

Test-Retest Reliability and Interrater Objectivity of Throwing Accuracy and Velocity Assessments in Female Handball Players

Roopathi.M¹, Dr S.M Prakash²

¹Research Scholar, Department of PG Studies and Research in Physical Education, Sports and Yoga, Kuvempu University, Shivamogga, Karnataka, India.

²Research Supervisor and former Director, Department of PG Studies and Research in Physical Education, Sports and Yoga, Kuvempu University, Shivamogga, Karnataka, India.

Abstract—The assessment of skills is essential for talent identification and performance monitoring in handball. This study examined test-retest reliability and interrater objectivity of nine skill assessments in female handball players using Pearson product-moment correlation. Twenty female handball players (age range 18–24 years) completed nine skill tests including defensive movement (D' test), wall passes, ground pass, dribbling, and four throwing assessments on two separate occasions rated by two independent raters. Test-retest reliability coefficients ranged from $r = .76$ to $r = .93$, demonstrating moderate to high reliability across all assessments. Interrater objectivity coefficients ranged from $r = .74$ to $r = .88$, indicating acceptable to high consistency between raters. The 22-meter Ground Pass ($r = .93$) and 10-meter Wall Pass ($r = .91$) demonstrated superior reliability, while Face Shooting exhibited the lowest reliability ($r = .76$). Results support the implementation of these assessments for performance evaluation in female handball players.

Index Terms—test-retest reliability, interrater objectivity, skill assessment, handball, Pearson correlation, female handball players, batteries.

I. INTRODUCTION

Handball performance requires a blend of technical and tactical skills, making accurate evaluation of fundamental abilities critical for talent development [1]. Reliable and valid measurement tools are essential to distinguish performance differences and track athlete progress over time [2]. Test-retest reliability and interrater objectivity are key properties that indicate whether assessments yield consistent

results across repeated trials and evaluators [3]. While prior studies have examined reliability of individual sport-specific tests, comprehensive evaluation of multiple skills within the same athlete cohort remains limited [4]. The Pearson product-moment correlation coefficient is widely used to assess score consistency across sessions or raters [5]. This study therefore investigated the test-retest reliability and interrater objectivity of nine handball skill tests, including defensive movement, passing, dribbling, and throwing tasks, in female players, providing practical evidence for applied performance assessment.

II. METHODOLOGY

2.1 Participants and Design

Twenty female handball players representing university, national, and state-level teams, as well as clubs and colleges across Karnataka, participated in this test-retest reliability investigation (mean age = 20.5 ± 2.3 years; competitive experience ≥ 2 years). Player performances were assessed on two occasions separated by a seven-day interval by two independent evaluators—a certified strength and conditioning coach and an experienced handball coach. All study procedures received approval from the institutional ethics committee, and written informed consent was obtained from each participant prior to data collection [6].

2.2 Skill Assessment Battery

Defensive Movement Test	22-Meter Ground Pass	Jump and Shooting
30-Second Wall Pass	Zigzag Dribble	Penalty Shooting
10-Meter Wall Pass	Face Shooting	Pivot and Shoot

2.3 Statistical Analysis

Pearson product-moment correlation coefficients (r) were calculated between test sessions (reliability) and between raters on first session (objectivity). Descriptive statistics (M, SD) and reliability indices are reported in Table 1. Correlation interpretation: $r \geq .80$ = high, $r = .70-.79$ = moderate, $r < .70$ = low [7].

III. RESULTS

Comprehensive data for all nine assessments are presented in Table 1. Test-retest reliability coefficients ranged from $r = .76$ (Penalty Shooting) to $r = .93$ (22-meter Ground Pass), with eight of nine assessments demonstrating high reliability ($r \geq .80$). Interrater objectivity coefficients ranged from $r = .74$ (Penalty Shooting) to $r = .88$ (22-meter Ground Pass), indicating moderate to high consistency between raters.

