ORIGINAL SCIENTIFIC PAPER

Construction and validation of a Beach Volleyball float service accuracy test

Marina Tsakiri[1], Manolis Adamakis[1], Eugenia Gortsila[1], Karolina Barzouka[1], Nikolaos Geladas[1], Konstantinos Sotiropoulos[1]

Affiliations: [1]National & Kapodistrian University of Athens, Greece

Correspondence: Marina Tsakiri. National & Kapodistrian University of Athens, Greece, NKUA, School of Physical Education & Sports Science, Ethnikis Antistasis 41, 17237 Dafni, Athens - Greece. E-mail: marinatsak@phed.uoa.gr

Abstract

The purpose of this study is to develop a specialized service test for beach volleyball and to evaluate reliability and construct validity. Fifty athletes (31 females and 19 males) participated in a test-retest protocol. Service velocity was analyzed using one-way repeated measures ANOVA and the non-parametric Friedman test within a gender-stratified framework, while accuracy scores across four target zones (Z1 to Z4) were evaluated using the Friedman test. Relative reliability was assessed via Intraclass Correlation Coefficient (ICC) and Spearman's rho, while absolute reliability was quantified through the Standard Error of Measurement (SEM) and the Minimal Detectable Change (MDC95). The results confirmed the protocol's construct validity, revealing a clear hierarchical progression of spatial difficulty. Zone 4 was the most challenging, yielding the lowest accuracy, followed by Z3, while Z1 and Z2 were the easiest (p<.001). Service velocity exhibited moderate-to-good relative reliability, while gender stratification uncovered distinct absolute reliability benchmarks, with females demonstrating lower intra-individual variability (MDC95: .25 to .55) and learning adaptations compared to males (MDC95: .36 to .69). Paradoxically, the results did not support a traditional speed-accuracy trade-off; athletes recorded their highest mean velocity in the most difficult zone (Z4) and lowest in Z2, indicating a risk-taking behavior under high spatial constraints. In conclusion, this custom-made protocol is a valid and reliable sport-specific instrument. It effectively bridges the gap between research and practice, providing coaches with an objective tool to monitor technical proficiency, simulate high-pressure decision-making and accurately quantify athlete progression on the sand.

Keywords: *Performance analysis, speed-accuracy relationship, target accuracy, construct validity*

@MJSSMontenegro
BEACH VOLLEYBALL FLOAT SERVICE TEST
http://mjssm.me/?sekcija=article&artid=330

Cite this article: Tsakiri, M., Adamakis, M., Gortsila, E., Barzouka, K., Geladas, N., Sotiropoulos, K. (2027) Construction and validation of a Beach Volleyball float service accuracy test. Montenegrin Journal of Sports Science and Medicine, 23 (1), Ahead of print. https://doi.org/10.26773/mjssm.270105

Introduction

Beach volleyball is a high intensity intermittent team sport (Kiraly and Shewman, 1999), requiring technical precision, tactical intelligence, team cohesion and physical endurance (Palao et al., 2015; Natali et al., 2018) under demanding environmental constraints. Among the six fundamental technical skills (service, reception, set, attack, block, and dig), the service is unique as it is the only "closed skill" -executed in

Received: 15 May 2026 | Accepted after revision: 09 September 2026 | Early access publication date: 01 September 2026 | Final publication date: 15 March 2026

Conflict of interest: None declared.

a self-paced manner free from immediate opponent interference (Homberg & Papageorgiou, 1995; Costa et al., 2024). Modern service has evolved from a simple game-initiation tool into an aggressive offensive weapon designed to score direct points (aces) or systematically disrupt the opponent's offensive structure (Wise, 2002; Altundag et al., 2024). Winning sets at competitive levels heavily correlates with serving efficacy and speed (Zetou et al., 2007; Marelic et al., 2004; Moras et al., 2008). Because defensive side-out transitions are highly optimized in elite circles, generating break points during one's own service is paramount breaking match statements (Laios & Kountouris, 2005). Parameters judging service speed and placement efficiency are critical indicators used by developmental and youth coaches to evaluate technical proficiency and future performance potential (Lidor et al., 2007; Bari et al., 2023).

Historically, the evaluation of service proficiency has relied on standardized skill batteries using target-based court layouts and multi-trial protocols to secure statistical power and reliability (Alnedral et al., 2020; Bartlett et al., 1991; Strand & Wilson 1993). Across the existing literature, psychometric integrity in these testing instruments has traditionally been established through distinct methodological validation procedures. Content and face validity have been confirmed using quantitative expert judgement indices such as Content Validity Index (CVI) and Aiken's V coefficient (Costa et al., 2024; Jariono et al., 2023), structural construct validity via factor analysis (Alnedral et al., 2020) and known-groups discriminant validity by differentiating performance across expertise levels or exertion conditions (French & Thomas 1987; Gabbett & Georgieff, 2006; Lidor et al., 2007). Regarding test stability, reliability has been primarily demonstrated through test-retest correlation coefficients (r), Intraclass Correlation Coefficients (ICC) using two-way mixed models, non-parametric Spearman's rank correlations (ρ) and inter-rater agreement indices (Costa et al., 2024; Giatsis, 2003; Valadares et al., 2018). Furthermore, contemporary target designs have evolved from historical parallel linear scoring zones (Bassett et al., 1937) toward tactical spatial difficulty maps (Costa et al., 2024) and high-precision X-Y coordinate tracking grids (Bieleke et al., 2019).

