












Differences in judging judo matches between judges with different qualifications, coaches and players

Maciej Kostrzewa ^{1ABCDE}, Jan Supiński ^{2ABCDE}, Sławomir Kownacki ^{2ABCD}, Lulińska Ewelina ^{13ABC}, Jarosław Markowski ^{4ABDE}, Jan Pilch ^{5ADE}, Marcin Zak ^{1ABDE}, Arkadiusz Marzec ^{6DE}, Adam Maszczyk ^{1ABCDE}

¹ Institute of Sport Sciences, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland

² Department of Psychology, Academy of Physical Education in Wrocław

³ Gdansk University of Physical Education and Sport, Gdańsk, Poland

⁴ Department of Laryngology, Faculty of Medical Sciences, Medical University of Silesia in Katowice, Poland

⁵ Department of Physiological and Medical Sciences, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland

⁶ Institute of Physical Education Science, Jan Dlugosz University in Czestochowa, Poland

Authors' Contribution: A – Study Design, B – Data Collection, C – Statistical Analysis, D – Manuscript Preparation, E – Funds Collection

Abstract: The conducted research aimed to ascertain disparities in the proficiency of appraising combat scenarios among referees varying in their officiating expertise, coaches, and judo athletes. Ninety individuals partook in the study, comprising international referees, national referees, coaches, and judo athletes. Participants were assigned the task of allotting one of six possible ratings to each presented action and categorizing it into one of the groups. A meticulous analysis reveals certain distinctions as well as resemblances in the evaluations of selected actions among referees, coaches, and athletes. A notable incongruity in the appraisal of actions was noted in the shido penalty, where referees demonstrated the most precise assessments, while coaches and athletes rated them similarly. Substantial differences were identified in the evaluation of the hansoku-make penalty, as the coaches' cohort markedly disagreed with the actions depicted in the video and failed to acknowledge that the demonstrated actions warranted the most severe penalty permissible in judo competitions, namely the disqualification of the athlete during the match. Among the three groups examined, referees exhibited the highest level of accuracy in evaluating judo actions. International referees, in comparison to national referees, demonstrated superior assessment of judo actions, and this discrepancy was statistically significant. International referees more effectively evaluated judo actions such as yuko, waza-ari, shido, ippon, and hansoku-make in comparison to union referees. In one instance, union referees exhibited superior assessment of judo actions, specifically in the case of no score actions.

Keywords: assessing judo matches, Ippon, Waza-ari, Yuko, Shido, Hansoku-make

Corresponding author: Adam Maszczyk, email: a.maszczyk@awf.katowice.pl; Maciej Kostrzewa, email: m.kostrzewa@awf.katowice.pl

Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Received: 22.04.2024; Accepted: 20.05.2024; Published online: 3.07.2024

Citation: Kostrzewa M, Supiński J, Kownacki S, Lulińska E, Markowski J, Pilch J, Zak M, Marzec A, Maszczyk A. Differences in judging judo matches between judges with different qualifications, coaches and players. Phys Act Rev 2024; 12(2): 86-94. doi: 10.16926/par.2024.12.24



INTRODUCTION

Over the years, the rules of combat sports have undergone multiple modifications, necessitating a different approach to training and changes in the preparation of athletes in terms of technical-tactical aspects [1-4], increasing the education of referees to objectify assessments and minimize incorrect judgments [5-7]. As indicated by Franchini [8], these changes do not always yield positive results. In order to improve the rules of combat, the International Judo Federation (IJF) constantly introduces corrections to existing ones [9-11].

In scientific research regarding refereeing and rules of combat, the majority of studies have focused on judo referees. Only a few studies have included judo athletes and coaches. Choi [12] analyzed differences in a selected refereeing decision "hantei" (a situation of a draw), comparing the subjective assessment of judo athletes and referees. The results indicate that there is a discrepancy in assessments between the subjects. Brito and colleagues [13] analyzed 1411 judo matches in 36 international tournaments. The matches were recorded on video, and the analysis included the following criteria: phases of the match (approach, attack, defense, groundwork), penalties awarded by referees, types of attacks. Research has demonstrated that judo referees tend to award significantly more points to competitors who are fighting in their home country, thereby increasing the likelihood of victory for the home athletes. This phenomenon, known as "home advantage," is influenced by several factors, including the presence of supportive crowds and potential unconscious biases among referees. Studies have shown that referees' decisions are swayed by crowd noise, which can lead to awarding more points to home competitors, thereby impacting the overall outcome of the match [14-16].

