



SCIENTIFIC JOURNAL OF SPORT AND PERFORMANCE

VOLUME

5

ISSUE

3

July 2026

EDITED & PUBLISHED BY



ARD
ASOCIACIÓN ESPAÑOLA

ASOCIACIÓN ESPAÑOLA DE ANÁLISIS DEL RENDIMIENTO DEPORTIVO

A conceptual integration of physiological and biomechanical load in distance running

 Thomas Sørensen  . Greenmoor Running Institute. Denmark.

ABSTRACT

The distinction between physiological and biomechanical load has become increasingly salient in both scientific and coaching discourse on endurance running. While physiological load has long been operationalised through metrics such as heart rate, lactate, and training impulse (TRIMP), biomechanical load remains conceptually and empirically less defined. This paper examines the current debate surrounding the measurement of biomechanical stress and proposes an integrative framework that situates both forms of load within a neuromuscular systems perspective. Drawing on recent work in biomechanics (Edwards, 2018; Matijevich et al., 2019) and exercise physiology (Seiler, 2010; Jones & Burnley, 2009), it is argued that the two domains cannot be adequately understood in isolation. Rather, they must be conceptualised as interacting dimensions of the same adaptive process. Physiological load governs metabolic strain and recovery kinetics, whereas biomechanical load shapes the structural tolerance and efficiency of movement. The paper outlines the limitations of current modelling approaches, such as device-derived impact scores, and advances a conceptual synthesis, an integrated model, which views neuromuscular recruitment as the mediating process through which metabolic and mechanical stress co-evolve.

Keywords: Physiological load, Biomechanical load, Running economy, Neuromuscular coordination, Training load, Injury risk.

Cite this article as:

Sørensen, T. (2026). A conceptual integration of physiological and biomechanical load in distance running. *Scientific Journal of Sport and Performance*, 5(3), 355-363. <https://doi.org/10.55860/WQCA8559>



Corresponding author. Greenmoor Running Institute. Denmark.

E-mail: tsoerensen@me.com

Submitted for publication November 13, 2025.

Accepted for publication January 09, 2026.

Published February 21, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/WQCA8559>

INTRODUCTION

In the past decade, endurance training has seen a growing interest in distinguishing between physiological and biomechanical load as partially independent dimensions of stress. Physiological load refers broadly to the energetic demand placed upon the cardiovascular and metabolic systems during exercise, often quantified through measures such as oxygen uptake, heart-rate derived training impulse (TRIMP), or sessional ratings of perceived exertion (Calvert et al., 1976; Foster et al., 2001). Biomechanical load, by contrast, concerns the mechanical strain experienced by tissues (muscles, tendons, bones, and ligaments) as a consequence of repetitive force application (Edwards, 2018).

Despite conceptual clarity, the empirical delineation between these forms of load remains elusive. Physiological stress is readily measurable through indirect metabolic proxies; biomechanical stress, however, is distributed across complex structures, making it difficult to estimate from external metrics such as speed or ground-reaction force (Matijevich et al., 2019). Recent attempts to derive biomechanical indicators from wearable devices (e.g. Stryd's Impact Loading Rate or Garmin's Running Tolerance) have popularised the notion that such stress can be numerically expressed. Yet the scientific validity of these algorithms is limited, as the highest predictors of tissue injury arise not from impact forces per se, but from active muscle contractions and their associated internal stresses (Edwards, 2018; Matijevich et al., 2019).

This conceptual confusion has created a methodological asymmetry in the endurance community: physiological load is tracked with precision, while biomechanical load remains inferred. The consequences are significant. Many athletes modulate their training primarily through heart-rate and pace metrics, assuming that mechanical stress scales linearly with metabolic demand. However, the relationship between force magnitude and tissue damage is highly non-linear (Edwards, 2018). Doubling the internal load within the Achilles tendon or tibia may produce an exponential increase in micro-damage per stride, even when the metabolic cost appears modest. The present paper therefore aims to clarify this relationship by situating both load types within a systems framework that recognises neuromuscular recruitment as the integrative mediator.