Despite these methodological advancements in indoor volleyball (Alnedral et al., 2020; Bieleke et al., 2019; Costa et al., 2024), existing testing batteries fail to account for the unique environmental and tactical dynamics of beach volleyball. Specifically, traditional protocols neglect sport-specific constraints such as smaller court dimensions (8×8m), unstable sand surface and aerodynamic-wind distortions, factors that fundamentally alter serving mechanics, ball velocity requirements and tactical execution (Häyrinen & Tampouratzis, 2012; Jiménez-Olmedo et al., 2012). While generalized technical batteries have occasionally been applied (Giatsis, 2003), there remains a complete absence of standardized validated instrument specifically calibrated to evaluate float service performance – integrating both accuracy and serving velocity – under beach volleyball conditions. This represents a critical gap, given that service velocity is a primary determinant of offensive efficacy and Float Serve (FS) accounts for up to 22% of serves in beach volleyball (Häyrinen & Tampouratzis, 2012; Jiménez-Olmedo et al., 2012). Therefore, a specialized valid testing framework is required to provide coaches with an objective tool for evaluating float service technical proficiency and monitoring performance adaptations on sand.

Materials & Methods

Participants and design

Fifty athletes (31 females, 19 males) volunteered for the study. Individual playing experience varied from recreational to competitive levels. Anthropometrics and experience characteristics are outlined in Table 1. All participants (and legal guardians for minors) signed informed consent forms. Procedures were institutional review board approved (NKUA: 1650/12-06-2024) and conducted in alignment with established ethical principles.

Table 1. Descriptive statistics and demographic characteristics of participants by gender (N=50).

Descriptive Statistics						
Gender		N	Min	Max	Mean	Std. Deviation
Male	Height (cm)	19	165.30	192.30	177.76	7.74
	Weight (kg)	19	57.80	93.80	79.45	9.43
	BMI	19	20.82	28.44	24.95	2.29
	Age (years)	19	15.00	49.00	31.68	9.38
	Experience (years)	19	1.00	23.00	8.53	6.77
Female	Height (cm)	31	154.00	188.00	166.14	7.11
	Weight (kg)	31	49.80	81.60	64.14	7.58
	BMI	31	18.23	28.55	23.35	2.59
	Age (years)	31	15.00	46.00	26.48	7.15
	Experience (years)	31	0.50	23.00	10.19	6.29

BMI=Body Mass Index

Service test

A custom, portable perforated sheet was secured over the sand court to delineate specific targeting areas matching official dimensions, as seen on Figure 1. Zones were categorized from 1 (easiest) to 4 (most difficult).

The spatial configuration was established based on expert panels (three very experienced elite coaches and university faculty specialists), studies with similar tests (Häyrinen & Tampouratzis, 2012) and official FIVB (Fédération Internationale de Volleyball) reception metrics (Casanova, 2015). Grounded in data showing that the majority of elite ball receptions occur in the center of the half-court ("reception hot zone"), Zone

DOI 10.26773/mjssm.270305

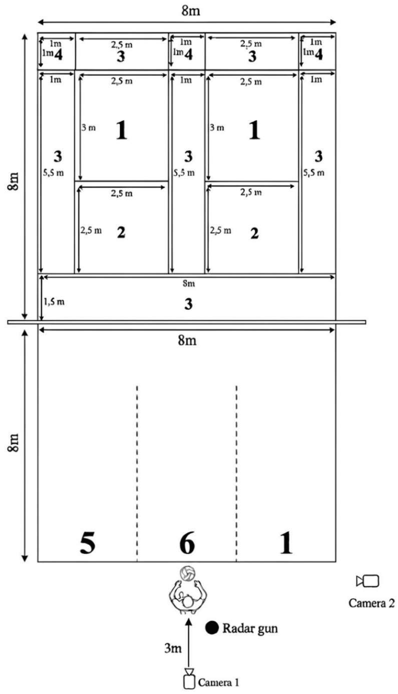

Figure 1. Target areas and dimensions (up), service zones (down) and equipment (cameras & radar gun) set-up.

1 (Z1) was placed exactly over this primary reception base (≈6m from the net, central), where the initial reception zone for the players is located, while Zone 2 (Z2) directly preceded it. Peripheral boundaries, deep baselines and the inter-player seam represented Zone 3 (Z3; moderate difficulty). Zones 4 (Z4) targeted deep, high-constraint boundary corners where trajectory judgement is heavily compromised.