Analyzing the available literature, there is a lack of reference to studies that would indicate factors modifying the correctness of assessment. Bar-Eli [17] conducted a review of studies concerning athletes, coaches, and referees, indicating their usefulness for psychologists, physical education teachers, and scientists. The authors, through their work, provide an opportunity for better understanding the complex nature of refereeing and decision-making in sports, while also pointing out the need for further research in this area. The purpose of this study is to fill this gap.

The conducted research primarily aimed to determine the differences in the ability to assess combat situations among referees with varying levels of officiating experience, coaches, and judo athletes. The secondary objective was to identify which of the studied groups most accurately evaluates sports combat situations. The inclusion of athletes and coaches in the study was prompted by observations that during competitions, signs of criticism and dissatisfaction are often seen from their side, which, as indicated by Souchon et al. [18-20], may influence the assessment of judo actions.

MATERIAL AND METHODS

Participants

Ninety individuals participated in the study, including international-level referees (15), national-level referees (15), coaches (30), and active judo athletes (30). The research was conducted separately for referees (R), coaches (C), and athletes (A). The mean years of expertise and standard deviation for each group were as follows: international-level referees (M = 12.4 years, SD = 2.2), national-level referees (M = 10.7 years, SD = 2.5), coaches (M = 15.1 years, SD = 4.1), and active judo athletes (M = 8.3 years, SD = 2.8). The studies obtained approval from the Bioethics Committee of the University of Physical Education in Katowice (no. 3/2021) and were carried out within the framework of grant N RSA2 025 52 and grant N RSA3 039 53.

Protocol

Each group underwent the same research procedure. Each participant was informed about the purpose of the study, its anonymity, and its scientific nature. Participants were informed of their right to withdraw from the study at any stage. The research was conducted in seminar rooms adapted for this purpose, with groups working separately.

Referees, coaches, and athletes assessed 60 actions based on video presentations, which they had to classify into one of six groups: no score (action should not be scored), yuko (5-point action), wazari (7-point action), ippon (10-point action, ending the match), shido (penalty for a minor infraction), hansoku-make (HSM), the highest penalty, resulting in disqualification.

Participants were tasked with assigning one of the six possible ratings to each presented action and classifying it into one of the groups. The assessment focused on the ability to correctly evaluate each of the 60 actions. Due to the various ways to achieve a favorable outcome in judo matches, the prepared material contained divisions into sequences related to scoring as well as penalty giving. Six groups of 10 actions each were distinguished, i.e., 10 actions assessed for Ippon, 10 for Waza-ari, 10 for Yuko, 10 for Shido, 10 for Hansoku-make, and 10 for No score (this situation should not be evaluated).

Statistical analysis

The obtained research results were analyzed using the PQStat ver. 1.4.2.324 statistical package. Data of nominal nature were compared using Pearson's chi-square test. For quantitative variables, the distribution was first checked for compliance with the normal distribution using the Shapiro-Wilk test. If a normal distribution was confirmed, comparison between two groups for quantitative variables was performed using the t-student test. For comparisons involving more than two groups, analysis of variance (ANOVA) was utilized. Upon identifying statistically significant differences in ANOVA analysis, post-hoc tests such as Scheffe, Bonferroni, and Duncan were employed. The presence of a significant relationship between quantitative variables (coping style with stress and judo actions) for each of the studied groups was verified using Spearman's rank-order correlation analysis. All statistical tests conducted were two-tailed, and p-values below 0.05 were considered statistically significant.

Scales were compared using the Mann-Whitney U test with continuity correction for the variable. A significance level of $p \leq 0.05$ was considered statistically significant, while a test probability level of $p \leq 0.01$ was considered highly significant.