PHYSIOLOGICAL LOAD: ENERGETICS AND SYSTEMIC STRESS

The study of physiological load originates in classical exercise physiology and has been refined through decades of work on energy systems, cardiovascular dynamics, and hormonal adaptation (Bassett & Howley, 2000). In endurance running, it refers to the metabolic cost incurred by sustaining submaximal or maximal oxygen consumption (VO_2) over time. The quantification of this load has evolved from simple heart-rate monitoring to complex composite indices such as TRIMP (Calvert et al., 1976), Edwards' modified TRIMP (1993), and the more recent Training Stress Score (Allen & Coggan, 2010), all of which attempt to link internal effort to external work.

Physiological load represents the stress applied to systems responsible for oxygen transport, substrate utilisation, and energy resynthesis. During running, oxygen consumption increases with speed until it reaches a plateau at maximal aerobic capacity (VO_{2max}). Although VO_{2max} is often cited as a primary determinant of endurance performance, evidence indicates that its ceiling is reached relatively early in an athlete's development, after which further improvement depends on submaximal efficiency and lactate kinetics (Joyner & Coyle, 2008; Seiler, 2010). Thus, while VO_{2max} defines the upper limit of aerobic metabolism, the capacity to sustain a high fraction of that maximum, so-called fractional utilization, is shaped by cumulative physiological load over time (Jones & Burnley, 2009).

The accumulation of physiological stress is essential to adaptation but also constrained by recovery capacity. Overexposure leads to maladaptive states such as non-functional overreaching or overtraining syndrome (Meeusen et al., 2013). Consequently, endurance programming seeks to balance the magnitude and frequency of load with adequate restitution. The polarised training model (Seiler & Tønnessen, 2009) has gained prominence for its efficiency in managing this balance: approximately 80% of volume is performed at low intensities (Zone 1–2), while the remaining 20% targets thresholds and $\text{VO}_{2\text{max}}$ stimuli. This distribution minimises chronic sympathetic stress while preserving high-quality stimuli for central adaptations.

However, these models remain predominantly metabolic. They quantify cardiovascular strain but say little about the structural stresses that accompany running. For example, two athletes may record identical TRIMP scores for a session, yet one may accumulate far more tissue load due to differences in body mass, stride mechanics, or surface hardness. Moreover, traditional physiological models assume a quasi-linear relationship between intensity and total load, whereas musculoskeletal damage increases exponentially with speed and ground-reaction forces (Matijevich et al., 2019).

Recent evidence has further complicated this picture by demonstrating that the metabolic and mechanical components of training do not always co-vary. Studies using inverse dynamics and musculoskeletal modelling (Farris & Sawicki, 2012; Vernillo et al., 2017) show that small increases in running speed disproportionately elevate the force-time integral on the lower limb, especially in the Achilles tendon and tibia. These increases occur even when the change in heart rate or oxygen uptake is minimal. Consequently, monitoring physiological load alone risks underestimating the true stress experienced by the tissues.

At the same time, the adaptive processes governing physiological load cannot be isolated from mechanical context. Capillary proliferation, mitochondrial biogenesis, and enzymatic adaptation all depend upon repeated muscle contractions that produce both metabolic and mechanical stimuli. In this sense, physiological and biomechanical load are not parallel phenomena but two expressions of the same neuromuscular act. This recognition provides the conceptual bridge toward an integrated model of load, developed further in the next section of this paper, in which the neuromuscular recruitment system is positioned as the critical interface between metabolic and structural adaptation.

FROM COMPETING LOADS TO AN INTEGRATED PRACTICE

The contemporary discourse on training stress in distance running has crystallised around two ostensibly incompatible vantage points. On the one hand stand physiological accounts that render training legible through global energetic markers: heart-rate dynamics, oxygen flux, lactate turn-points, and their various derivatives (e.g., TRIMP, TSS). On the other, biomechanical accounts insist that injury risk and durability are better understood as the history of forces borne by tissues with finite fatigue lives. Each, when pursued to its limit, becomes reductionist. The former struggles to explain why identical “load” as indexed by heart-rate-based metrics can precipitate divergent injury outcomes. The latter cannot, on its own, specify how an identical pattern of tissue forces can elicit radically different metabolic adaptations depending on the surrounding milieu of recovery and systemic stress. The practical task, therefore, is not to pick a winner but to elucidate the higher-order unity in which both perspectives become moments of a single training logic.