Procedure

Following a supervised, standardized motor-skill warm-up, athletes executed 20 standing FS individually. Participants selected their preferred execution base among standard serving positions (1, 6 or 5). The protocol required 5 successive trials targeting Z1, followed sequentially by 5 trials each of Z2, Z3 and Z4 (20 total). A measurement assistant positioned behind the server logged ball velocity using a validated (Makar et al., 2024) Bushnell radar gun (speed range: 15-180km/h; accuracy: ±2km/h; frequency: 24.115GHz). Two synchronized digital cameras (Sony Handycam-DCR-SR32E) captured sagittal and posterior motion planes to verify target acquisition and technical conformity (Figure 1). Serving faults (net touches, out-of-bounds) were logged based on official FIVB rules. To analyze temporal stability, the identical testing structure was replicated after one-hour rest window.

Statistics

Normality was tested via the Shapiro-Wilk metric (N=50), validating parametric models for service velocity data and non-parametric structures for accuracy distributions for total sample and for the gender-stratified breakdown (males=19 & females=31). Velocity difficulty variance was assessed for the total sample and for females via one-way repeated measures ANOVA, with Greenhouse-Geisser corrections on total sample (e=.715) and Friedman test for males. For the female group Mauchly's test verified that the assumption of sphericity was satisfied (p=.052, e=.809). Accuracy variances across zones were mapped via the Friedman test for the total sample and both gender subgroups, supplemented by post-hoc pairwise Bonferroni adjustments to evaluate the spatial difficulty hierarchy. Relative reliability was quantified via two-way mixed-effects absolute agreement single-measures ICC for velocity variables maintaining normal distributions (total sample and females), while Spearman's rank correlation (ρ) was utilized for non-normal velocity outputs (males) and all target accuracy datasets.

Systemic bias was tracked via the Wilcoxon signed-rank test. Absolute reliability parameters were defined by the Standard Error of Measurement ($SEM=SD\times\sqrt{(1\text{-reliability coefficient})}$) and the Minimal Detectable Change ($MDC95=SEM\times1.96\times\sqrt{2}$) (Haley & Fragala-Pinkham, 2006; Weir, 2005). Significance was set at $p<.05$ via IBM SPSS Statistics v30.0.

Results

Descriptive assessments revealed that 62.0% (68.4% of male & 58.1% of female) of the services were executed from zone 1, 6% (5.3% of male & 6.5% of female) from zone 5 and 32% (26.3% of male & 35.5% of female) from zone 6. Normality markers verified normal data profiles for all velocity outputs ($p>.05$), contrasting with non-normal accuracy distributions across testing increments ($p<.001$).

Repeated measures ANOVA demonstrated a significant main effect for target zone on service velocity during both testing sessions for the total sample (pre: $F(3, 147)=277.50$, $p<.001$, $\eta^2=.85$; post: $F(3, 147)=296.84$, $p<.001$, $\eta^2=.86$). Gender-stratified analysis revealed that for women, all service velocities maintained normality ($p>.05$) and a one-way ANOVA confirmed a highly significant main effect across zones ($F(3, 90)=138.56$, $p<.001$, e=.809). Bonferroni-adjusted post-hoc pairwise comparisons for the female cohort quantified these velocity trajectories, revealing that all target zones differed significantly from one another ($p<.001$). Zone 4 prompted the highest execution velocities, scaling significantly all other zones, followed sequentially by Z1 and Z3, while Z2 elicited the baseline minimum. This established a strict, robust and statistically significant velocity hierarchy mapped as Z2<Z3<Z1<Z4 (p<.001 across all pairs).

Conversely, men demonstrated a localized normality violation in velocity of Z3 (p=.024). A non-parametric Friedman test applied to the male cohort validated significant shifts across velocity zones both in the first ($\chi2(3)$=48.91, p<.001) and the second measurement ($\chi2(3)$=43.04, p<.001).

Bonferroni adjustment for the total sample showed that Z2 prompted significantly lower velocities compared to all other zones (p<.001), whereas Z4 recorded the highest velocity marks (p<.001). No velocity variance emerged between Z1 and Z3 (p=.085). Single-measures ICC indexes proved statistically significant (p<.001) for total sample, revealing moderate-to-good relative velocity consistency across all zones (Table 2). For male cohort, relative reliability tracked via Spearman's coefficients (ρ) confirmed high stability in Z1 (ρ=.649, p=.003) and Z4 (ρ=.610, p=.006).

Table 2. Pre & post - test service velocity and test-retest reliability across target zones stratified by gender (N=50)

Target Zone	Cohort	Pre (km/h)	Post (km/h)	ICC (95%)	Spearman's ρ
	Total (N=50)	46.38 ± 5.92	46.93 ± 5.86	.703 (.530–.820)	-
Zone 1	Males (n=19)	47.38 ± 5.43	47.90 ± 5.76	-	.649**
	Females (n=31)	45.76 ± 6.20	46.33 ± 5.94	.708 (.479–.849)	-
	Total (N=50)	36.51 ± 6.22	37.06 ± 6.14	.712 (.538–.827)	-
Zone 2	Males (n=19)	36.31 ± 5.56	36.90 ± 5.27	-	.123
	Females (n=31)	36.63 ± 6.69	37.16 ± 6.71	.853 (.719–.926)	-
	Total (N=50)	45.38 ± 5.82	45.75 ± 5.42	.790 (.659–.875)	-
Zone 3	Males (n=19)	46.40 ± 5.41	47.28 ± 4.96	-	.215
	Females (n=31)	44.75 ± 6.06	44.82 ± 5.53	.681 (.440–.833)	-
	Total (N=50)	49.68 ± 6.21	50.46 ± 6.56	.790 (.649–.877)	-
Zone 4	Males (n=19)	51.65 ± 5.42	51.78 ± 6.51	-	.610**
	Females (n=31)	48.47 ± 6.46	49.65 ± 6.54	.729 (.511–.861)	-