RESULTS

Assessment of the validity of judo actions among referees of different classes, coaches, and athletes

The collective assessment of sixty selected judo actions by judges, coaches, and players indicates which actions were judged most accurately and which least accurately (Figure 1). Analytical data suggest that the surveyed groups most accurately evaluated shido and ippon actions. These significantly differed statistically from the others, with respective p-values of 0.002 and 0.003. Conversely, all surveyed groups least accurately evaluated the HSM action ($p=0.001$ vs. Ippon and Shido, and $p=0.446$; $p=0.551$; $p=0.614$ respectively vs. No Score, Yuko, and Wazari).

Figure 2 depicts the evaluations of individual actions by each group separately and the assessment of selected 60 judo actions. Detailed analysis reveals variability in the accuracy of assessment among the different groups. Judges most accurately and statistically significantly evaluated actions as Shido with $p=0.001$ compared to other groups, and least accurately evaluated Wazari. Coaches most accurately assessed actions as Ippon with $p=0.004$ compared to other groups, and least accurately assessed HSM

actions. Players most accurately evaluated No Score and Wazari with $p=0.003$ compared to other groups, and least accurately evaluated HSM.

Analyzing individual actions sequentially, it can be observed that players and judges rated No Score most accurately, while coaches rated it least accurately (statistical significance for coaches $p=0.002$). Conversely, all three groups rated Yuko and Wazari at similar levels (corresponding differentiating values $p=0.0876$ and $p=0.543$, respectively). In the assessment of Ippon actions, coaches and players achieved the highest accuracy ratings, while judges rated them least accurately (differentiating value $p=0.109$). Significant discrepancies in assessment were found for Shido actions, where judges rated them most accurately (as previously mentioned, $p=0.001$), while coaches and players rated them similarly ($p=0.786$). Significant differences occurred in the assessment of HSM actions, as the accuracy ratings of coaches and players statistically significantly differed from those of referees ($p=0.003$).

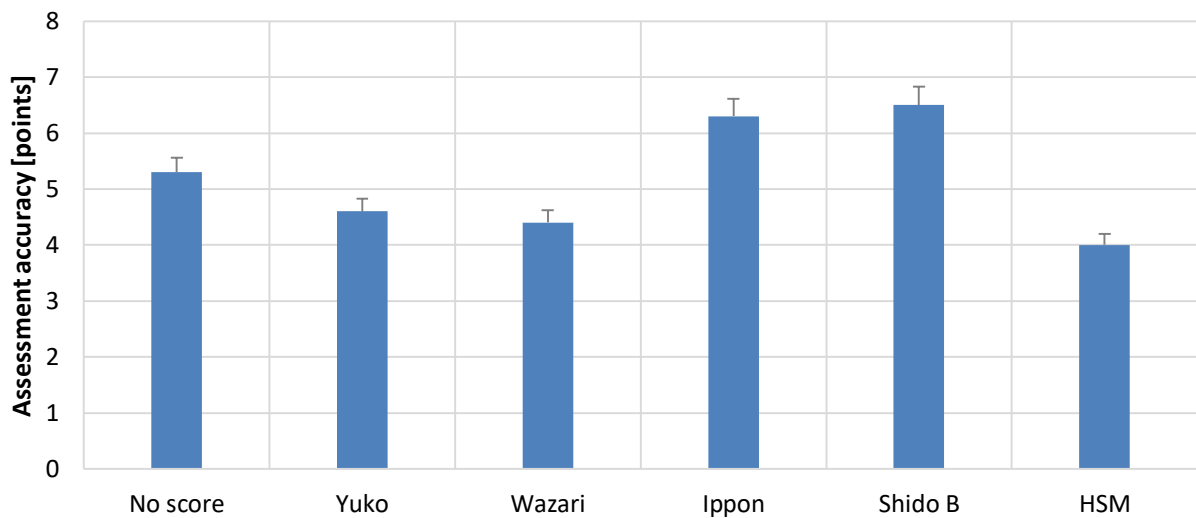


Figure 1. Global assessment accuracy level of individual judo actions by referees, coaches, and athletes

