ORIGINAL ARTICLE
doi: 10.16926/par.2027.15.01

Fairness and Reliability Issues in Taekwondo Poomsae Competition: A One-Year Study

Jusik Park [1], **Hyeongjun Kim** [1]

[1] Keimyung University

Article information

Received: 2026-10-05

Accepted: 2026-10-07

Published: 2027-01-06

Keywords:

Taekwondo; Poomsae; judging reliability; scoring consistency; fairness; Cronbach's alpha; Presentation; Accuracy.

Copyright: © 2027 by the authors. Submitted for possible open access publication under the terms and conditions of Creative Commons Attribution 4.0 International.

Abstract:

Background: Recognized Taekwondo Poomsae competitions rely on judges' evaluations of technical execution and presentation. Because competition outcomes depend on subjective assessments, ensuring the reliability and fairness of scoring is essential. This study investigated the reliability of scoring components and score differences across competitions organized by the Korea Taekwondo Association (KTA), the Korean University Taekwondo Federation (KUTF), and World Taekwondo (WT). Methods: Official score sheets from 211 male individual Recognized Poomsae matches conducted between January and December 2024 were collected during KTA, KUTF, and WT competitions. Scores awarded independently by five judges for the Accuracy, Speed and power, Rhythm and tempo, and Expression of energy components were analyzed. Internal consistency was evaluated using Cronbach's α coefficients. Differences in scores among organizing bodies were examined using one-way ANOVA or Welch's ANOVA with appropriate post hoc tests. Results: The Accuracy component demonstrated consistently high reliability across all organizing bodies and most Poomsae types (Cronbach's α = 0.754-0.935). In contrast, the Presentation-related components exhibited substantially lower and more variable reliability, with most α values below 0.70. Significant differences in scores were observed among organizing bodies. WT competitions generally produced higher Accuracy scores but lower Presentation scores than KTA and KUTF competitions. These findings indicate that Presentation-related components are more susceptible to subjective interpretation and greater inter-judge variability than Accuracy. Conclusions: The findings demonstrate that scoring reliability in Recognized Taekwondo Poomsae differs substantially between technical and presentation components. While the Accuracy component showed satisfactory reliability, Presentation-related components require improved evaluation criteria, enhanced judge education, and greater standardization across organizing bodies. These findings may contribute to improving fairness and consistency in Taekwondo Poomsae judging and support the development of technology-assisted evaluation systems.

Corresponding author: Jusik Park, email: parkjs@kmu.ac.kr

How to cite: Park J, Kim H. Fairness and Reliability Issues in Taekwondo Poomsae Competition: A One-Year Study. Phys. Act. Rev. 2027;15(1):e1501. doi: 10.16926/par.2027.15.01

INTRODUCTION

Sport is founded on the principle of fairness, whereby competition outcomes should be determined by athletes' performance, skill, and effort rather than by external factors [1,2]. Fairness is widely recognized as a fundamental element for ensuring the credibility of competition results and the sustainable development of sport [3]. When athletes, coaches, and spectators perceive competition outcomes as unfair, confidence in the sport may decline, which can ultimately hinder its long-term growth and development [4]. Therefore, ensuring fairness is recognized as a critical responsibility shared by all sports stakeholders, including athletes, coaches, judges, and governing organizations.

To enhance fairness and objectivity, many sports have adopted technological and institutional innovations. For example, football has implemented the Video Assistant Referee (VAR) system [5], tennis utilizes the Hawk-Eye system [6], and baseball has introduced the Automated Ball-Strike System (ABS) [7]. Similarly, Taekwondo Kyorugi has adopted the Protector Scoring System (PSS), which has improved judging objectivity and reduced controversies associated with referee decisions [8]. These examples demonstrate that the establishment of fair and reliable evaluation systems is essential for the development and globalization of sport.

Particularly in judged sports such as figure skating, gymnastics, and diving, fairness is of even greater importance because competition outcomes are largely determined by judges' evaluations [9-11]. To minimize subjectivity, these sports have developed detailed scoring systems that clearly define technical difficulty, performance criteria, bonus points, and deduction rules. Such systems contribute to greater consistency among judges and enhance the credibility of competition outcomes.

Taekwondo Poomsae has become one of the major disciplines of Taekwondo, alongside Kyorugi and has experienced continuous growth over the past two decades. Since the inaugural World Taekwondo Poomsae Championships in 2006, Poomsae has evolved into a competitive discipline through various international events, including the Asian Games, the Universiade, and the World Championships [12]. In response to this development, the Korea Taekwondo Association (KTA) and World Taekwondo (WT) have continuously revised and improved competition rules and evaluation criteria to promote the advancement of the discipline [13,14].

Despite these efforts, concerns regarding fairness and objectivity in Poomsae judging continue to be raised by athletes, coaches, and judges [15-18]. Unlike figure skating or artistic gymnastics, where technical difficulty and execution are independently quantified through standardized scoring frameworks, Poomsae does not provide sufficiently detailed evaluation criteria for distinguishing between technical quality and performance execution [14,19,20]. As a result, competition outcomes may be influenced by judges' subjective interpretations, particularly in Presentation-related components.