Data for service velocity are presented as Mean ± Standard Deviation. Due to deviations from normality in the male subgroup, test-retest reliability was evaluated using Spearman's rank correlation coefficient (ρ) for males (n=19), whereas Intraclass Correlation Coefficients (ICC, two-way mixed effects, absolute agreement) with 95% Confidence Intervals (95% CI) were reported for the total sample (N=50) and female group (n=31); *p<.05, **p<.01

The Friedman established a highly significant main effect for target zones on accuracy aiming for the total sample ($\chi2(3)$=85.05, p<.001), males ($\chi2(3)$=36.67, p<.001) and females ($\chi2(3)$=54.28, p<.001), validating a definitive difficulty hierarchy and confirming that the experimental zones were accurately structured from the easiest to the most difficult target. For all cohorts, Z4 was confirmed as the most challenging target area, Z3 acted as intermediate, while Z1 & Z2 represented the easiest spatial challenges as seen in Table 4.

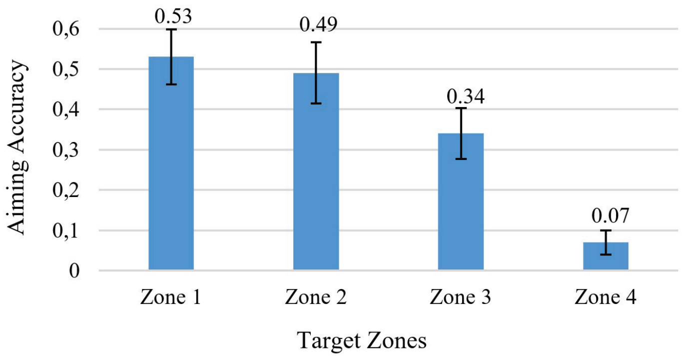

Figure 2. Aiming accuracy across zones (Zones 1-4) for the total sample (N=50). Bars represent mean aiming accuracy rates. Error bars denote 95% confidence intervals (95% CI). Exact mean values with 95% CI were 0.53 (0.46, 0.60) for Zone 1, 0.49 (0.41, 0.56) for Zone 2, 0.34 (0.28, 0.40) for Zone 3, and 0.07 (0.04, 0.10) for Zone 4.

Bonferroni post-hoc test demonstrated significant accuracy deltas between Z4 and all other zones (p<.001). Z1 and Z2 displayed no significant differences across all groups (p=1.000), confirming shared difficulty dynamics at the baseline of the spatial hierarchy.

To track systemic bias, Wilcoxon signed-rank tests evaluated pre-to-post accuracy changes. The total sample revealed no significant systematic deltas between test and retest sessions for Z1 (p=.175), Z3 (p=.140), and Z4 (p=.707). However, the gender stratified breakdown unveiled localized adaptations. For males, a significant test-retest accuracy increase emerged exclusively in Z1 (Z=-2.012, p=.044). For females, profound systemic adjustments were evident, as significant precision increases occurred across three target zones: Z1

 DOI 10.26773/mjssm.270305

Table 3. Reliability and stability indices for target accuracy.

Target Zone	Cohort	Systemic Bias (Wilcoxon Z)	Accuracy Spearman's ρ	Accuracy SEM	Accuracy MDC95
Zone 1	Total (N=50)	-1.357	.720**	0.16	0.51
	Male (n=19)	-2.012*	.584**	0.21	0.58
	Female (n=31)	-2.607**	.764**	0.11	0.31
Zone 2	Total (N=50)	-2.234*	.751**	0.21	0.58
	Male (n=19)	-0.584	.753**	0.22	0.61
	Female (n=31)	-2.463*	.761**	0.20	0.55
Zone 3	Total (N=50)	-1.474	.791**	0.21	0.59
	Male (n=19)	-0.835	.794**	0.25	0.69
	Female (n=31)	-2.308*	.771**	0.17	0.47
Zone 4	Total (N=50)	-0.375	.802**	0.11	0.30
	Male (n=19)	-0.065	.744**	0.13	0.36
	Female (n=31)	-0.407	.835**	0.09	0.25

SEM = Standard error mean, MDC95 = Minimal detectable change (Confidence Interval 95%); *p<.05, **p<.001

(Z=-2.607, p=.009), Z2 (Z=-2.463, p=.014) and Z3 (Z=-2.308, p=.021) Absolute and relative reliability coefficients are presented in Table 3.

Friedman testing mapping velocity against spatial difficulty constraints revealed a clear breakdown of the traditional speed-accuracy trade-off across the entire pool ($\chi2(3)=114.944$, $p<.001$), a behavioral trend closely paralleled within both male and female subgroups (Table 4).

Table 4. Friedman's test results on accuracy and velocity.