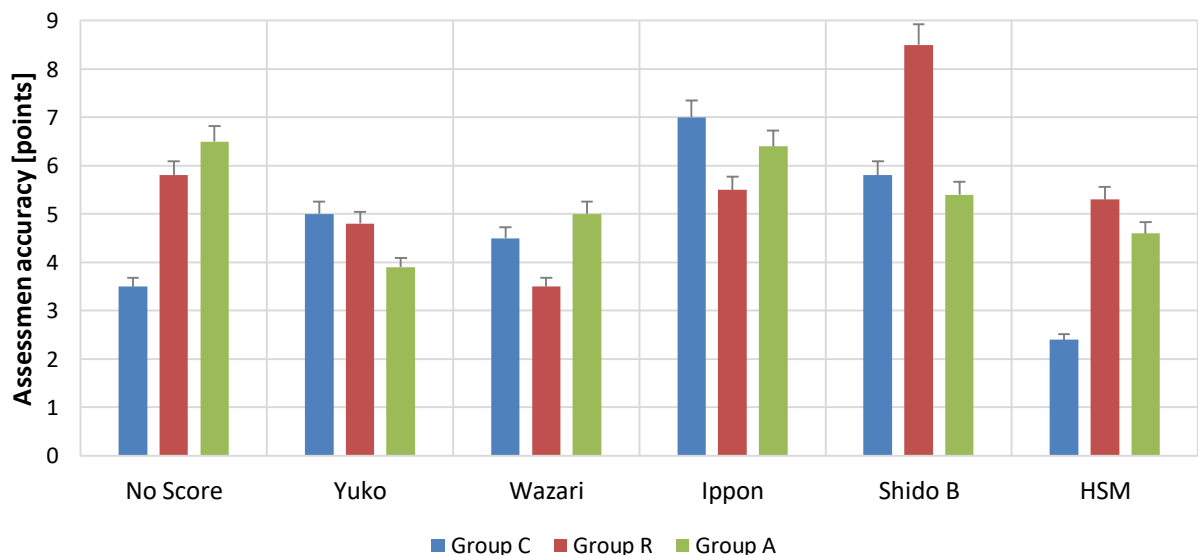


Figure 2. Detailed assessment accuracy level of individual judo actions by referees, coaches, and athletes

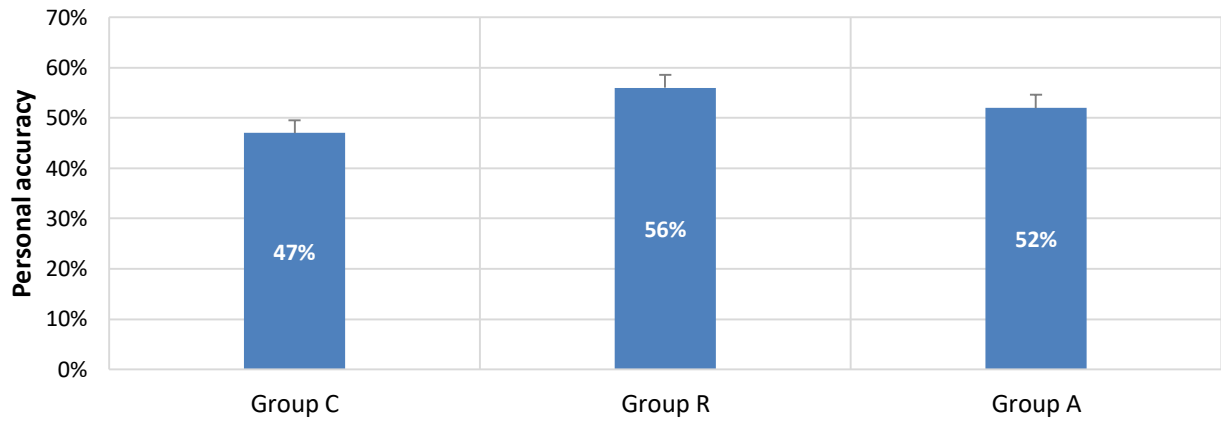


Figure 3. Percentage distribution of accuracy ratings made by referees, coaches, and athletes

Data presented in Figure 3 indicate that the surveyed groups differed only slightly in their percentage assessments of the 60 judo actions. The percentage distribution of assessments in Figure 3 shows that judges had 56% accurate responses, players 52%, and coaches rated the least accurately at 47%. Further analysis revealed that judges most accurately assessed Shido actions (96% of them), No Score (85%), and HSM (81%); coaches rated Ippon (93%) and Shido (84%) most accurately, while players rated Ippon (85%) and No Score (79%) most accurately.

