 0.71 for agility-offense up to 0.81 for agility-defense. The CV values showed relatively small within-subject variations when athletes' offensive and defensive performances were evaluated by different judges, ranging from 3% for efficacy-offense, to 7% for agility-defense (Table 1).

Table 2 Discriminant canonical analysis between qualitative groups of players (starters vs. Non-starters) in six performance indicators.

	Can Root
D_Poly (score)	0.11
D_Agil (score)	0.54
D_Effic (score)	0.66
O_Poly (score)	0.33
O_Agil (score)	0.41
O_Effic (score)	0.55
Can R	0.62
Wilk's Lambda	0.45
p	0.01
C: Starters	0.76
C: Nonstarters	-0.68

Legend: D_Poly – athlete's ability to accomplish different playing-duties in defense; D_Agil – athlete's ability to efficiently change the playing position during defense; D_Effic – athlete's ability to perform his primary position tasks in defense; O_Poly – athlete's ability to accomplish different playing-duties in offense; O_Agil – athlete's ability to efficiently change the playing position during offense; O_Effic – athlete's ability to perform his primary position tasks in offense; Can R- canonicam correlation coefficient; Can Root – structure of canonical root; C – position of centroid for each group.

Independent samples t-test showed significant differences between qualitative groups of players (starter vs. non-starters) in five of six observed performance-indicators. The high-quality group achieved significantly higher scores than low-quality group for all variables but polyvalence in defensive duties. Therefore, the validity of the evaluation for the performance-indicators was partially confirmed (Figure 1).

Discussion

Though the commonly employed statistical and performance-related methods in water polo allows identification of efficacy for the entire team, less emphasis has been placed on defining the real-game performance of individual players (Escalante et al., 2011; Escalante et al., 2012). Therefore, the results of appropriate reliability of the evaluation-method evaluated in this study are highly encouraging. The level of consistency was most likely due to the evaluators being highly objective and familiar with each athlete's qualities. Familiarity is known to be important for the reliability of testing (Sinclair, Hebron, & Taylor, 2014). Therefore, the fact that the evaluators followed each athlete throughout the previous competitive season allowed them to be well aware of their specific qualities which consequently led to accurate and reliable scoring. It is also important that the judges avoided subjectivity. Because the team of evaluators consisted of three highly experienced, they were well aware of necessity and importance of objective evaluation. Differences in reliability of evaluation among six observed performance indicators are actually a logical consequence of water polo game statistics. Namely, the lowest reliability is observed for performance indicator of offensive-agility (ICC: 0.71). This quality is highly influenced by overall team –tactics. More precisely, if one's team prefers dynamic offensive actions, players are easily recognized as being highly agile in offensive duties. However, in other situations one's offensive agility is possible not recognized because his team is mostly oriented toward static tactical solutions during offense.

Consequently, players are frequently scored according to team-tactics and not according to their true-performance. Although may seem surprising, the highest reliability is evidenced for another agility-performance, defensive-agility (ICC: 0.81). However, this finding is explainable knowing the characteristics of modern water polo game, where approximately one-third of all attacks are performed in power-play situations (Uljevic et al., 2014). Briefly, during the power-play (i.e. extra man in offense), players who perform defensive duties must present agility, and this is not influenced by eventual team-tactics. Indeed, they have to cover a larger space, have to swim faster, and change directions quickly and accurately. As a result, judges (scorers) easily recognized individual qualities of different players, which resulted even in high reliability of evaluation for offensive-agility-performance. Generally, scores on performance indicators are lower than those presented in previous studies on junior players, but this is a natural consequence of differences in competitive level (Kontic et al., 2017). Namely, it is clear that junior players who were studied previously show better performance qualities, and therefore earn higher scores than their younger peers (i.e. cadet players). Although discriminative validity was not confirmed for all variables, we may say that the system of performance indicators applied at cadet players is valid, mainly because discriminative analysis showed significant differentiation between performance levels (starters vs. non-starters).

However, it must be emphasized that some performance indicators doesn't contribute to differentiation between performance-groups. For example, low discriminative validity and therefore questionable applicability in evaluation of real-game performance in cadet water polo players is evidenced for offensive and defensive polyvalence. However, it is again a logical consequence of players' age. At this age group polyvalence in defensive duties is not so profound, and therefore the nonsignificant differences between starters and non-starters in this variable are logical. On the other hand it seems that factors of offensive- and defensive-efficacy hold the strongest validity of all

performance indicators. Again, it is not surprising since "game-efficacy" is a most vivid performance indicator of all included in this study. Even averagely informed evaluator will be able to evidence the efficacy of the water polo players. In offensive duties it is related to scored goals and/or earned exclusions (for Centers for example), while in defensive duties, it is evidenced oppositely (i.e. number of successful blocks, and earned offensive faults) (Kontic et al., 2017; Sekulic et al., 2016). Therefore, highly skilled professionals easily evaluated these performance indicators in observed players. Meanwhile, objective and consequently valid evaluation of players' agility and/or polyvalence in offensive and defensive duties is possible only for those with profound knowledge on a game. As a result, discriminative validity of these two performance indicators is somewhat lower.