Zone	Velocity Mean Rank	Accuracy Mean Rank	Result
Zone 1	Male=2.50	Male=3.39	High velocity - high accuracy
	Female=2.94	Female=3.18	
Zone 2	Male=1,05	Male=3.18	Lowest velocity - high accuracy
	Female=1.03	Female=3.11	
Zone 3	Male=2.47	Male=2.11	High velocity - moderate accuracy
	Female=2.39	Female=2.52	
Zone 4	Male=3.97	Male=1.32	Highest velocity - minimum accuracy
	Female=3.65	Female=1.19	

Discussion

The primary objective of this research was to establish the reliability and construct validity of a custom-engineered beach volleyball service test. In sports testing, construct validity evaluates the degree to which an instrument measures the targeted underlying theoretical trait – defined here as spatial service precision under constraints – led demands. This construct encompasses an athlete's capacity to adjust motor control, trajectory dynamics and velocity when facing special targets under sand and environmental constraints. The analytical matrix validates the framework's construct validity. Rather than relying solely on arbitrary spatial boundaries, the protocol demonstrated a systematic, statistically verified degradation in aiming accuracy as spatial target constraints intensified (Z1, Z2>Z3>Z4; $\chi2(3)=85.05$, $p<.001$). This linear performance degradation conforms to the constraints-led paradigm (Davids et al., 2008), confirming that contracting spatial parameters and increasing court-boundary risks systematically stress the underlying motor control system. Combined with the protocol's ability to capture zone-specific velocity adjustments and tactical risk-taking behaviors, this cross-cohort sensitivity explicitly validates that the test measures true spatial motor control within beach volleyball constraints.

Crucially, the Friedman test across the total sample, males and females explicitly confirm the step-by-step difficulty progression (Z1, Z2, Z3, Z4). This cross-cohort congruence quantitatively demonstrates that the experimental target zones were accurately structured and calibrated from the easiest to the most difficult spatial challenge, establishing robust structural validity.

Regarding the ecological and tactical validity of the protocol, the structural alignment of the test parameters directly corresponds with the environmental and strategic demands of competitive beach volleyball. While traditional instruments like the AAHPERD Service test (1969) and the Volleyball Skills Test Battery (Bartlett et al., 1991) have long focused on indoor volleyball, there is distinct lack of specialized, validated tools for the unique constraints of beach volleyball. The selection of target zones in this study directly reflects the tactical realities of the sport. Interestingly, while float service is rarely used in indoor volleyball (≈3%), its prevalence is significantly higher in beach volleyball, ranging from 11.6% in males, to 22% in females (Häyrinen & Tampouratzis, 2012; Jiménez-Olmedo et al., 2012), due to flight distortions and trajectory inconsistencies that disrupt opposition reception. The findings showed that 62% of the services were executed from Z1, a departure from match-play observations, where Z6 often dominates (41%; Häyrinen & Tampouratzis, 2012), suggesting that under the specific constraints of this protocol, athletes prioritize different serving origins to achieve target accuracy.

This strategic adjustment during service execution heavily interacts with the spatial properties and mechanical output required for each target accuracy. Conversely, the lack of significant difference between Z1 and Z2 across both genders confirms a shared, highly accessible difficulty dynamic at the baseline of the spatial hierarchy. This is heavily attributed to the reduced target depth in Z2, which prompts athletes to adjust their motor output. Specifically, while Z1 permits a powerful execution with higher baseline velocities (≈47.3km/h for males and 45.7km/h for females), the baseline constraint of Z2 forces athletes to adopt a significantly more conservative, low-velocity approach (≈36.3km/h for males and 36.6km/h for females) to ensure drop-trajectory placement and avoid serving out of bounds. This mechanical adjustment explains why at this level of expertise, these two target options present an identical accuracy challenge under baseline conditions.

In terms of internal data behavior, the gender-stratified breakdown uncovered critical statistical and behavioral nuances that were previously masked within the aggregate data. The divergent statistical behavior of velocity distributions-where females exhibited robust normality, while males introduced a localized parametric violation at Z3 ($p=.024$)-is heavily linked to the smaller male sample size ($n=19$) interacting with distinct motor execution strategies or biomechanical modifications. More profoundly, the Bonferroni-adjusted post-hoc pairwise comparisons for the female cohort quantified these velocity trajectories with absolute precision, revealing that all target zones differed significantly from one another ($p<.001$). This establishes a strict, statistically absolute liner velocity hierarchy mapped as Z2<Z3<Z1<Z4, showcasing a highly structured and distinct tactical pacing strategy unique to female group.

Despite these distributional variations, both groups displayed an identical, atypical interaction between performance metrics, demonstrating a clear deviation from the traditional speed-accuracy trade-off across the entire pool ($\chi2(3)=114.944$, $p<.001$). Contrary to expectations and Fitts' law (Fitts, 1954), velocity was not reduced to enhance precision in the most difficult areas (Z4). Instead, athletes recorded their highest mean velocity in Z4. A possible explanation lies in the spatial location of these targets, since Z4 is located in the deep corners of the court, a higher trajectory or higher velocity is biomechanically necessary to reach the baseline. Nonetheless, this represents a gender-independent "risk-taking" behavior, where the athletes prioritize maximum power when they perceive the probability of spatial success to be low, aligning with elite performance patterns, where high velocity pressures the opposition even when placement suboptimal (Medeiros et al., 2014).