Evaluation of judo actions by international class referees (IR) and national referees (NR)

The group of referees was divided into 15 international referees (IR) and 15 national referees (NR), respectively, and then the actions were analyzed in relation to all assessed segments of the fight. Analyzing the overall assessment of 60 actions by international and national referees, it can be observed that international referees significantly obtained more correct assessments with $p=0.041$ than national referees (Figure 4).

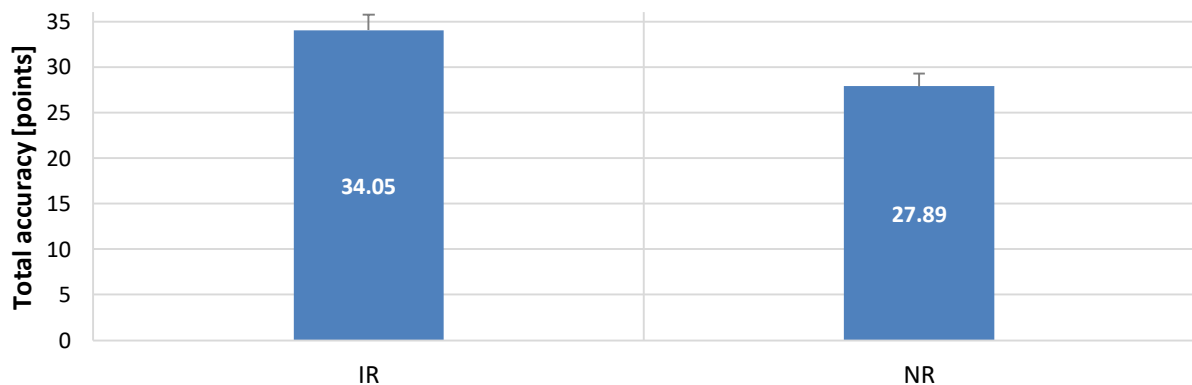


Figure 4. Overall assessment of the accuracy of 60 actions by international (IR) and national referees (NR)

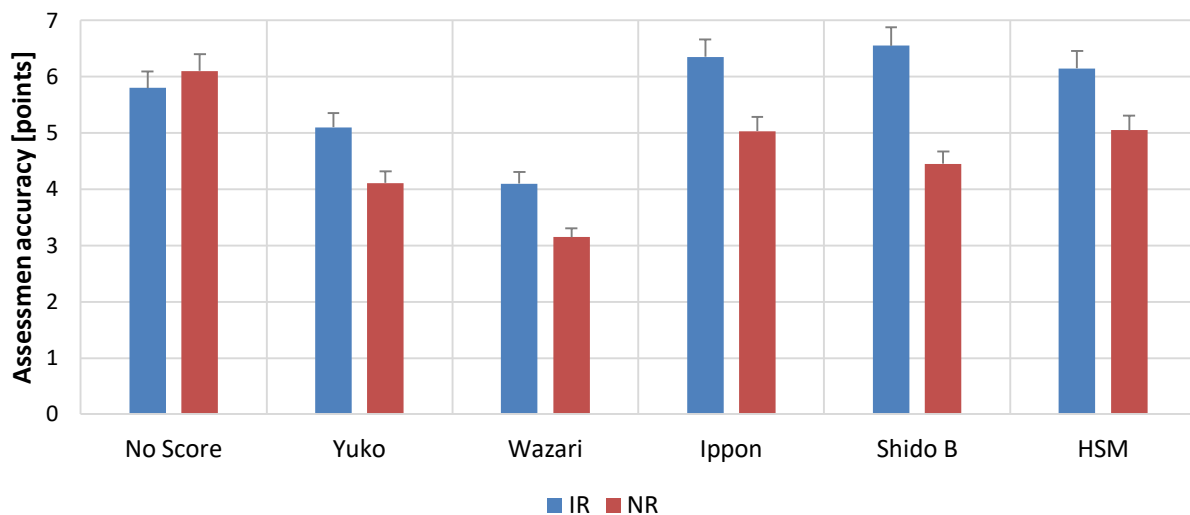


Figure 5. Evaluation of individual judo actions international (IR) and national referees (NR)