Previous studies conducted by the present authors consistently reported relatively high reliability for the Accuracy component, whereas the Presentation component demonstrated substantially lower reliability [4,12,21,22]. These findings suggest that inter-judge variability is more pronounced in evaluation components requiring qualitative judgment.

Ensuring fairness and reliability in judging is an essential prerequisite for the development of any sport into a global discipline [12,23]. Although the importance of fairness and objectivity in Poomsae competitions has been increasingly recognized, studies that systematically quantify competition results and examine inter-judge reliability remain limited. Furthermore, most previous studies have analyzed only a single competition, thereby limiting the generalizability of their findings. Therefore, this study aims

to address this gap by quantitatively examining the reliability of judging in Taekwondo Poomsae competitions held in Korea over a one-year period, thereby enabling a more comprehensive and robust evaluation of judging reliability. This study seeks to provide empirical evidence to support the development of a more objective, standardized, and reliable judging system in Taekwondo Poomsae.

Based on previous findings and the existing literature, the following hypotheses are proposed: (H1) The Accuracy component will demonstrate higher inter-judge reliability than the Presentation-related components, (H2) Inter-judge reliability will differ among organizing bodies, (H3) Scores for both Accuracy and Presentation-related components will differ among organizing bodies.

MATERIAL AND METHODS

Data Sources

In this study, score sheets from Recognized Poomsae individual competitions held in Korea between March and September 2025 were used as the primary data for analysis. The competitions included national-level Poomsae tournaments organized by the Korea Taekwondo Association (KTA) and the Korean University Taekwondo Federation (KUTF), as well as an international Poomsae competition organized by World Taekwondo (WT).

For the national competitions organized by KTA and KUTF, only matches from the university division involving athletes enrolled in four-year universities were included. University athletes enrolled in four-year institutions were selected because they represent a relatively homogeneous population of specialized Poomsae athletes who undergo systematic training. Restricting the sample to this population minimized heterogeneity in competitive levels and improved comparability across competitions.

Athletes from two-year colleges were excluded to minimize potential differences in competition level, as these participants are less likely to represent specialized Poomsae athletes. For the international competition organized by WT, Under 30 division matches were selected to ensure comparability with the university level athletes. The WT Under 30 division was selected because its age range and competitive levels most closely correspond to those of university athletes competing in domestic competitions.

A total of 11,455 matches were analyzed, including 6,985 matches from seven KTA competitions, 3,990 matches from two KUTF competitions, and 480 matches from one WT competition. The unit of analysis was a single official score sheet corresponding to one athlete's performance of one Recognized Poomsae in an individual match. Each score sheet contained the scores assigned independently by all five judges for the Accuracy and Presentation-related components (Table 1).

Inclusion and Exclusion Criteria

To ensure the consistency and validity of the dataset, the following criteria were applied. First, Recognized Poomsae competitions are generally conducted using a tournament system. However, some recognized Poomsae competitions in 2025 adopted a cut-off system. To maintain consistency in competition format, Recognized Poomsae matches conducted under the cut-off system were excluded.

Table 1. Overview of Recognized Poomsae competitions and number of matches analyzed

Organizing Body	Division	Classifications	Gender	Number of Matches
KTA	University division	Individual match	Male	6,985
KUTF	University division	Individual match	Male	3,990
WT	Under 30	Individual match	Male	480
Total				11,455