Crucially, tracking pre-to-post intervention parameters highlighted distinct adaptation profiles between genders. Relative reliability of service velocity revealed moderate-to-good ICC and Spearman stability markers across testing sessions. These parameters suggest that the custom protocol maintains a consistent ranking of participants despite the inherent variability of a complex athletic skill, aligning with the stringent reliability standards established by Zetou et al. (2005) for beach volleyball and Costa et al. (2024) for indoors service accuracy. The slight increase in terms of the target achievement found in the second measurement (post) across cohorts may be attributed to a minor learning effect or an optimization of the athletes' warm-up status, echoing the methodological insights of Lidor et al. (2007), who highlighted that elite and developmental athletes should undergo repeated trials of sensorimotor adaptation and strategy optimization.

However, the stratified analysis revealed that men maintained stable metrics with an isolated spatial accuracy improvement only in Z1. Conversely, women exhibited widespread systemic shifts, significantly increasing their accuracy across three target zones (Z1, Z2, Z3). This multi-zone accuracy optimization in women, combined with their superior pre-to-post relative reliability indices (velocity ICC up to .853; accuracy ρ up to .835), indicates a more uniform sensorimotor adaptation and strategy optimization compared to men (Lidor et al., 2007).

Regarding absolute reliability, the female cohort recorded significantly lower MDC95 values across all zones (.25 to .55) compared to men (.36 to .69). This demonstrates that male athletes possess a higher degree of intra-individual variability in aiming performance on sand, where slight technical inconsistencies lead to larger fluctuations in outcome. The lower MDC observed in Z4 across all cohorts (.30 for total pool, .36 for males and .25 for females) should be interpreted cautiously; it likely reflects a "floor effect" (McBee, 2010) due to the extreme difficulty of the target, rather than higher precision. Given that approximately 36% of error points in a match are attributed to service errors (Häyrinen & Tampouratzis, 2012), these gender-specific MDC benchmarks provide essential knowledge for coaches, establishing the exact boundaries an athlete must exceed to demonstrate true skill acquisition, rather than random performance noise.

The observed sex-specific disparities in reliability parameters – characterized by superior relative stability and consistently lower MDC95 values in females (.25 – .55) compared to males (.36 – .69) – warrant deeper biomechanical and methodological contextualization. Methodologically, the smaller male sample size ($n=19$ vs $n=31$) naturally increases susceptibility to sampling variability and statistical sensitivity (Hopkins, 2000). Biomechanically and tactically, however, these differences point toward distinct motor execution strategies on sand. Male players routinely execute services at higher absolute velocities (≈47.3 to 51.7km/h), where slight micro-adjustments in striking angle, wind-gust interactions or foot placement on the unstable sand surface amplify the outcome. In accordance with signal-noise theory (Harris & Wolpert, 1998; Schmidt et al., 1979), higher biomechanical power output inherently escalates neuromuscular variability, weakening correlation coefficients in specific zones. Conversely, female players displayed a highly structured, linear velocity hierarchy and more uniform sensorimotor adaptation, suggesting a prioritized focus on technical repeatability and spatial placement over maximal force production. Consequently, the female cohort achieved tighter clustering of trials around their mean performance, yielding superior measurement precision. Recognizing these sex-specific reliability profiles is critical for practitioners, as diagnostic sensitivity threshold tests must account for the higher inherent motor noise present in high-velocity male service dynamics (Atkinson & Nevill, 1998; Medeiros et al., 2014).

Lastly, a deeper understanding of service execution under variable environmental conditions, along with the development of valid, reliable testing instruments, will enhance both coaching practice and empirical research in beach volleyball.

DOI 10.26773/mjssm.270305

The findings of this study can help players and coaches find the "safe" aiming zones, while being under pressure in score or fatigue and work target accuracy in submaximal velocity, so players can refine their tactical decision-making. As serving accounts for 9% of total points scored and 39% of break-point situations (Häyrinen & Tampouratzis, 2012), this integration of scientific exploration with practical application underlines the study's value for the beach volleyball community, aiming to bridge the gap between research and practice (Altundag et al., 2024).

Practical Applications

The practical use of this study lies in the development of a specialized and validated assessment tool for beach volleyball service, addressing a significant gap in athletic coaching and performance analysis. Coaches and practitioners can use this testing protocol to accurately quantify athlete progression, leveraging the established MDC values to distinguish between genuine skill acquisition and random performance fluctuations. Furthermore, the findings regarding the "risk-taking" behavior in difficult zones and the absence of a traditional speed-accuracy trade-off provide a new empirical basis for tactical training. Rather than relying on generic technical cues, this test enables coaches to simulate high-pressure decision-making, helping players identify "safe" target zones and refine their serving strategy during critical break-point situations or periods of high fatigue. Ultimately, the implementation of this sport-specific instrument bridges the gap between scientific research and field application, offering an objective means to maximize service efficacy - a primary determinant of success in modern beach volleyball.

