AUTHENTICITY AND DEPENDABILITY OF THE VOLLEYBALL SERVE PRECISION EXAMINATION

Nguyen Van Hoa¹
Nguyen Quoc Tram ²
Chau Hoang Cau³
Nguyen Huu Tri ⁴

ABSTRACT

Background: Assessing the precision of a volleyball overhead serve is essential for understanding players' skill levels and development. Existing methods may lack consistency and perceptiveness, warranting further investigation into a more reliable assessment tool. Objectives: This study aimed to scrutinize the substance, facade, perceptiveness, and consistency of a volleyball overhead serve for precision assessment. Specifically, it sought to evaluate the validity and reliability of the assessment tool across different skill levels, ranging from novice to experienced players.

Method: To achieve these objectives, a multi-faceted approach was employed. Fifty-five volleyball mentors, instructors, and athletes participated in the facade validity examination, while six academic experts assessed the substance validity. Additionally, thirty-eight university students, comprising both seasoned and inexperienced players, were involved to gauge the test's capability to differentiate between performance tiers. Perceptiveness was evaluated by comparing disparities in service precision across skill levels using the t-test for independent samples. A hierarchical clustering analysis was conducted using the Euclidean quadratic distance metric and the among-groups linkage method. Furthermore, the consistency of repeated testing was analyzed using standard error and intraclass correlation coefficients.

Result: The study yielded promising results regarding the validity, reliability, and perceptiveness of the volleyball overhead serve precision assessment. Facade validity demonstrated high agreement among participants, with consensus ranging from 81% to 88%. The test's substance validity and reliability were strong, as shown by good results in the evaluations by academic experts and high intraclass correlation coefficients for repeated testing (ICC = 0.85 for experienced participants and ICC = 0.82 for novices). Significant differences in performance were observed between experienced and novice players, further affirming the tool's ability to differentiate skill levels effectively.

Conclusion: The volleyball serves precision assessment emerged as a reliable and valid tool for evaluating players' performance across different skill levels. Its ability to discern skill nuances and provide consistent results makes it suitable for educational contexts, particularly in assessing novice players' development and guiding instructional strategies. This study emphasizes the importance of rigorous assessment methods in improving volleyball skills comprehension and advancement.

Keywords: Motor Learning, Repeated Testing, Performance Assessment, Athletic Achievement.

¹ Physical Education Department, Can Tho University, Vietnam. E-mail: nguyenvanhoactu@gmail.com
Orcid: https://orcid.org/0009-0003-6604-0021

² Physical Education Department, Phu Yen University, Vietnam. E-mail: tramnguyenq2019@gmail.com
Orcid: https://orcid.org/0009-0008-4479-0135

³ Physical Education Department, Can Tho University, Vietnam. E-mail: chauhoangcauctu@gmail.com
Orcid: https://orcid.org/0009-0004-4790-9825

⁴ Physical Education Department, Can Tho University, Vietnam. E-mail: tringuyenhuu22@gmail.com
Orcid: https://orcid.org/0009-0008-1751-5058
AUTENTICIDAD Y FIABILIDAD DEL EXAMEN DE PRECISIÓN DEL SERVICIO DE VOLEIBOL

RESUMEN

Antecedentes: La evaluación de la precisión de un servicio aéreo de voleibol es esencial para comprender los niveles de habilidad y el desarrollo de los jugadores. Los métodos existentes pueden carecer de coherencia y percepción, lo que justifica una investigación más profunda sobre la herramienta de evaluación más fiable. Objetivos: El objetivo de este estudio fue examinar la sustancia, fachada, percepción y consistencia de una sobre-carga de voleibol para una evaluación de precisión. Específicamente, se buscó evaluar la validez y confiabilidad de la herramienta de evaluación en diferentes niveles de habilidad, variando de jugadores iniciantes a experimentados.

Método: Para lograr estos objetivos, se utilizó un enfoque multifacético. Cincuenta y cinco mentores, instructores y atletas de voleibol participaron en el examen de validez de fachada, mientras que treinta estudiantes universitarios, incluyendo jugadores experimentados e inexpertos, fueron involucrados para evaluar la capacidad de la prueba para diferenciar entre niveles de rendimiento. La percepción se evaluó comparando las disparidades en la precisión de los servicios de diferentes jugadores para un examen de precisión. Específicamente, se realizó un análisis de agrupamiento jerárquico utilizando la métrica de distancia cuadrática euclidiana y el método de vinculación entre grupos. Además, se analizó la consistencia de las pruebas repetidas utilizando coeficientes de error estándar y de correlación intraclase.