Examining the differences in the assessment of judo actions by international and national referees, it can be noted that only in one instance did national referees provide better assessments (not statistically significantly) of the presented actions - this pertained to the No Score action (Figure 5). In the remaining assessments, international referees significantly outperformed national referees. Specifically, for Yuko ($p=0.014$), Wazari ($p=0.036$), Ippon ($p=0.009$), Shido ($p=0.004$), and HSM ($p=0.010$). In summary, the assessment of actions by referees unequivocally indicated that international referees significantly better evaluated judo actions than national referees.

DISCUSSION

The research results indicate that statistically significant differences in the accuracy of judging judo match situations among referees of different classes, coaches, and athletes only occur in the case of referees of different classes. No such differences were observed in the other groups. Similar results, indicating that referees with higher qualifications are more effective in their assessments, have been obtained by other researchers. For example, in the study by Johansen [21], the level of anxiety was analyzed using the State-Trait Anxiety Inventory and an online decision-making questionnaire among top and lower-level soccer referees. The results of this study indicate that both in terms of anxiety and competence, top-class referees achieved better results compared to lower-class referees. In judo, Hayashi et al. [22] investigated factors influencing the assessment of referees with licenses A, B, C, and lower classes who evaluated four different throwing techniques executed in various directions. It was found that the average scores awarded for throwing techniques by referees with licenses A and B were significantly higher than those awarded by referees with license C. This suggests that more qualified referees assess throwing techniques more accurately. The authors proposed that less qualified referees may not fully appreciate the subtleties of throwing techniques. It should be noted that the classification of referees as A, B, and C, as mentioned by the authors, does not align with the current referee nomenclature, which uses different categories to describe referee qualifications. Furthermore, the similarity in the lettering system may cause confusion regarding the qualifications and experience levels of referees. Currently, there is a division into national, continental, and world referees, which corresponds to the possibility of officiating competitions of a certain rank.

The overall assessment of the analyzed actions indicates that referees judged judo actions best, but the percentage differences are slight. International and national referees achieved 56% correct responses, coaches 47%, and athletes 52%. Statistical analysis did

not reveal significant correlations between the groups studied. It is noteworthy that coaches achieved a lower percentage of correct assessments compared to athletes. Comparing individual judo actions, such as no score, yuko, wazari, ippon, shido, and hansoku-make, it can be seen that referees most accurately assessed actions such as shido and hansoku-make, coaches ippon and shido, and athletes ippon and no score. It is also important to mention that the average scores awarded for throwing techniques by referees with licenses A and B were significantly higher than those by referees with license C. This suggests that more qualified referees assess throwing techniques more accurately. The authors proposed that less qualified referees may not fully appreciate the subtleties of throwing techniques. Additionally, the classification of referees as A, B, and C, as mentioned by the authors, does not align with the current referee nomenclature, which uses different categories to describe referee qualifications. The similarity in the lettering system may cause confusion regarding the qualifications and experience levels of referees.

When the analysis was separately applied to international and national referees, the results showed differences in the assessment of sixty judo actions between these groups. The research results confirm that international referees assess judo actions presented on video best. Also, international referees accurately assessed the most judo actions compared to national referees, coaches, and athletes.

The obtained research results are consistent with studies in other sports disciplines, which have shown that referees of higher classes achieve better results than referees of lower classes. Similar results among ice hockey referees were obtained by Hancock [23], although a completely different research procedure was applied in their study. Fifteen higher-level referees and fifteen lower-level referees, wearing motion recorders on their heads during penalty decision-making, were analyzed, which was recorded on a computer screen. Visual behaviors and decision accuracy were recorded. Higher-level referees made decisions significantly more accurately than lower-level referees. In conclusion, the authors state that this may be because higher-level referees process decision-making information more effectively. Dosseville et al. [24], based on video recordings, analyzed how referees assessed actions on the mat and whether it was related to previous refereeing experience. The results indicated that previous experience influenced decisions and judgments made, highlighting the need for further improvement of refereeing skills.