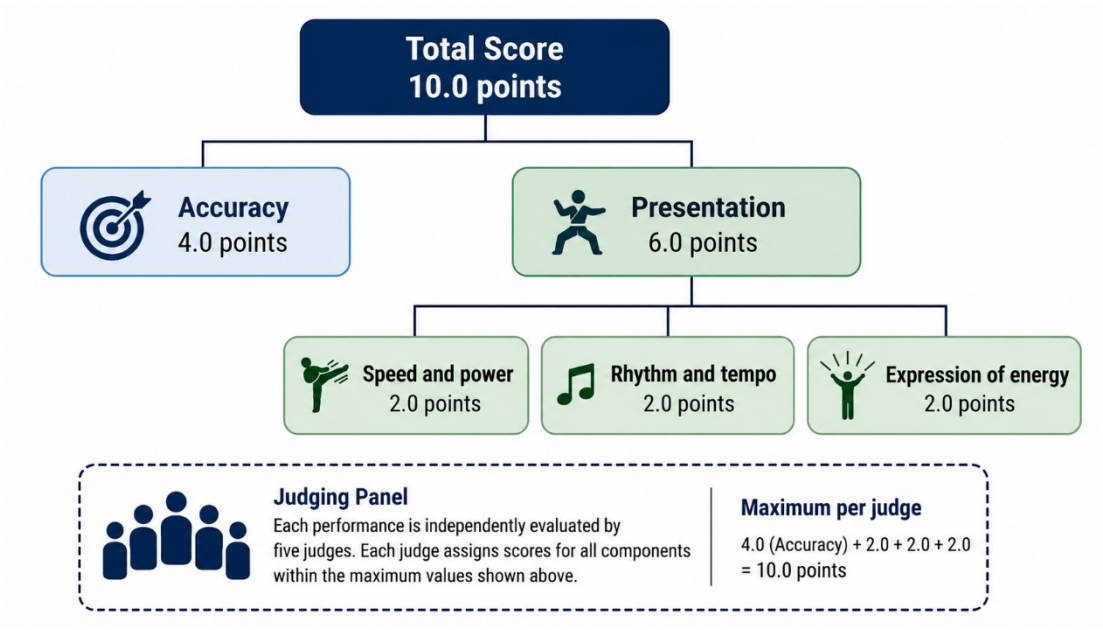


Figure 1. Scoring system used in Recognized Poomsae competitions. Each performance was independently evaluated by five judges. Each judge assigned a maximum of 4.0 points for the Accuracy component and 2.0 points for each Presentation component (Speed and power, Rhythm and tempo, and Expression of energy), resulting in a maximum total score of 10.0 points per judge.

Second, most Poomsae competitions are evaluated using a five-judge panel; however, some competitions employed a seven-judge system, or switched to a seven-judge system starting from the quarterfinals or finals. Given that variations in judging panels may influence scoring criteria and averaging methods, all matches evaluated under a seven-judge system were excluded.

Third, as the Jitae Poomsae is not included in WT international competitions, analyses involving Jitae were limited to comparisons between KTA and KUTF competitions.

Additionally, matches were excluded if video recordings were unavailable due to athlete withdrawal or if scoring results could not be verified due to technical errors during video transmission.

Data Collection and Instruments

Data were collected by the researchers through direct on-site visits to competition venues identified from the official annual schedules published by the Korea Taekwondo Association (KTA), the Korean University Taekwondo Federation (KUTF), and World Taekwondo (WT). Official score sheets were collected at each event and organized in Microsoft Excel (Microsoft Corp., Redmond, WA, USA), with separate datasets created for each organizing body. Scores assigned independently by all five judges for every match and Recognized Poomsae were recorded according to the scoring system shown in Figure 1.

To ensure data accuracy and completeness, the recorded scores were cross-checked against official match brackets and results published by the respective organizations. In addition, publicly available competition videos from YouTube were used to verify the recorded scores and minimize potential transcription errors.

Ethical Considerations

This study analyzed anonymized official scoring data obtained from publicly conducted Taekwondo Poomsae competitions. No human participants were recruited, no intervention or interaction with participants was performed, and no personally identifiable information was collected or analyzed. Therefore, according to institutional research ethics policies for studies using publicly available anonymous data, formal Institutional Review Board approval and informed consent were not required.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 28.0 (IBM Corp., Armonk, NY, USA). The analytical procedures were as follows:

First, the unit of analysis was a single official score sheet corresponding to one athlete's performance of one Recognized Poomsae in an individual match. Each score sheet contained the detailed scores assigned independently by five judges and served as the basic unit for all statistical analyses. Each performance was treated as an independent observation because the objective of this study was to evaluate judging outcomes at the performance level rather than to examine repeated measurements of individual athletes or judges across competitions.

Second, Cronbach's α coefficients were calculated for each scoring component across all Recognized Poomsae matches included in the analysis, regardless of organizing body. Cronbach's α was used to evaluate the internal consistency of judges' scoring across the five judges for each performance. A Cronbach's α value of 0.70 was adopted as the criterion for acceptable reliability.

Third, Cronbach's α coefficients were calculated separately for each scoring component according to organizing body (KTA, KUTF, and WT).

Fourth, homogeneity of variance was assessed using Levene's test prior to group comparisons. Given the large sample size of the present study, parametric analyses were considered appropriate based on the Central Limit Theorem, which supports the robustness of ANOVA to minor deviations from normality. Accordingly, one-way analysis of variance (ANOVA) was performed when the assumption of homogeneity of variance was satisfied, whereas Welch's ANOVA was applied when the assumption was violated. For Jitae Poomsae, which was not included in WT competitions, independent samples t-tests were conducted to compare only KTA and KUTF competitions. Statistical significance was set at $p < 0.05$.

Finally, when significant group differences were identified, Bonferroni post hoc tests were conducted following one-way ANOVA, whereas Games–Howell post hoc tests were performed following Welch's ANOVA. These post hoc procedures were selected to reduce the risk of Type I errors associated with multiple comparisons across Poomsae types and scoring components.

RESULTS

Reliability Analysis of Scoring Components (Overall)

Overall and organizing body-specific reliability of the scoring components is presented in Table 2. The Accuracy component demonstrated consistently high reliability

across all Poomsae types and organizing bodies (Cronbach's α = 0.754–0.935). In contrast, the Presentation-related components (Speed and power, Rhythm and tempo, and Expression of energy) exhibited lower and more variable reliability (Cronbach's α = −0.238–0.739). Across most Poomsae types, the WT competitions showed lower and less consistent reliability for the Presentation-related components than the KTA and KUTF, whereas the Accuracy component remained consistently high across all organizing bodies.