Resultado: Se presentaron resultados promisорios en relación a la validez, confiabilidad y percepción de la sobre-carga de voleibol para una evaluación de precisión. La validez de la fachada demostró alto acuerdo entre los participantes, con un consenso vario en los 81% a 88%. La validez y confiabilidad de la sustancia testada fueron consistentes, con buenos resultados tanto en los especialistas académicos como en los estudiantes universitarios. Diferencias significativas en el rendimiento fueron observadas entre jugadores iniciantes y experimentados, lo que confirma la capacidad de la prueba para diferenciar los niveles de habilidad de forma eficaz.

Conclusión: El voleibol servicio es una herramienta confiable y válida para evaluar el desempeño de los jugadores en diferentes niveles de habilidad. Su capacidad de discernir nuances de habilidades y ofrecer resultados consistentes o en contextos educacionales, en particular la evaluación del desarrollo de jugadores iniciantes y estrategias instruccionales orientadoras. Este estudio enfatiza la importancia de métodos de evaluación rigurosos para mejorar la comprensión y el avance de las habilidades de voleibol.

Resultado: El estudio arrojó resultados prometedores con respecto a la validez, confiabilidad y perceptividad de la evaluación de precisión del servicio aéreo de voleibol. La validez de la fachada demostró un alto acuerdo entre los participantes, con un consenso que oscila entre el 81% y el 88%. La validez de la sustancia y la fiabilidad de la prueba fueron fuertes, como lo demuestran los buenos resultados en las evaluaciones de expertos académicos y los altos coeficientes de correlación intraclase para pruebas repetidas (ICC = 0,85 para participantes experimentados e ICC = 0,82 para novatos). Se observaron diferencias significativas en el rendimiento entre los jugadores experimentados y novatos, lo que confirma aún más la capacidad de la herramienta para diferenciar los niveles de habilidad de manera efectiva.

Conclusión: El voleibol sirve para la evaluación de precisión surgido como una herramienta confiable y válida para evaluar el rendimiento de los jugadores en diferentes niveles de habilidad. Su capacidad para discernir matices de habilidades y proporcionar resultados consistentes lo hace adecuado para contextos educativos, particularmente para evaluar el desarrollo de los jugadores novatos y guiar las estrategias de instrucción. Este estudio enfatiza la importancia de los métodos de evaluación rigurosos para mejorar la comprensión y el avance de las habilidades de voleibol.

Palabras clave: Aprendizaje Motor, Pruebas Repetidas, Evaluación del Rendimiento, Logro Atlético.

1 INTRODUCTION

Elaborate motor tasks have been the central focus of inquiries conducted by researchers into motor learning. A typical motor endeavor, such as tossing salon darts, has been employed in numerous studies (23). Varied motor skills from diverse sports, encompassing baseball (11), volleyball (5), badminton (10), tennis (28), and golf (1), have been explored in other investigations. The predominant motor skill scrutinized in volleyball has been the serve, whereas additional studies (4) have included open motor abilities such as the bump and two-hand volley/set. The stability of the environment during the volleyball serve contributes to robust generalizability and practical applicability. A lot of research on motor learning has used the volleyball serve (13; 7; 8), with performance accuracy as the main variable of interest. Other studies have looked at other volleyball skills, like spike or bump (9). However, few assessments have undergone a rigorous validation process. Ensuring a test's ecological validity for educational contexts is paramount (21; 24). The AAHPERD Serve Test (2) and the Serve Test within the Volleyball Skills Test Battery (3) stand as two established assessments for evaluating volleyball serves. Their results shed light on the areas where opponents encounter the greatest challenges during real-game scenarios. However, while integral to the game's dynamics, the court's boundaries pose limitations for assessing changes in accuracy attributes over time. For instance, achieving the highest score when the ball lands on the back edge of the court may compromise the validity and consistency of the test, as skilled players may aim for this area but
narrowly miss, resulting in a zero score, as highlighted by Bartlett J et al (3). Conversely, to ensure a minimum score, beginners ought to target regions with lower scores, as suggested by Morrow Jr. et al (20). Low-sensitivity testing thus makes it arduous to assess the learning trajectory of the student during the initial learning phase, which is characterized by inaccuracy and inconsistency in performance (27). To appraise learning trajectories and discern performance disparities linked to novice accuracy, practitioners must devise novel assessments and ascertain their validity and reliability. Researchers exploring motor learning should take heed of this, as there is a dearth of validated tools adept at evaluating intricate skill accuracy with enough sensitivity to detect subtle performance nuances. This study scrutinizes the accuracy of a volleyball overhand serve test, encompassing its substance, facade, and sensitive validity, as well as its test-retest reliability.