LIMITS OF SCALAR PHYSIOLOGY

The attraction of scalar physiological metrics is obvious: they compress heterogeneous experience into a single number that appears amenable to arithmetic periodisation (Calvert et al., 1976; Allen & Coggan, 2010).

Yet such scalars presuppose smooth mappings from internal energetic state to performance and from cumulative arithmetic to adaptive trajectory. Two problems follow.

First, equifinality: distinct session architectures can yield comparable integrated heart-rate or lactate profiles while imposing dissimilar neuromuscular and coordinative demands. A 40-minute continuous sub-threshold run and a session of short cruise intervals separated by extensive recovery may tally similarly in TRIMP, yet the latter preserves stride length and stiffness under repeated accelerations; the former rehearses economical monotony. Second, opacity to local damage: physiology-first metrics track systemic strain but remain insensitive to the distribution of peak forces per step that governs micro-damage accrual in load-bearing tissues (Edwards, 2018). Consequently, they are weak predictors of overuse injury despite their utility for gauging global fatigue.

BIOMECHANICAL INSISTENCE AND ITS ASYMMETRIES

Biomechanical treatments rectify this opacity by tracing how internal forces accumulate across steps, runs, and weeks. Crucially, the relevant loads are internal (e.g., tendon tensile force, tibial bending moment), not simply external surrogates such as ground-reaction force peaks or impact transients. Evidence increasingly demonstrates that common external proxies correlate only modestly with the internal stresses that drive damage (Matijevich et al., 2019). At the same time, tendon- and joint-level analyses show that substantial mechanical work can occur within the musculotendinous system with only subtle changes in external kinematics (Zelik & Franz, 2017), and that joint-level mechanical work can serve as an indirect measure of the energetic and coordination costs that shape movement strategies (Zelik & Kuo, 2012). Two asymmetries then become decisive.

The relationship between peak force and damage is highly nonlinear: small increases in peak or rate of loading can produce orders-of-magnitude increases in fatigue damage per cycle, particularly in bone and tendon (Edwards, 2018). Tissues adapt at different tempos and with different transfer properties. Muscle remodels quickly, tendon slowly (Kjaer et al., 2009; Magnusson et al., 2007); cortical bone responds to strain stimulus with a delayed and threshold osteogenic effect (Hughes et al., 2020). A training week that marginally elevates internal Achilles loads may be metabolically innocuous yet materially hazardous if repeated without allowance for tendon remodelling.

FATIGUE, FORM, AND THE ECOLOGY OF COORDINATION

A common rejoinder is that fatigue alters mechanics, so longer repeats (e.g., 4 × 2000 m at 95% 5K) must, *ceteris paribus*, be more injurious than shorter ones (10 × 800 m) of equal total fast volume. The empirical picture is subtler. Changes in running kinematics under fatigue are typically modest, individual-specific, and not uniformly deleterious; some are plausibly protective (e.g., slightly increased cadence reducing per-step vertical excursion) (van Hooren et al., 2024). Moreover, the time actually spent in a highly maladaptive mechanical state during longer repetitions may be a small fraction of total work, particularly in trained runners whose coordinative patterns are stable. What does follow is not a universal ranking of set-shapes by danger, but a methodological caution: when progressing training, increasing intensity or duration changes damage far more than scalar physiology implies, and those changes are mediated by how coordination adapts under speed-and-fatigue constraints.

AN INTEGRATED CONCEPT OF LOAD

The foregoing suggests that “load” is neither a single scalar (physiology) nor an exhaustive vector of internal forces (biomechanics). Rather, load is a structured composite whose meaning is only realised in the interaction among three moments:

1. Metabolic moment: the global energetic signal that conditions mitochondrial biogenesis, capillary proliferation, and substrate use (Calvert et al., 1976; Allen & Coggan, 2010).
2. Mechanical moment: the spatiotemporal distribution of internal tissue forces across steps that governs fatigue damage and remodelling thresholds (Edwards, 2018; Matijevich et al., 2019).
3. Coordinative moment: the stability and plasticity of the motor solution under task constraints (speed, gradient, surface, footwear), which in turn modulates both metabolic cost and mechanical loading pathway (van Hooren et al., 2024).