Table 2. Cronbach's α coefficients for the scoring components in Recognized Poomsae competitions (overall and by organizing body)

Poomsae type	Group	n	Accuracy	Speed and power	Rhythm & tempo	Expression of energy
Taeguk 7 Jang	Overall	2,760	0.834	0.652	0.639	0.655
	KTA	1,530	0.769	0.695	0.683	0.688
	KUTF	1,120	0.818	0.601	0.596	0.630
	WT	110	0.916	0.648	0.629	0.657
Taeguk 8 Jang	Overall	3,170	0.858	0.610	0.584	0.600
	KTA	1,960	0.830	0.678	0.656	0.666
	KUTF	1,070	0.856	0.648	0.658	0.668
	WT	140	0.924	0.340	0.150	0.205
Koryo	Overall	2,630	0.842	0.702	0.694	0.706
	KTA	1,660	0.817	0.739	0.727	0.728
	KUTF	800	0.852	0.635	0.636	0.657
	WT	170	0.824	0.096	0.208	0.344
Keumgang	Overall	3,110	0.869	0.627	0.608	0.614
	KTA	1,860	0.850	0.657	0.631	0.643
	KUTF	960	0.812	0.640	0.631	0.640
	WT	290	0.935	0.705	0.694	0.691
Taebaek	Overall	2,840	0.855	0.671	0.662	0.668
	KTA	1,770	0.799	0.718	0.706	0.707
	KUTF	950	0.869	0.704	0.704	0.712
	WT	120	0.880	–0.407	–0.428	–0.422
Pyeongwon	Overall	2,850	0.833	0.669	0.641	0.652
	KTA	1,770	0.822	0.713	0.679	0.683
	KUTF	1,060	0.810	0.604	0.593	0.614
	WT	20	0.754	0.027	0.096	–0.238
Sipjin	Overall	2,730	0.865	0.689	0.680	0.687
	KTA	1,610	0.850	0.711	0.696	0.698
	KUTF	1,010	0.843	0.668	0.662	0.673
	WT	110	0.932	0.668	0.665	0.701

Poomsae type	Group	n	Accuracy	Speed and power	Rhythm & tempo	Expression of energy
Jitae	Overall	2,830	0.838	0.672	0.662	0.675
	KTA	1,820	0.800	0.641	0.630	0.635
	KUTF	1,010	0.860	0.672	0.663	0.683
	WT	–	–	–	–	–

Differences in Accuracy Scores by Organizing Body

Levene's test indicated that the assumption of homogeneity of variance was violated for all Accuracy comparisons. Therefore, Welch's ANOVA was used for all analyses presented in Table 5. Significant differences among the organizing bodies were observed across all Poomsae types from Taeguk 7 Jang to Sipjin. Overall, WT competitions produced the highest Accuracy scores, followed by KTA and KUTF competitions. For Jitae Poomsae, Accuracy scores were significantly higher in KTA competitions than in KUTF competitions.

Differences in Presentation-related Scores by Organizing Body

Levene's test indicated that the assumption of homogeneity of variance was violated for all Presentation-related comparisons except for the Expression of energy component of Pyeongwon Poomsae. Accordingly, Welch's ANOVA was applied to all Presentation-related analyses presented in Table 3, except for the Expression of energy score of Pyeongwon Poomsae, which was analyzed using a conventional one-way ANOVA. Significant differences among the organizing bodies were observed for all three Presentation-related components across all Poomsae types (Table 4). Overall, KTA competitions tended to receive higher scores than KUTF and WT competitions for most Poomsae types, although the ranking patterns varied slightly across Poomsae types and Presentation components. For Jitae Poomsae, KTA competitions consistently yielded significantly higher scores than KUTF competitions.

Table 3. Differences in Accuracy scores by organizing body

Poomsae Type	Organizing Body	M ± SD	Statistic	*p*	Effect size	Post-hoc
Taeguk 7 Jang	KTA (n=1,530)	3.25 ± 0.10	F = 174.85	<0.001	η^2 = 0.193	c>a>b
	KUTF (n=1,120)	3.18 ± 0.13				
	WT (n=110)	3.47 ± 0.21				
Taeguk 8 Jang	KTA (n=1,960)	3.24 ± 0.11	F = 80.80	<0.001	η^2 = 0.113	c>a>b
	KUTF (n=1,070)	3.18 ± 0.16				
	WT (n=140)	3.43 ± 0.29				
Koryo	KTA (n=1,660)	3.22 ± 0.13	F = 80.18	<0.001	η^2 = 0.063	c>a>b
	KUTF (n=800)	3.15 ± 0.15				
	WT (n=170)	3.27 ± 0.13				
Keumgang	KTA (n=1,860)	3.17 ± 0.18	F = 108.18	<0.001	η^2 = 0.098	c>a>b
	KUTF (n=960)	3.12 ± 0.17				
	WT (n=290)	3.35 ± 0.26				
Taebaek	KTA (n=1,770)	3.24 ± 0.12	F = 111.68	<0.001	η^2 = 0.093	c>a>b
	KUTF (n=950)	3.15 ± 0.16				
	WT (n=120)	3.31 ± 0.28				
Pyeongwon	KTA (n=1,770)	3.21 ± 0.14	F = 50.00	<0.001	η^2 = 0.040	c>a>b
	KUTF (n=1,060)	3.18 ± 0.16				
	WT (n=20)	3.53 ± 0.16				