2 METHODS

Procedure: Encompassed Employing a 4-point Likert scale, six academic experts (four specialists in motor behavior and two in sports pedagogy with a specialization in volleyball) assessed the applicability, lucidity, and relevance of the test to ascertain its validity. To further scrutinize the facial validity of the volleyball overhand serve precision examination, 60 professionals, comprising coaches of youth and adult volleyball teams (N = 25), school-employed volleyball instructors (N = 15), and volleyball athletes (N = 20), independently participated in the face validity assessment, utilizing the same Likert scale. In addition to possessing extensive professional backgrounds (5–10 years = 33%; 10–20 years = 22%; 20–30 years = 35%; 30–40 years = 10%), all coaches and instructors also boasted prior athletic experience (18% professionals, 82% amateurs). Thirty-eight individuals, undergraduates identifying as right-handed and aged between 18 and 34, engaged in the study to evaluate sensitivity and validity. Two groups, differing in experience levels, were formed. Eighteen volunteers from the experienced cohort were recruited from a nearby institution and underwent regular volleyball training, averaging two sessions per week over the preceding two years. The novice group comprised fifteen volunteers who had never undergone any formal volleyball instruction. The study received approval from Vietnam National University Ho Chi Minh City in Research, and all procedures adhered to the 7th Edition of the APA ethical guidelines.

Proceedings: Participants in the volleyball serve assumed positions five meters from the court's "A" side, with both feet firmly planted and facing a target located on the court's "B" side.
The objective was to accurately strike the target's bull's eye, positioned four meters from the net on the "B" side of the court, when executing an overhand serve in volleyball.

**Figure 1**

*Depiction of the designated work area.*

The target area consisted of four circular zones with diameters measuring 1, 2, 3, and 4 meters, respectively. Scoring accuracy toward the bull's eye within the target ranged from 2 to 22 points. A triangle, delineated by adhesive tape and originating from the serving point, extended through the side of the target. Each line had a width of five centimeters, with higher scores awarded for landing on these lines. In addition to standard trials, serves hitting the central region of the target scored 22 points, with scores decreasing by 2 points as they moved away from the center. Serves landing within the court but beyond the triangular region of the target earned 14 points, while those making contact with the upper band of the net and falling towards the target garnered 12 points. Serves that made contact with the net and traversed it without reaching the intended zone received eight points; those that landed on the court but failed to land within the triangular area were granted ten points. Serves outside the court were scored at six points. Additionally, four points were allotted for the triangular region on the "A" side of the court, while two points were assigned outside of this region. The experiment was conducted on an official volleyball court with the net positioned at a height of 2.24 meters, utilizing adhesive tape to demarcate the target area, and employing standard volleyball balls.

**Process:** Specialists were called upon to participate in assessing the content's validity. Subsequently, the intended audience, comprising volleyball coaches, educators, and athletes, was enlisted to evaluate the face's validity following the final version's receipt. Upon this, individuals were invited to partake in the research endeavor. Before commencing the trials, participants were briefed on the study's objectives, methodologies, and associated risks,
following which they completed informed consent forms. Participants were then instructed on the task's objective and provided with precise descriptions of the required serve—specifically, the overhand serve. After this initial instruction, each participant received guidance on serving positioning, with performance data kept confidential. A sole researcher stationed themselves to the left of the participant, issuing them the ball, managing the six to eight-second intertrial intervals, and signaling the participant to act by using cues like "prepare" or "go" prior to each trial. Post-trial, a second researcher positioned themselves courtside at side "B," proximate to the net and triangle line, to score each serve. The landing spot of the ball was documented for each trial, with no feedback provided to the participants. A retest was conducted using the same methodology after a seven-day interval. One novice participant and three from the expert group failed to attend the retest without explanation, resulting in the exclusion of their data from analysis.