Table 1: Descriptive Statistics and Pearson Correlation Coefficients for Skill Assessments (N = 20)

Assessment	Mean	Std. Deviation	Reliability (r)	Objectivity (r)	Interpretation
D' Defensive Movement	7.49	.80	.81	.82	High
30 sec Wall Pass	41.80	3.40	.88	.81	High
10 m Wall Pass	18.15	2.87	.91	.85	High
22 m Ground Pass	6.85	3.56	.93	.88	High
Zigzag Dribble	10.31	1.21	.92	.83	High
Face Shooting	4.75	1.94	.82	.80	High
Jump and Shooting	4.90	1.80	.85	.82	High
Penalty Shooting	5.45	1.50	.76	.74	Moderate
Pivot and Shoot	4.15	1.50	.85	.85	High

Table 1: Descriptive statistics and Pearson correlation coefficients demonstrating test-retest reliability and interrater objectivity.

Passing-based evaluations produced the strongest consistency outcomes, with the 22-meter ground pass recording the highest reliability coefficient ($r = .93$). Overall, passing skills demonstrated uniformly high reliability ($r = .88-.93$) and strong interrater

objectivity ($r = .81-.88$). The dribbling test also showed excellent test-retest reliability ($r = .92$) alongside moderate to high objectivity ($r = .83$). Assessment of defensive movement yielded acceptable reliability ($r = .81$) and objectivity ($r = .82$). For shooting tasks, both jump shooting and pivot shooting displayed high reliability values ($r = .85$), whereas penalty shooting showed comparatively

lower reliability ($r = .76$) and objectivity ($r = .74$), although these values remained within acceptable limits for performance assessment applications [8].

IV. DISCUSSION

Assessment of nine handball-specific skill tests demonstrated moderate to high reliability and interrater objectivity across diverse motor tasks, supporting their application in performance evaluation and talent identification for female athletes. Passing tests showed particularly strong reliability ($r = .81-.93$) and objectivity ($r = .81-.88$), likely due to more standardized execution compared with dynamic throwing tasks. Throwing assessments exhibited acceptable reliability ($r = .76-.85$) and objectivity ($r = .74-.85$), despite increased complexity. High interrater agreement reflects effective evaluator training, while slightly lower values for penalty shooting may result from individual technique or psychological factors. Overall, these tests provide a comprehensive technical performance profile.

V. CONCLUSION

The test-retest reliability and interrater objectivity of nine handball skill assessments were evaluated in female players using Pearson product-moment correlation analysis. Results showed moderate to high reliability for all tests, with eight out of nine meeting high reliability standards. Interrater objectivity also demonstrated acceptable to high agreement between independent evaluators. These outcomes support the practical and research use of this comprehensive assessment battery for accurately measuring technical skill performance in female handball Players.

REFERENCES

[1] Srhoj, V., Milanovic, D., & Jukic, I. (2006). Relationships between general anthropometric and motor variables and technical skills in handball. *Collegium Antropologicum*, 30(3), 551–558.

[2] Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of*

Chiropractic Medicine, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

[3] Portney, L. G., & Watkins, M. P. (2008). *Foundations of clinical research: Applications to practice* (3rd ed.). Prentice Hall.

[4] Hafelinger, U. (2018). Performance testing and evaluation of handball players. In T. Reilly & F. M. Drust (Eds.), *Science and football VI* (pp. 201–210).

[5] Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Philosophical Transactions of the Royal Society*, 187(A), 253–318.

[6] Frstrup, B., Larsen, S., & Petermann-Rocha, F. (2024). Test-retest reliability of lower limb muscle strength, jump and sprint performance tests in elite female team handball players. *European Journal of Applied Physiology*. Advance online publication. <https://doi.org/10.1007/s00421-024-05488-8>

[7] Jarnig, G., Kerbl, R., & van Poppel, M. N. M. (2024). Reliability, objectivity, validity, and reference levels of the Austrian pole climbing test (APCT)—A novel monitoring tool for assessing general fitness in children, adolescents, and young adults. *Sports*, 12(9), 258.

[8] Nagorka, R., & Chmura, J. (2024). Building confidence in motor skill testing for youth handball players: Validity and reliability analysis. *Sports, Activity, and Yoga Quarterly*, 13, 775–786.