Poomsae Type	Organizing Body	M ± SD	Statistic	*p*	Effect size	Post-hoc
Sipjin	KTA (n=1,610)	3.23 ± 0.12	F = 126.63	<0.001	$\eta^2 = 0.135$	c>a>b
	KUTF (n=1,010)	3.16 ± 0.18				
	WT (n=110)	3.44 ± 0.22				
Jitae	KTA (n=1,820)	3.22 ± 0.15	t = 9.361	<0.001	d=0.373	a>b
	KUTF (n=1,010)	3.17 ± 0.15				
	WT (n=0)	-				

Table 4. Differences in Presentation-related scores by organizing body

C	Poomsae type	KTA (M ± SD)	KUTF (M ± SD)	WT (M ± SD)	Statistics
Speed and power	Taeguk 7 Jang	1.62 ± 0.06	1.57 ± 0.07	1.57 ± 0.10	F=178.29, p<0.001, η^2=0.114, a>b, c
	Taeguk 8 Jang	1.62 ± 0.08	1.58 ± 0.07	1.56 ± 0.12	F=70.77, p<0.001, η^2=0.045, a>b, c
	Koryo	1.61 ± 0.07	1.58 ± 0.08	1.60 ± 0.08	F=69.87, p<0.001, η^2=0.053, b<a, c
	Keumgang	1.60 ± 0.07	1.56 ± 0.08	1.55 ± 0.11	F=97.31, p<0.001, η^2=0.064, a>b, c
	Taebaek	1.62 ± 0.07	1.57 ± 0.08	1.54 ± 0.08	F=159.40, p<0.001, η^2=0.112, a>b>c
	Pyeongwon	1.61 ± 0.07	1.57 ± 0.07	1.53 ± 0.07	F=110.90, p<0.001, η^2=0.072, a>b>c
	Sipjin	1.61 ± 0.07	1.58 ± 0.07	1.60 ± 0.11	F=87.68, p<0.001, η^2=0.057, b<a, c
	Jitae	1.62 ± 0.07	1.58 ± 0.07	—	t=13.096, p<0.001, d=0.528, a>b

C	Poomsae type	KTA (M ± SD)	KUTF (M ± SD)	WT (M ± SD)	Statistics
Rhythm and tempo	Taeguk 7 Jang	1.60 ± 0.07	1.55 ± 0.07	1.56 ± 0.11	F=182.39, p<0.001, η^2=0.115, a>b, c
	Taeguk 8 Jang	1.60 ± 0.08	1.55 ± 0.07	1.54 ± 0.11	F=113.44, p<0.001, η^2=0.045, a>b, c
	Koryo	1.59 ± 0.07	1.55 ± 0.08	1.59 ± 0.07	F=85.35, p<0.001, η^2=0.066, b<a, c
	Keumgang	1.58 ± 0.08	1.54 ± 0.07	1.53 ± 0.12	F=97.81, p<0.001, η^2=0.061, a>b, c
	Taebaek	1.60 ± 0.07	1.55 ± 0.08	1.54 ± 0.09	F=156.09, p<0.001, η^2=0.108, a>b, c
	Pyeongwon	1.59 ± 0.07	1.55 ± 0.07	1.55 ± 0.05	F=109.35, p<0.001, η^2=0.070, a>b, c
	Sipjin	1.60 ± 0.08	1.55 ± 0.08	1.58 ± 0.11	F=93.36, p<0.001, η^2=0.061, a>b
	Jitae	1.59 ± 0.07	1.55 ± 0.08	—	t=12.936, p<0.001, d=0.526, a>b
Expression of energy	Taeguk 7 Jang	1.58 ± 0.07	1.53 ± 0.07	1.55 ± 0.11	F=160.17, p<0.001, η^2=0.098, a>b, c
	Taeguk 8 Jang	1.58 ± 0.08	1.54 ± 0.07	1.52 ± 0.11	F=94.38, p<0.001, η^2=0.058, a>b, c
	Koryo	1.58 ± 0.07	1.54 ± 0.08	1.58 ± 0.06	F=60.91, p<0.001, η^2=0.048, b<a, c
	Keumgang	1.56 ± 0.08	1.53 ± 0.07	1.52 ± 0.12	F=77.84, p<0.001, η^2=0.048, a>b, c
	Taebaek	1.58 ± 0.07	1.54 ± 0.08	1.52 ± 0.10	F=131.36, p<0.001, η^2=0.090, a>b, c
	Pyeongwon	1.57 ± 0.07	1.54 ± 0.07	1.53 ± 0.06	F=85.37, p<0.001, η^2=0.057, a>b, c
	Sipjin	1.58 ± 0.08	1.54 ± 0.08	1.57 ± 0.11	F=79.93, p<0.001, η^2=0.054, b<a, c
	Jitae	1.58 ± 0.07	1.54 ± 0.07	—	t=11.264, p<0.001, d=0.450, a>b

C – Component. Values are presented as mean ± standard deviation (M ± SD). Welch's ANOVA was applied for all comparisons except the Expression of energy score for Pyeongwon, which was analyzed using one-way ANOVA. Jitae Poomsae was compared only between KTA and KUTF using the independent-samples t-test because WT did not include this Poomsae. η^2 = eta squared; d = Cohen's d. Post hoc comparisons were performed using Games–Howell or Bonferroni tests, as appropriate. a = KTA, b = KUTF, c = WT.