Assessment: To gauge the test's relevance, clarity, and applicability, both experts and representatives of the target audience—coaches, instructors, and athletes—evaluated it using a Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). The Content Validity Coefficient (CVC) was calculated (26), with a value exceeding .80 considered satisfactory (15). The experts' responses were examined for face validity using the percentage agreement method. Descriptive statistics such as the mean and standard deviation were provided. The Shapiro-Wilk test confirmed data normality, with parametric inferential statistics applied thereafter. Independent t-tests were conducted to compare service accuracy differences between experienced and novice groups, assessing the test's discriminant validity. Hierarchical cluster analysis was performed using the Euclidean quadratic distance measure and the between-groups linkage approach, yielding two clusters via the single solution method. Chi-square tests were utilized to confirm group-cluster relationships. Intraclass correlation coefficients (ICC) and standard error of measurement were employed to assess test-retest reliability (25). All analyses were executed using SPSS for Windows (version 22.0), with a significance threshold set at p < .050.

3 RESULTS

Face and Content Validity: Initially, specialists evaluated the test's content validity. The original test underwent two evaluations, with amendments made to its explanations and scoring based on expert recommendations. Following these iterations, the final version was achieved. Experts unanimously agreed that the test effectively measures a volleyball player's accuracy,
demonstrating high levels of agreement for its applicability (CVC = .91), relevance (CVC = 1.0), and clarity (CVC = .91). Among the 60 experts assessing face validity, concordance rates were 81% for relevance, 88% for applicability, and 78% for clarity.

Sensitive Validity: The impact of experience levels, serving as the independent variable, on sensitive validity was assessed, revealing the test's accuracy in distinguishing service accuracy levels. The t-test revealed a strong effect size, with the experienced group outperforming the novice group in both initial and follow-up tests (6). Based on Cohen's d index, the level of experience exhibited a noteworthy influence on the accuracy of volleyball serves, as evidenced by both the initial test (d = 2.15) and the subsequent retest (d = 1.56).

Table 1

Comparative Analysis between Groups: Averages, deviations from the mean, p-values, and Cohen's measure of effect size (d).

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Cohen’s d</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>18</td>
<td>15.87</td>
<td>1.42</td>
<td>2.15#</td>
<td>0.001</td>
</tr>
<tr>
<td>Experienced</td>
<td>20</td>
<td>18.81</td>
<td>1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>16</td>
<td>16.02</td>
<td>1.85</td>
<td>1.56#</td>
<td>0.001</td>
</tr>
<tr>
<td>Experienced</td>
<td>18</td>
<td>18.13</td>
<td>1.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Cohen's d indicating a substantial effect size; p-values from Student's t-test.
Source: Prepared by Authors (2024)

Table 2 illustrates that most participants were organized according to the cluster analysis in a manner consistent with the original experience-based groupings. Ninety-nine percent of the cases in the cluster analysis based on experience level were accurately classified. Within the novice group, two individuals were assigned to cluster 2, while only one member of the expert group was assigned to cluster 1. There was a significant correlation between groups and clusters (Chi2 p=0.001).

Table 2

Reliability of Test-Retest.

<table>
<thead>
<tr>
<th>Groups</th>
<th>SEM</th>
<th>%SEM</th>
<th>CI(95%)</th>
<th>N</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>0.72</td>
<td>4.63</td>
<td>0.43-0.92</td>
<td>16</td>
<td>0.83</td>
</tr>
<tr>
<td>Experienced</td>
<td>0.47</td>
<td>2.64</td>
<td>0.51-0.94</td>
<td>18</td>
<td>0.85</td>
</tr>
<tr>
<td>Total</td>
<td>0.57</td>
<td>4.31</td>
<td>0.80-0.94</td>
<td>34</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Note: N = sample size; ICC = Intraclass Correlation Coefficient; CI = Confidence Interval; SEM = Standard Error of Measurement; % SEM = percentage of SEM in relation to the mean.
Source: Prepared by Authors (2024)

Reliability: The intraclass correlation coefficients, accompanying confidence intervals, and standard error of measurement are shown in Table 3. Test-retest reliability was noteworthy
for both the experienced and beginner groups. The consistency of test-retest was notably stronger for experienced individuals and the overall sample, although it remained acceptable for novices. This is evident as the confidence interval for temporal stability was narrower for the experienced groups and the overall sample compared to the novice groups.