Training planning, then, becomes the art of co-shaping these moments such that each supports the others: metabolic development proceeds under mechanical distributions that stay within safe fatigue budgets, and coordinative practice improves the runner’s ability to express economy and preserve form under race-specific constraints.

This three-moment model can be read as a neuromuscularly-focused reformulation of existing load capacity frameworks (Gabbett, 2016; Bertelsen et al., 2017), in which structure-specific capacity and cumulative load are understood as emergent properties of interacting metabolic, mechanical, and coordinative processes.

PRACTICAL COROLLARIES

From this integrated concept follow several practice-relevant corollaries that refine, rather than reject, both camps.

Progression by constraint, not arithmetic

Instead of increasing “load” by a fixed scalar each microcycle, progress by constraint-aware adjustments:

- Intensity ceilings for sessions with high internal load risk (downhill running, firm surfaces in aggressive shoes, very fast strides): cap the number of steps above a clinically meaningful internal-force proxy (e.g., steps faster than 3K pace or on -2% grade) even when the metabolic budget allows more.
- Duration floors for metabolic work: ensure sufficient continuous time near LT1/LT2 for stimulus while maintaining mechanical safety by surface, footwear, and cadence choices. This is progression as shaping a feasible region, not stepping up a single dial.

Distribute threshold and localise peak force

Physiologically, splitting threshold work (e.g., $2 \times 20 \text{ min} \rightarrow 4 \times 10 \text{ min}$) maintains stimulus with lower systemic strain; mechanically, it also caps per-bout force drift that sometimes emerges late in continuous tempos. Conversely, if one wishes to practise race-specific economy at near-continuous intensities, mechanical risk can be localised by flattening course profiles, using more compliant footwear, and biasing cadence upward to reduce peak vertical excursion, all coordinative manipulations that alter internal force without blunting the metabolic aim.

Preserve form under speed with low fatigue

Short, relaxed fast running (extended strides, diagonals, light hill sprints) delivers high coordinative and elastic-system practice at minimal metabolic dose. Properly dosed (few minutes of total fast time), such work maintains the ability to recruit high-threshold motor units and to coordinate stiffness without meaningfully perturbing recovery (cf. your neuromuscular recruitment rationale). In this framework, these sessions are neither decoration nor “*speed for speed’s sake*,” but the coordinative moment’s weekly anchor that helps the metabolic and mechanical moments reconcile late-race demands.

Recognise tissue tempos and seasonality

Because tendon and bone remodel slowly, mechanical-risk features (aggressive downhill; spikes in very fast steps; abrupt footwear changes) should be introduced early and held constant over mesocycles, letting tissues catch up while the metabolic moment cycles through base, specific, and taper emphases. This is a guarded reading of “*keep workout volume constant*” at the level of mechanically salient step-counts rather than merely total metres at pace.

Individualise by coordinative signature

Since fatigue-driven kinematic shifts are idiosyncratic, periodic mechanics audits (high-frame-rate sagittal footage at easy, threshold, and closing-rep speeds) can identify each runner’s decompensation attractor (e.g., posterior pelvic tilt and over-stride; collapsing ankle stiffness). Training then includes micro-constraints to keep the runner near their economical solution: modest cadence prompts, light uphill finishes to dampen braking, or footwear selected for the athlete’s preferred stiffness interaction. The aim is not to impose a universal technique but to stabilise the individual’s efficient pattern as intensity rises.

Reconciling specific disputes

Equal fast-metres imply equal biomechanical load. False if read literally; plausible if read under strong constraints. Equalising only the distance at a target pace ignores distributional features (surfaces, gradients, footwear, cadence) and end-of-bout states. However, if those features are held within narrow bands, then progressing from many short reps to fewer, longer reps need not precipitate a large increase in damage relative to increasing pace or total fast metres, because damage is more sensitive to peak force than to modest extensions of steady-state work. The safer practical move, then, is often to lengthen reps first, before adding pace or extra fast-metres, while keeping mechanical constraints constant. And of course, this is a rule of thumb, but provided surface, footwear, and relative intensity are tightly constrained, a common progression heuristic is to slightly lengthen repetitions at a given pace before either adding total fast-metres or increasing pace.