DISCUSSION

Reliability Analysis of Scoring Components in Recognized Poomsae Competitions

The present study demonstrated that the Accuracy component exhibited consistently higher reliability than the Presentation-related components across organizing bodies and most Poomsae types. These results are consistent with previous studies analyzing male individual Recognized Poomsae competitions, thereby supporting the present findings [4,12,21,22].

Furthermore, prior research has reported that the Accuracy component in Recognized Poomsae is based on clearly defined criteria, including movement trajectories, prescribed techniques, and clearly established deduction rules, allowing for relatively objective scoring and smaller inter-judge variability [24]. For instance, specific deduction elements such as heel lifting can be assessed using observable and discrete criteria, similar to error-based scoring systems employed in judged sports such as figure skating and artistic gymnastics, where execution or grade-of-execution deductions are systematically applied [25]. In practice, both the Korea Taekwondo Association and World Taekwondo have continuously refined judging criteria by standardizing aspects such as movement names, spatial positions, distances, gaze direction, weight transfer, and deduction factors. These characteristics may partially explain the relatively high reliability observed in the Accuracy component. However, because the effects of judge education and training were not directly examined in the present study, this interpretation should be considered with caution.

In contrast, the Presentation-related components, Speed and power, Rhythm and tempo, and Expression of energy, showed lower overall reliability. Although some cases exceeded the threshold of 0.7, most analyses stratified by organizing body yielded reliability values below 0.7. These findings are consistent with previous studies, further supporting the present results [4,12,21,22].

One possible explanation is related to the composition of judging panels in international competitions. Because WT competitions involve judges from different countries and regions, variability in judging backgrounds and interpretations may influence scoring consistency. However, the present study did not examine judges' backgrounds, experience, or panel composition directly. Therefore, this explanation should be interpreted as a plausible hypothesis rather than a confirmed mechanism.

In addition, the relatively low reliability of the Presentation components may be associated with recent changes in the competition structure. In Recognized Poomsae, Presentation has gained increasing importance, with scoring weights currently allocated as 4.0 points for Accuracy and 6.0 points for Presentation. However, previous research has shown that judges may place even greater emphasis on Presentation in practice [26, 27]. As the relative importance of Presentation increases, subjective judgment plays a more critical role in determining competition outcomes. Despite this shift, standardized evaluation criteria and judge training systems for Presentation components remain insufficient. This may contribute to inconsistencies in judging.

It is also important to consider the relatively small sample size of World Taekwondo competitions included in this study. Therefore, caution should be exercised when generalizing these findings, and the results should be interpreted as indicating the potential for instability in scoring at the international level rather than definitive conclusions.

Additionally, previous studies have suggested that accumulated judging workload may reduce concentration and scoring consistency by increasing cognitive fatigue [27]. Even experienced judges may experience difficulty maintaining consistent evaluations when assessing multiple performances consecutively [21]. Although judging workload and

fatigue were not directly measured in the present study, these factors may represent possible explanations for the relatively low reliability observed in the Presentation-related components. Future studies should examine these variables directly to better understand their potential influence on judging reliability.

Taken together, these findings suggest that improving the reliability of Presentation-related components requires not only clearer evaluation criteria but also more standardized judge education and further investigation of judge- and competition-related factors that may influence scoring consistency.

Analysis of Score Differences in Recognized Poomsae Competitions

The present study found systematic differences in Accuracy scores among organizing bodies, with WT competitions generally producing higher scores than domestic competitions. As international interest in Taekwondo Poomsae has recently increased, the overall performance level of athletes has become more standardized and elevated both domestically and internationally [21].

In addition, international Poomsae competitions organized by WT are structured into tiers ranging from G1 to G10 depending on the scale of the event. Athletes earn ranking points based on competition tier and placement, and these rankings directly influence seeding in subsequent competitions [14]. The WT competition analyzed in this study corresponded to a G2-level event, which typically attracts a large number of national team-level or elite athletes.

One possible explanation is that WT G-ranked competitions generally attract a larger proportion of elite athletes, which may contribute to the relatively higher Accuracy scores observed in the present study. However, athletes' competitive level was not directly evaluated; therefore, this interpretation should be considered with caution.

In contrast, the analysis of differences in Presentation-related components across organizing bodies revealed that scores in WT competitions were generally lower than those observed in competitions organized by the KTA and the KUTF across most Poomsae types.

In WT competitions, judges from various countries participate because of the international nature of the events. Differences in technical background, judging experience, and cultural perspectives may lead to variation in the interpretation of Presentation-related criteria. However, because judges' backgrounds and panel composition were not analyzed in the present study, this explanation should be regarded as speculative.