Table 3

*illustrates the distribution of clusters according to experience level.*

<table>
<thead>
<tr>
<th></th>
<th>Grouping 1</th>
<th>Grouping 2</th>
<th>Total</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inexperienced</td>
<td>15(83.3%)</td>
<td>3(16.7%)</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Experienced</td>
<td>3(15.0%)</td>
<td>17(85.0%)</td>
<td>20</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Note. * The Pearson's Q-square value is 22.012. Eighty-five percent of the initially grouped cases were accurately categorized.

Source: Prepared by Authors (2024)

4 DISCUSSION AND CONCLUSIONS

This study explored the content, face, sensitive validity, and test-retest reliability of the volleyball overhand serve accuracy assessment. The results demonstrate that this accuracy test for serving is a feasible and reliable approach for monitoring player progress and investigating motor learning. All experts found the test's ability to evaluate correctness to be satisfactory, and its validity was affirmed with Content Validity Coefficients (CVCs) for utility, clarity, and significance exceeding acceptable thresholds (26; 15). Additionally, the expert panel validated the test's utility in providing accurate and clear guidance through relevant evaluation data. Sensitivity analysis, a crucial aspect of psychometric standards, also revealed the volleyball serve accuracy test's capability to differentiate performance levels across varying experience levels (18).

As evidenced in this study, it is anticipated that experienced individuals will exhibit superior performance in precision tasks such as volleyball serves and other sports skills. A plausible explanation lies in the enhanced ability of seasoned individuals to perceive, retain, and manipulate relevant environmental information (19). Furthermore, cluster analysis revealed consistent categorization of nearly every participant with their original group, and the accuracy of the serve test facilitated differentiation of participants' skill levels. Volleyball coaches have commonly utilized skill assessments to evaluate young players' performance in fundamental abilities (16). In terms of serve skill, accuracy testing has focused on game requirements, assigning higher scores in areas where receiving is more challenging (2; 3; 22). These evaluations are particularly relevant for monitoring athletes' performance and aligning with the
specific demands of the sport. Improvements in parameterization, including ball strength, speed, and direction, are crucial for motor learning. The current test fulfills the requirement for instruments assessing motor skill development in the context of sports to detect improvements in performance accuracy. It is noteworthy that while the Bartlett et al. (3) test differentiates athletes' performance accuracy during a volleyball match, it faces challenges in differentiating performance during early learning due to the one-centimeter difference between minimum and maximum scores, as previously mentioned. Hence, future research could explore the relationship between test sensitivity and experience level, as investigated in this study. For instance, the fact that elite and near-elite players perform similarly in terms of accuracy (16) may be due to the instrument's reduced sensitivity. The accuracy test proposed in this study not only demonstrates appropriate validity and reliability indices but also addresses these limitations.

The results of test-retest reliability were considered good for the entire sample and adequate for both groups (14). As per findings by Robertson et al (21), scant evidence exists for achieving commendable to outstanding levels of reliability, akin to those observed in this study, across various evaluations of sports motor skills performance. This is underscored by the fact that approximately 41% of the 22 studies included in the systematic review yielded comparable outcomes. The metrics obtained in this study surpass those of the NCSU assessment (3) and mirror closely the findings reported by Gabbett and Georgieff (9) in their proposal for evaluating the abilities of young volleyball players. However, the latter investigation omitted the use of a precise target; instead, it solely gauged whether the ball traversed to the court's opposite side to ascertain serve accuracy. Consequently, akin to the NCSU Volleyball Skills Test Battery (3), the assessment devised by Gabbett and Georgieff (9) carries the potential of overestimating novice performance while underestimating the performance of players exhibiting superior technical skills.

One strength of this study lies in the applicability of the test to a teaching-learning environment. The instrument serves as a suitable tool for assessing a wide range of performances, from novice to expert levels. A limitation of the current investigation was the relatively small sample size in terms of sensitive validity, which constrained the generalizability of the findings.

The accuracy test for volleyball serves exhibits considerable validity and test-retest reliability, rendering it a viable alternative to laboratory tasks in studies focusing on intricate motor skills. This avenue of study holds promise for yielding greater advantages in sports interventions compared to laboratory endeavors (24). Given the challenges associated with
crossing the court due to structural and functional constraints, we propose that testing be conducted in closer proximity to the net during the initial learning phases. Further research is warranted to explore the psychometrics of this approach. Additionally, it is recommended to delve into the criterion validity of the serve accuracy test and investigate the correlation between test accuracy and the quality of movement patterns in terms of psychometrics.

REFERENCES