Device impact metrics (ILR, LSS) are useful proxies

They are sometimes useful trend indicators, but they mislead when reified. Impact-centred metrics capture a narrow slice of external dynamics and correlate weakly with internal forces that matter for tissue damage (Matijevich et al., 2019). They may be included in a dashboard as flags (e.g., “*today’s downhill repetitions produced atypically high impact surrogates, consider reducing similar exposures this week*”), but planning should privilege mechanism-grounded constraints and athlete-specific responses over opaque composite scores.

For coaches and self-coached runners, the integrated approach can be rendered as a short grammar for designing and auditing a week:

1. Name the metabolic target (e.g., “*two hours near LT1 on Sunday; 30–40 minutes within LT2 across split sessions mid-week*”).

2. Bind the mechanical risk (e.g., “< 300 steps faster than 3K pace; no downhills steeper than –2%; one very-fast exposure day only”).
3. Install the coordinative anchor (e.g., “one low-fatigue fast-coordination session: 6–8 relaxed 120–150 m strides, or 8–10 × 10-s hill sprints”).
4. Progress within the feasible region by first extending bout duration at a given pace, then adding bouts, and only then increasing pace – re-auditing mechanical constraints at each change.
5. Monitor with two ledgers: a metabolic ledger (TRIMP/TSS, session RPE, HRV, subjective freshness) and a mechanical-exposure ledger (counts of very-fast steps, downhills, firm-surface minutes), plus periodic form spot-checks.

This grammar neither abolishes physiological scalars nor fetishises forces. It makes them co-determinant under coordinative governance.

Implications for performance and injury

Performance gains in the marathon are typically attributed to expanded aerobic capacity and improved substrate utilisation. Yet late-race collapse is often recognisable as a coordinative–mechanical failure: stride length decays, cadence drifts, posture folds, despite heart-rate remaining stable. By reserving a weekly slot for fast but fresh coordinative work, distributing threshold to maintain quality, and constraining mechanical hazards while metabolic work scales, one creates conditions in which the athlete can express the physiology already acquired. Injury risk, in turn, becomes a function not of “*too much load*” in the abstract, but of mis-shaped load – the wrong internal-force distributions rehearsed under the wrong coordinative states. The integrated practice pushes against that mis-shaping.

Finally, evaluation should match theory. Group-level correlations between heart-rate scalars and injury are predictably weak because they average away the very distributional and coordinative features that mediate risk. Within-athlete designs that log both metabolic scalars and mechanical-exposure primitives alongside simple kinematic markers (step length, cadence at fixed speeds; qualitative posture codes) will better detect the thresholds at which each athlete’s feasible region narrows. In this way, practice generates its own evidence without presuming that a single number can tell the whole story.

SUPPORTING AGENCIES

No funding agencies were reported by the author.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

REFERENCES

- Allen, H., & Coggan, A. R. (2010). Training and racing with a power meter (2nd ed.). VeloPress.
- Bassett, D. R., Jr., & Howley, E. T. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine & Science in Sports & Exercise*, 32(1), 70-84.
<https://doi.org/10.1097/00005768-200001000-00012>