Furthermore, these findings are consistent with the results of the reliability analysis in this study, which showed that most Presentation-related components had Cronbach's α values below 0.7. In other words, the observed differences in Presentation scores among organizing bodies do not simply reflect differences in score magnitude but rather indicate that Presentation components are highly dependent on subjective interpretation by judges.

In contrast, in figure skating, an event with competitive characteristics similar to Taekwondo Poomsae, athletes are assigned base values according to a pre-submitted technical element sheet, and judges evaluate performances based on Grade of Execution (GOE) scores ranging from ±5 (up to ±50% of the base value), along with detailed deduction criteria [28]. In other words, regardless of the organizing body, a standardized scoring system is established to ensure that the same technical difficulty and execution quality receive consistent scores.

This structure is designed so that final rankings are determined by both technical composition and quality of execution. As a result, athletes and coaches can reasonably predict potential score variations, and trust in judging outcomes can be maintained at a relatively high level.

Taken together, the present findings suggest that scores in Taekwondo Poomsae may reflect not only athletes' performances but also characteristics of the judging environment. Although the present study did not directly examine these factors, further research is needed to clarify how they influence score interpretation across organizing bodies.

Therefore, future efforts should focus on developing more standardized scoring criteria, particularly for Presentation-related components, while strengthening judge education and establishing evaluation guidelines that facilitate consistent score interpretation across organizing bodies.

CONCLUSION

The findings of this study demonstrate that Recognized Poomsae competitions exhibit differences in judging reliability and scoring according to evaluation components and organizing bodies. In particular, the Presentation-related components exhibited substantially lower inter-judge reliability than the Accuracy component.

The present findings suggest that the observed variability in scoring is associated with differences across evaluation components and organizing bodies. Specifically, the relatively low reliability of the Presentation-related components indicates the need for clearer evaluation criteria, greater consistency in the application of scoring guidelines, and continued efforts to strengthen judge education. These improvements may contribute to enhancing fairness and reliability in Taekwondo Poomsae competitions.

However, this study has several limitations. First, the data from World Taekwondo competitions were limited in sample size and were based on a single competition, which restricts the generalizability of the findings. Future research should incorporate longitudinal data from multiple international competitions, including G1–G8 level events, to further examine changes in scoring patterns and reliability across organizing bodies and competition levels.

Second, Cronbach's α was used to evaluate the internal consistency of judges' scoring across the five judges for each performance. Although this approach provides useful information regarding the consistency of scoring patterns, it does not directly assess inter-rater agreement. Therefore, future studies should consider complementary reliability measures, such as the Intraclass Correlation Coefficient (ICC) or Generalizability Theory, to provide a more comprehensive assessment of judging reliability.

Third, although each performance was treated as an independent unit of analysis, some athletes and judges may have participated in multiple competitions during the study period. Consequently, complete statistical independence of observations cannot be guaranteed. This potential pseudoreplication should be considered when interpreting the present findings. Future studies should employ longitudinal or multilevel analytical approaches that explicitly account for repeated participation by athletes and judges.

Overall, the present findings indicate that judging reliability and scoring patterns in Recognized Poomsae vary according to evaluation components and organizing bodies. Most notably, the relatively lower reliability observed for the Presentation-related components highlights the need for continued refinement of evaluation criteria and more consistent application of judging standards. These findings provide empirical evidence that may support future efforts to improve the fairness and consistency of judging in Taekwondo Poomsae competitions.

Limitations and Future Directions

Future studies should also examine whether prolonged judging sessions are associated with physiological or cognitive fatigue using objective indicators, similarly to recent investigations of fatigue dynamics in elite taekwondo athletes [29].

Funding Statement: The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of Interests: The authors declare no conflict of interest regarding the material discussed in the manuscript.

Authors' Contribution: *Jusik Park:* Conceptualization, Methodology, Project Administration, Supervision, Validation, Writing – Review & Editing; *Hyeongjun Kim:* Data Curation, Investigation, Writing – Original Draft Preparation