Although the research results confirmed the hypothesis that international (continental and world) referees best assessed judo actions compared to coaches and athletes, it is remarkable that only referees and athletes exceeded the 50% threshold of correctly assessed judo actions, while coaches scored below 50% of correct assessments. However, when the assessment of actions by international referees was analyzed compared to national referees, the result was more favorable for international referees. An analysis of judo matches at the Grand Slam tournament in Paris in 2016 showed that shido penalties and ippon techniques were most frequently awarded by referees. This indicates that the frequent occurrence of these elements is significant in assessing judo actions [25].

It should be noted that in this study, referees, coaches, and athletes were not tested for their knowledge of current judo regulations. They assessed only actions presented in video recordings, not in natural mat conditions. Additionally, any referee wishing to obtain a referee license must pass an exam on the currently applicable Judo Contest Rules, Referee Rules, and practical refereeing exam. Passing these three parts of the exams is necessary to obtain referee licenses. Referees and coaches are required to attend a training and verification course annually to renew their licenses for the next year. To improve refereeing and coaching competencies, the authors propose the introduction (in addition to video action assessments) of a Referee Competence Test, which would include all three elements mentioned above, i.e., assessment of judo actions based on videos, knowledge of combat regulations, and referee and sports regulations. Based on this, referees would obtain further referee grades, and coaches would verify coaching licenses. It should be noted that after the training and verification course, updated materials are

available annually on the website of the Polish Judo Federation, which each of the mentioned groups can use at any time. During the conducted research, all groups had their first encounter with the presented video material.

LIMITATION

The obtained research results cannot be directly related to analogous ones in judo or other sports disciplines, as such research has not been conducted. Due to the changing judo regulations, it is not possible to repeat exactly such studies to check the ability to assess judo actions by referees, coaches, and athletes. However, they indicate certain trends among the studied groups. Some of the analyzed actions are not present in currently applicable regulations. Therefore, similar studies should be repeated to see if the results differ from those presented and if the trends are maintained.

CONCLUSION

In conclusion, the topic of the study is not fully explored due to the dynamics of changes in regulations and the actions themselves, which are subject to assessment by referees. However, statistically significant differences among individual groups are evident. An interesting topic for future research may also be a comparison of the accuracy of assessment in a group of referees who were former athletes compared to others. It is worth considering the authors' proposal for a training program for all interested parties to increase the level of understanding of regulations regarding the assessment of actions in combat by athletes, coaches, and referees.

Among the three groups studied - referees, coaches, and athletes, referees are the most accurate in evaluating judo actions. International referees, compared to national referees, assessed judo actions better, and this difference was statistically significant.

International referees better evaluated judo actions such as yuko, waza-ari, shido, ippon, and hansoku-make compared to national referees. In one instance, national referees assessed judo actions better, and it was the no score action.

Funding Statement: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

1. Gutiérrez-Santiago JA, Prieto-Lage I, Paramés-González A, Argibay-González JC, Reguera-López-de-la-Osa X, Gutiérrez-Santiago A. Should Blind Judokas and Partially Sighted Judokas Compete Together? A Reflection from the Study of the Temporal Structure of Combat. *J Hum Kinet* 2023; 89: 279-288.
2. Párraga Montilla, JA, Latorre Román, PA, Serrano Huete, V, Cabrera Linares, JC, Lozano Aguilera, E, Jiménez Reyes P. Differences in the Force-Velocity profile between Judoist and freestyle wrestlers. *Phys Act Rev* 2022; 10(1): 141-149. doi: 10.16926/par.2022.10.15
3. Serrano-Huete V, Latorre- Román P, García- Pinillos F, Morcillo-Losa J, Jiménez-Reyes P, Párraga-Montilla J. The variability of strength production capacities during a judo contest. *Phys Act Rev* 2023; 11(1): 1-10. doi: 10.16926/par.2023.11.01
4. Pronczuk M, Trybek G, Terbalyan A, Markowski J, Pilch J, Krzysztofik M, Kostrzewa M, Mostowik A, Maszczyk A. (2023). The Effects of EEG Biofeedback Training on Visual Reaction Time in Judo Athletes. *J Hum Kinet* 2023; 89: 247-258. doi: 10.5114/jhk/174272.
5. Hayashi H, Okada H, Masuchi K, Ishikawa Y, Ishii T, Sakamoto M, Komata K. Study o factors affecting different opinions expressed by Judo referees and judges regarding throwing techniques. *IJSHS* 2010; 55(2): 363-378.
6. Yoffie DB, Kwak M. Judo strategy. Harvard Business School, Harvard, 2001.