Conclusion

This investigation confirmed appropriate reliability and discriminative validity of novel method aimed at evaluation of water polo performance in offensive and defensive tasks. However, when applying the evaluation system proposed herein, special attention must be paid on quality of judges. Namely, herein players were evaluated by objective and skilled evaluators which almost certainly positively influenced the reliability of the measurement. In this study we have applied proposed method of evaluation in male cadet players and further researches are needed to define reliability and applicability of this procedure in female water polo players. In doing so possible modification of the performance indicators and scores are warranted.

References

Escalante, Y., Saavedra, J.M., Mansilla, M., & Tella, V. (2011). Discriminatory power of water polo game-related statistics at the 2008 Olympic Games. *Journal of Sports Sciences, 29*(3), 291-298.

Escalante, Y., Saavedra, J.M., Tella, V., Mansilla, M., Garcia-Hermoso, A., & Dominguez, A. M. (2012). Water Polo Game-Related Statistics in Women's International Championships: Differences and Discriminatory Power. *Journal of Sports Science and Medicine, 11*(3), 475-482.

Escalante, Y., Saavedra, J.M., Tella, V., Mansilla, M., Garcia-Hermoso, A., & Dominguez, A.M. (2013). Differences and discriminatory power of water polo game-related statistics in men in international championships and their relationship with the phase of the competition. *Journal of Strength and Conditioning Research, 27*(4), 893-901.

Huck, S.W. (2012). *Reading Statistics and Research*. Boston: Pearson.

Idrizovic, K., Uljevic, O., Ban, D., Spasic, M., & Rausavljevic, N. (2013). Sport-specific and anthropometric factors of quality in junior male water polo players. *Collegium Antropologicum, 37*(4), 1261-1266.

Kontic, D., Zenic, N., Uljevic, O., Sekulic, D., & Lesnik, B. (2017). Evidencing the association between swimming capacities and performance indicators in water polo: a multiple regression study. *Journal of Sports Medicine and Physical Fitness, 57*(6), 734-743.

Melchiorri, G., Viero, V., Triossi, T., De Sanctis, D., Padua, E., Salvati, A., Galvani, C., Bonifazi, M, Del Bianco, R., & Tancredi, V. (2015). Water polo throwing velocity and kinematics: differences between competitive levels in male players. *Journal of Sports Medicine and Physical Fitness, 55*(11), 1265-1271.

Sekulic, D., Kontic, D., Esco, M.R., Zenic, N., Milanovic, Z., & Zvan, M. (2016). Sport-Specific Conditioning Variables Predict Offensive and Defensive Performance in High-Level Youth Water Polo Athletes. *Journal of Strength and Conditining Research, 30*(5), 1316-1324.

Sinclair, J., Hebron, J., & Taylor, P.J. (2014). The influence of tester experience on the reliability of 3D kinematic information during running. *Gait and Posture, 40*(4), 707-711.

Trninic, S., & Dizdar, D. (2000). System of the performance evaluation criteria weighted per positions in the basketball game. *Collegium Antropologicum, 24*(1), 217-234.

Trninic, S., Dizdar, D., & Dezman, B. (2000). Empirical verification of the weighted system of criteria for the elite basketball players quality evaluation. *Collegium Antropoloicum, 24*(2), 443-465.

Trninic, S., Perica, A., & Dizdar, D. (1999). Set of criteria for the actual quality evaluation of the elite basketball players. *Collegium Antropologicum, 23*(2), 707-721.

Uljevic, O., Esco, M. R., & Sekulic, D. (2014). Reliability, validity, and applicability of isolated and combined sport-specific tests of conditioning capacities in top-level junior water polo athletes. *Journal of Strength and Conditioning Research, 28*(6), 1595-1605.

*** Hopkins, W.G. (2005). A new view of statistics. http://www. sportsci. org/resource/stats.

POUZDANOST I DISKRIMINATIVNA VALJANOST METODE ZA PROCJENU IZVEDBE U VATERPOLU KOD IGRAČA KADETSKOG UZRASTA

Sažetak

Ovo istraživanje imalo je za cilj provjeriti specifični model procjene sportske izvedbe kod vaterpolista kadetskog uzrasta. Uzorak ispitanika činili su mladi vaterpolisti (N=61; 15-16 godina starosti) iz Hrvatske i Crne Gore. Svi igrači procijenjeni su na šest indikatora specifične vaterpolo uspješnosti u igri, a procjenu su provela tri vrhunska eksperta iz vaterpola, na ljestvici od 1 do 5. Kako bi se utvrdila pouzdanost procjene izračunati su koeficijenti varijacije (CV) i intra-klasni koeficijenti (ICC), a kako bi se utvrdila diskriminativna valjanost primijenjenih varijabli utvrđene su razlike između dvije kvalitativne skupine igrača (igrači koji započinju utakmicu i igrači koji ne započinju utakmicu u svojoj ekipi). Pouzdanost indikatora sportske izvedbe bila je zadovoljavajuća (ICC: 0.71 do 0.81), a najbolja pouzdanost utvrđena je za indikator agilnosti u obrani. Diskriminativna analiza i t-test za nezavisne uzorke ukazali su na zadovoljavajuću diskriminativnu valjanost indikatora u procjenu napadačke i obrambene sportske izvedbe.

Ključne riječi: *vaterpolo, kadeti, polivalentnost, efikasnost, agilnost.*

Received: August 02, 2017
Accepted: August 16, 2017
Correspondence to:
Šime Veršić
Faculty of Kinesiology,
University of Split
Teslina 6, Split 21000, Croatia
E-mail: simeversic@gmail.com