REFERENCES

1. Abad D. Sportsmanship. Sport Ethics Philos. 2010; 4(1): 27–41. doi: 10.1080/17511320903365227
2. Agnew D, Henderson P, Woods C, Agnew D, Henderson MP. Ethics, integrity and well-being in elite sport: A systematic review. Sport J. 2017; 19: 1–19.
3. Costa SJ, Silva A, Dias TA, Marinho D, Batalha N, Roque R, et al. Integrity and transparency in sports: A survey review. Open Sports Sci J. 2025; 18: e1875399X353976. doi: 10.2174/011875399X353976250203045235
4. Kim HJ, Park JS. Reliability analysis of domestic and international Taekwondo Poomsae competition results. J Coach Dev. 2023; 25(6): 177–184. doi: 10.47684/jcd.2023.11.25.6.177
5. Holder U, Ehrmann T, König A. Monitoring experts: Insights from the introduction of video assistant referee (VAR) in elite football. J Bus Econ. 2022; 92(2): 285–308. doi: 10.1007/s11573-021-01058-5
6. Liu C, Wang W, Liu H, Wang J. Application of Hawk-Eye technology to sports events. In: Proc 2nd Int Conf Inf Technol Contemp Sports (TCS). 2022. p. 1-5. doi: 10.1109/TCS56119.2022.9918811
7. Lee K, Han K, Ko J. Analyzing the impact of the automatic ball strike system in professional baseball through a case study on KBO league data. Sci Rep. 2025; 15: 44459. doi: 10.1038/s41598-025-28142-y
8. Jang WJ, Lee KK, Lee WJ, Lim SH. Development of an inertial sensor module for categorizing anomalous kicks in taekwondo and monitoring the level of impact. Sensors (Basel). 2022; 22(7): 2591. doi: 10.3390/s22072591
9. Heiniger S, Mercier H. Judging the judges: Evaluating the accuracy and national bias of international gymnastics judges. J Quant Anal Sports. 2021; 17(4): 289–305. doi: 10.1515/jqas-2019-0113
10. Koh S. Study on judgment recognition of figure skating coaches. Kinesiology. 2021; 6(4): 1–11. doi: 10.22471/kinesiology.2021.6.4.01
11. McGee M. Maximizing diver score by examining discrepancies in diver competency and judges' marks. J Sports Anal. 2025; 11: 22150218251403307. doi: 10.1177/22150218251403307
12. Kim HJ, Park JS. Reliability analysis of Taekwondo Poomsae competition results. J Coach Dev. 2023; 25(1): 231–237. doi: 10.47684/jcd.2023.01.25.1.231

13. Korea Taekwondo Association. Reglement Technique. Available from: https: //www.koreataekwondo
.co.kr/d008 (accessed 2026 Jul 8).
14. World Taekwondo. WT Poomsae Competition Rules & Interpretation. Available from: https: //m.worldtaekwondo.org/rules-wt/rules.html?sc=02 (accessed 2026 Jul 8).
15. Chaâbane MA, Said IB, Chaari A. Deep learning-based autoencoder for objective assessment of Taekwondo Poomsae movements. In: Proc 13th Int Conf Sport Sci Res Technol Support. 2025. p. 170-177. doi: 10.5220/0013706100003988
16. Choi T, Heo S. Towards automated Taekwondo Poomsae scoring: Evaluating modern neural network techniques. J Int Acad Phys Ther Res. 2025; 16(1): 3401–3409. doi: 10.20540/JIAPTR.2025.16.1.3401
17. Kim TM, Kwak TY, Lee JH. Exploring measures to enhance fairness through the improvement of Taekwondo Poomsae competition rules. Taekwondo J Kukkiwon. 2025; 16(3): 33–46. doi: 10.24881/tjk.2025.16.3.33
18. Rafique MA, Khaskheli F, Hassan MT, Naseer S, Jeon M. Employing automatic content recognition for teaching methodology analysis in classroom videos. PLoS One. 2022; 17(2): e0263448. doi: 10.1371/journal.pone.0263448
19. International Skating Union. ISU Sports Rules. Available from: https: //www.isu.org/figure-skating-rules/ (accessed 2026 Jul 8).
20. International Gymnastics Federation. Code of Points and Rules. Available from: https: //www.gymnastics.sport/site/rules/ (accessed 2026 Jul 8).
21. Kim HJ, Park JS. Reliability analysis for each scoring item in Taekwondo Poomsae competition. J Coach Dev. 2023; 25(4): 190–197. doi: 10.47684/jcd.2023.07.25.4.190
22. Kim HJ, Park JS. Analysis of scoring differences by scoring item according to the competition stage of Taekwondo official Poomsae. J Coach Dev. 2026; 28(3): 224–231. doi: 10.47684/jcd.2026.03.28.3.224
23. Lee SM, Kim KN. Awareness and improvement plan for the fairness of Taekwondo Free Poomsae scoring regulations. Taekwondo J Kukkiwon. 2022; 13(3): 23–32. doi: 10.24881/tjk.2022.13.3.23
24. Jo YM, Ryu JK, Kim YS. Kinematic analysis of accuracy evaluation of Pyeongwon Poomsae. Korean J Sport Biomech. 2015; 25(4): 431–440. doi: 10.5103/kjsb.2015.25.4.431
25. Chu SW, Kwak JH. Improvement of Taekwondo Freestyle Poomsae competition rules: Focusing on the technical domain. J Martial Arts. 2025; 19(3): 129–157. doi: 10.51223/kosoma.2025.05.19.3.129-157
26. Koo BJ, Lee YK. Analysis of the priority of the scoring items of the standing referees of the Taekwondo Poomsae competition. Korean J Converg Sci. 2025; 14(3): 1–13. doi: 10.24826/kscs.14.3.1
27. Shin SJ. A qualitative case analysis on the stress of referees of permanent Poomsae referees in Korea. Asian J Phys Educ Sport Sci. 2021; 9(4): 29–57. doi: 10.24007/ajpess.2021.9.4.003
28. Korea Skating Union. ISU General Regulations. Available from: https: //skating.or.kr/faq/ (accessed 2026 Jul 9).
29. Kaya MZ, Shahidi SH. Unraveling lactate and fatigue dynamics in elite Taekwondo athletes during the 30-second continuous jump. Phys Act Rev. 2025; 13(2): 12–21. doi: 10.16926/par.2025.13.16