Limitations

This study is subject to several limitations that should be acknowledged. First, the sample was uneven in terms of gender distribution (31 females and 19 males) and included players of varying skill levels, ranging from recreational to semi-professional, which may limit the generalizability of the findings. Second, the analysis focused exclusively on the float serve, excluding other types (e.g., jump spin serve and jump float serve); therefore, the results may not fully represent the effectiveness of all serving techniques used in beach volleyball. Additionally, the testing procedure was conducted prior to training and without the presence of an opposing receiver, meaning that players performed serves without competitive pressure. Finally, the technical execution of the serves was not evaluated, which may have provided further insight into performance differences.

Conclusions

The present study confirms that the proposed beach volleyball service protocol is a valid and reliable tool for assessing service accuracy. The hierarchical structure of the test is validated by a significant decline in aiming accuracy from Zone 1 to Zone 4. While the protocol demonstrated moderate-to-good relative reliability, the calculated MDC values provide essential benchmarks for coaches to distinguish true skill acquisition from performance noise. Notably, the analysis revealed a departure from the typical speed-accuracy trade-off, as athletes adopted a 'risk-taking' strategy by maximizing service velocity in the most demanding zones. This behavioral shift suggests that under high spatial constraints, athletes prioritize power over precision as a compensatory tactical adjustment. Overall, this protocol offers a robust framework for monitoring service's technical proficiency and understanding tactical decision-making in beach volleyball athletes.

Conflicts of Interest

The authors declare no conflict of interest.

References

Ajemian, R., D'Ausilio, A., Moorman, H., & Bizzi, E. (2010). Why professional athletes need a prolonged period of warm-up and other peculiarities of human motor learning. *Journal of Motor Behavior*, *42*(6), 381–388. https://doi.org/10.1080/00222895.2010.528262

Alnedral, A., Jariono, G., & Rinaldi, R. (2020). Development of volleyball serve accuracy test instrument. *Journal of Physical Education and Sports*, *9*(3), 245–251.

Altundag, E., Soylu, C., & Akyildiz, Z. (2024). Multidimensional analysis of serving speed in volleyball players by position, sets, and league types: Interactions and statistical differences. *BMC Sports Science, Medicine and Rehabilitation*, *16*(1), Article 103. https://doi.org/10.1186/s13102-024-01031-z

American Alliance for Health, Physical Education, Recreation and Dance. (1969). *Volleyball skills test manual*. AAHPERD Publications.

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Bari, M. A., García-de-Alcaraz, A., & Peñas, C. L. (2023). Does serve speed matter in elite volleyball? Analysis of the relationship between serve speed, serve efficacy, and the opponent's reception. *Journal of Human Kinetics*, *88*, 221–231. https://doi.org/10.26773/jhk.2023.045

Bartlett, J., Smith, L., Davis, K., & Peel, J. (1991). Development of a valid volleyball skills test battery. *Journal of Physical Education, Recreation & Dance*, *62*(2), 19–21. https://doi.org/10.1080/07303084.1991.10606554

Bassett, G., Glassow, R., & Locke, M. (1937). Studies in testing volleyball skills. *Research Quarterly. American Physical Education Association*, *8*(4), 60–72. https://doi.org/10.1080/23267402.1937.10761851

Bieleke, M., Gowert, M., & Wolff, W. (2019). High-precision X-Y coordinate tracking grids in indoor target skill evaluation. *Frontiers in Psychology*, *10*, Article 1823. https://doi.org/10.3389/fpsyg.2019.01823

Casanova, R. M. (2015). *Beach volleyball*. Fédération Internationale de Volleyball (FIVB).

Costa, C., Valentini, N., Nascimento, E., & Ugrinowitsch, H. (2024). Validity and reliability of the volleyball serve accuracy-test. *Journal of Human Sport and Exercise*, *19*(2), 570–578. https://doi.org/10.14198/jhse.2024.192.12

Davids, K., Button, C., & Bennett, S. (2008). *Dynamics of skill acquisition: A constraints-led approach*. Human Kinetics.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. https://doi.org/10.1037/h0055392

French, K. E., & Thomas, J. R. (1987). The relation of knowledge development to children's basketball performance. *Journal of Sport Psychology*, *9*(1), 15–32. https://doi.org/10.1123/

jsp.9.1.15

Gabbett, T., & Georgieff, B. (2006). Physiological and anthropometric characteristics of Australian junior national, state, and novice volleyball players. *Journal of Strength and Conditioning Research*, *20*(1), 29–35. https://doi.org/10.1519/R-16814.1

Giatsis, G. (2003). Reliability and validity of beach volleyball skill test instruments in adolescent athletes. *Kinesiology*, *35*(2), 154–162.

Haley, S. M., & Fragala-Pinkham, M. A. (2006). Interpreting change scores of tests and measures used in physical therapy. *Physical Therapy*, *86*(5), 735–743. https://doi.org/10.1093/ptj/86.5.735

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*(6695), 780–784. https://doi.org/10.1038/29528

Häyrinen, M., & Tampouratzis, K. (2012). *Technical and tactical game analysis of elite female beach volleyball*. KIHU: Research Institute for Olympic Sports.