7. Feletti F, Bracco C, Takeko Molisso M, Bova L, Aliverti A. Analysis of Fluency of Movement in Parkour Using a Video and Inertial Measurement Unit Technology. *J Hum Kinet* 2023; 89: 5-18. doi: 10.5114/jhk/166581
8. Franchini E, Takito MY, Calmet M. European Judo Championships: impact of the new rule changes on points and penalties. *Inter Perform Anal Sport* 2013;
9. IJF Refereeing Rules, 2013; www.ijf.org.accessed.
10. IJF Refereeing Rules, 2017; www.ijf.org.accessed
11. Miyake K, Matsui T, Sato T, Ishikawa Y, Yokoyama T, Takezawa T, Kawabata K, Akimoto H, Sato S. Effects of the International Judo Federation Referee Rules on competition contents in the All- Japan Judo Championships: A 2008-2013 comparison. *Japan Acad Budo Res J Budo* 2015; 48(1): 17-27
12. Choi Gwan Young. An Analysis on the Different Views about Judo Hantei Between Referees and Players. *Kor J Physic Edu.* 1999; 38(1): 698-710.
13. Brito CJ, Miarka B, Díaz de Durana AL, Fukuda DH. Home Advantage in Judo: Analysis by the Combat Phase, Penalties and the Type of Attack. *J Hum Kinet* 2017; 57: 213-220.
14. Ferreira UJ, Goncalves Panisse VL, Measure B, Takito MY, Franchini E. Home advantage in judo: A study of the world ranking list. *J Sports Sci* 2013; 31(2): 213-218.
15. Jones MB. The home advantage in individual sports: An augmented review. *Psycho Sport Exer* 2013; 14(3): 397-404.
16. Krumer A. On Winning Probabilities, Weight Categories, and Home Advantage in Professional Judo. *J Sport Econ* 2014; 18(1): 77-96.
17. Bar-Eli M, Plessner H, Raab M. Judgment, Decision-making and Success in Sport. *Sport Exerc Psychol* 2011; 1st Edition. doi: 10.1002/9781119977032
18. Souchon N, Cabagno G, Rasclé O, Traclét A, Dosseville F, Maio GR. Referees' decision making about transgressions: The influence of player at the highest regional level. *Psychol Women Quart* 2009; 33(4): 445-452. doi: 10.1111/j.1471-6402.2009.01522.x
19. Souchon N, Cabagno T, Traclét A, Dosseville F, Livingstone A, Jones M. Referees' Decision-Making and Player Gender: The Moderating Role of the Type Situation. *JASP* 2010; 22(1): 1-16.
20. Souchon N, Fontayne P, Livingstone A, Maio GR, Mellac N, Genolini Ch. External Influences on Referees' decisions in Judo: The Effects of Coaches' Exclamations During Throw Situations. *JASP* 2013; 25(2): 223-233.
21. Johansen Bjorn Tore, Haugen T. Anxiety level and decision-making among Norwegian top-class soccer referees, *IJSEP* 2013; 11(2): 215-226.
22. Hayashi H, Aoyagi O, Takeuchi Y, Nakamura R, Okada H, Nanjo M. A study of IJF refereeing rules. *Japan Acad Budo Res J Budo* 2002; (35): 29-38
23. Hancock DJ, Ste-Marie DM. Gaze behaviors and decision making accuracy of higher-and lower-level ice hockey referees. *Psychol Sport Exer* 2013; 14 (1): 66-71.
24. Dosseville F, Laborde S, Raab M. Contextual and Personal Motor Experience Effects in Judo Referees' Decisions. *Sport Psycho* 2011; 25(1): 67-81.
25. Ceylan B, Şükrü SB. The Impact of New Rule Changes In Judo: A Comparison of Points and Penalties during Grand Slam Paris between 2016 And 2017. *Inter J Adv Sport Manag* 2017; 2: 91-94.