Homberg, S., & Papageorgiou, A. (1995). *Handbook for volleyball*. Meyer & Meyer Sport.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Jang, B., Rusdiana, A., & Sudin, S. (2022). Three-dimensional kinematical analysis of jump serve in volleyball: Muscle fatigue effects. *Annals of Applied Sport Science*, *10*(3), Article e1034. https://doi.org/10.52547/aassjournal.1034

Jariono, G., Alnedral, A., & Rinaldi, R. (2023). Content validity index (CVI) and Aiken's V analysis for sports testing instruments. *Journal of Human Sport and Exercise*, *18*(4), 812–821.

Jiménez-Olmedo, J. M., Penichet-Tomás, A., Saiz-Colomina, S., Martínez-Carbonell, J. A., & Jove-Tossi, M. A. (2012). Serve analysis of professional players in beach volleyball. *Journal of Human Sport and Exercise*, *7*(3), 706–713. https://doi.org/10.4100/jhse.2012.73.10

Kiraly, K., & Shewman, B. (1999). *Beach volleyball*. Human Kinetics.

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

Laios, A., & Kountouris, P. (2005). Association between selected skill factors and results in women's professional volleyball in Greece. *International Journal of Performance Analysis in Sport*, *5*(2), 1–5. https://doi.org/10.1080/24748668.2005.11868322

Lidor, R., Arnon, M., Hershko, Y., Maayan, G., & Falk, B. (2007). Accuracy in a volleyball service test in rested and physical exertion conditions in elite and near-elite adolescent players. *Journal of Strength and Conditioning Research*, *21*(3), 937–942. https://doi.org/10.1519/R-20236.1

Makar, P., Silva, A. F., Silva, R. M., Janusiak, M., Smoter, M., & Clemente, F. M. (2024). The agreement between Bushnell and Stalker radar guns for measuring ball speed in throwing and kicking. *Applied Sciences*, *14*(22), Article 10476. https://doi.org/10.3390/app142210476

Marelic, N., Reic, Z., & Kuric, P. (2004). Influence of some parameters of conditioning, morphological and psychological characteristics on the efficiency of volleyball players. *Kinesiology*, *36*(1), 79–87.

McBee, M. T. (2010). Modeling outcomes with floor or ceiling effects: An introduction for educational researchers. *Gifted Child Quarterly*, *54*(4), 316–320. https://doi.org/10.1177/0016986210379095

Medeiros, A. I. A., Palao, J. M., Marcelino, R., & Mesquita, I. (2014). Revisão sistemática sobre a performance desportiva no voleibol de praia a partir da análise do jogo. *Revista Brasileira de Cineantropometria e Desempenho Humano*, *16*(6), 698–708. https://doi.org/10.5007/1980-0037.2014v16n6p698

Moras, G., Buscà, B., Peña, J., Rodríguez, S., Vallejo, L., Tous-Fajardo, J., & Mujika, I. (2008). A objective means of measuring the speed of the serve in beach volleyball. *International Journal of Performance Analysis in Sport*, *8*(2), 9–21. https://doi.org/10.1080/24748668.2008.11868430

Natali, S., Ferioli, D., La Torre, A., & Bonato, M. (2018). Physical and technical demands of elite beach volleyball according to playing position and gender. *The Journal of Sports Medicine and Physical Fitness*, *59*(1), 48–55. https://doi.org/10.23736/S0022-4707.17.07972-5

Palao, J. M., López-Martínez, A. B., Valadés, D., & Ortega, E. (2015). Physical actions and work-rest time in women's beach volleyball. *International Journal of Performance Analysis in Sport*, *15*(1), 424–429. https://doi.org/10.1080/24748668.2015.11868803

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., & Quinn, J. T., Jr. (1979). Motor output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, *86*(5), 415–451. https://doi.org/10.1037/0033-295X.86.5.415

Strand, B. N., & Wilson, R. (1993). *Assessing sport skills*. Human Kinetics.

Valadares, N., Nascimento, E., & Ugrinowitsch, H. (2018). Inter-rater reliability and stability of motor skill evaluation tools in team sports. *Brazilian Journal of Motor Behavior*, *12*(2), 45–52.

Weir, J. P. (2005). Quantifying test-retest reliability using the Intraclass Correlation Coefficient and the SEM. *Journal of Strength and Conditioning Research*, *19*(1), 231–240. https://doi.org/10.1519/15184.1

Wise, M. (2002). Serving. In D. Shondell & C. Reynaud (Eds.), *The volleyball coaching bible* (pp. 163–175). Human Kinetics.

Zetou, E., Tzetzis, G., & Giatsis, G. (2005). *Validation and reliability of beach volleyball skill test instruments* (Ref. No. W-40-7553). Aristotle University of Thessaloniki.

Zetou, E., Moustakidis, A., Tsigilis, N., & Komninakidou, A. (2007). Determining the effectiveness of serve and reception in volleyball. *International Journal of Performance Analysis in Sport*, *7*(3), 1–12. https://doi.org/10.1080/24748668.2007.11868404