- Bertelsen, M. L., Hulme, A., Petersen, J., Brund, R. K., Sørensen, H., Finch, C. F., Parner, E. T., & Nielsen, R. O. (2017). A framework for the etiology of running-related injuries. *Scandinavian journal of medicine & science in sports*, 27(11), 1170-1180. <https://doi.org/10.1111/sms.12883>
- Calvert, T. W., Banister, E. W., Savage, M. V., & Bach, T. (1976). A systems model of the effects of training on physical performance. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(2), 94-102. <https://doi.org/10.1109/TSMC.1976.5409179>
- Edwards, S. (1993). High performance training and racing. In S. Edwards (Ed.), *The Heart Rate Monitor Book* (pp. 113-123). Fleet Feet Press.
- Edwards, W. B. (2018). Modeling overuse injuries in sport as a mechanical fatigue phenomenon. *Exercise and Sport Sciences Reviews*, 46(4), 224-231. <https://doi.org/10.1249/JES.0000000000000163>
- Farris, D. J., & Sawicki, G. S. (2012). The mechanics and energetics of human walking and running: A joint level perspective. *Journal of the Royal Society Interface*, 9(66), 110-118. <https://doi.org/10.1098/rsif.2011.0182>
- Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., Doleshal, P., & Dodge, C. (2001). A new approach to monitoring exercise training. *Journal of strength and Conditioning Research*, 15(1), 109-115. <https://doi.org/10.1519/00124278-200102000-00019>
- Gabbett, T. J. (2016). The training-injury prevention paradox: Should athletes be training smarter and harder? *British Journal of Sports Medicine*, 50(5), 273-280. <https://doi.org/10.1136/bjsports-2015-095788>
- Hughes, J. M., Castellani, C. M., Popp, K. L., Guerriere, K. I., Matheny, R. W., Jr, Nindl, B. C., & Bouxsein, M. L. (2020). The Central Role of Osteocytes in the Four Adaptive Pathways of Bone's Mechanostat. *Exercise and sport sciences reviews*, 48(3), 140-148. <https://doi.org/10.1249/JES.0000000000000225>
- Jones, A. M., & Burnley, M. (2009). Oxygen uptake kinetics: an underappreciated determinant of exercise performance. *International journal of sports physiology and performance*, 4(4), 524-532. <https://doi.org/10.1123/ijsp.4.4.524>
- Joyner, M. J., & Coyle, E. F. (2008). Endurance exercise performance: The physiology of champions. *Journal of Physiology*, 586(1), 35-44. <https://doi.org/10.1113/jphysiol.2007.143834>
- Kjaer, M., Langberg, H., Heinemeier, K., Bayer, M. L., Hansen, M., Holm, L., & Doessing, S. (2009). From mechanical loading to collagen synthesis, structural changes and function in human tendon. *Scandinavian Journal of Medicine & Science in Sports*, 19(4), 500-510. <https://doi.org/10.1111/j.1600-0838.2009.00986.x>
- Magnusson, S. P., Hansen, M., Langberg, H., Miller, B., Haraldsson, B., Westh, E. K., Koskinen, S., Aagaard, P., & Kjaer, M. (2007). The adaptability of tendon to loading differs in men and women. *International journal of experimental pathology*, 88(4), 237-240. <https://doi.org/10.1111/j.1365-2613.2007.00551.x>
- Matijevich, E. S., Branscombe, L. M., Scott, L. R., & Zelik, K. E. (2019). Ground reaction force metrics are not strongly correlated with tibial bone load when running across speeds and slopes. *PLOS ONE*, 14(1), e0210000. <https://doi.org/10.1371/journal.pone.0210000>
- Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., Raglin, J., Rietjens, G., Steinacker, J., Urhausen, A., European College of Sport Science, & American College of Sports Medicine (2013). Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Medicine and science in sports and exercise*, 45(1), 186-205. <https://doi.org/10.1249/MSS.0b013e318279a10a>
- Seiler, S. (2010). What is best practice for training intensity and duration distribution in endurance athletes? *International Journal of Sports Physiology and Performance*, 5(3), 276-291. <https://doi.org/10.1123/ijsp.5.3.276>
- Seiler, S., & Tønnessen, E. (2009). Intervals, thresholds, and long slow distance: The role of intensity and duration in endurance training. *Sportscience*, 13, 32-53.

- van Hooren, B., Jukic, I., Cox, M. et al. The Relationship Between Running Biomechanics and Running Economy: A Systematic Review and Meta-Analysis of Observational Studies. *Sports Med* 54, 1269-1316 (2024). <https://doi.org/10.1007/s40279-024-01997-3>
- Vernillo, G., Giandolini, M., Edwards, W. B., Morin, J.-B., Samozino, P., Horvais, N., & Millet, G. Y. (2017). Biomechanics and physiology of uphill and downhill running. *Sports Medicine*, 47(4), 615-629. <https://doi.org/10.1007/s40279-016-0605-y>
- Zelik, K. E., & Franz, J. R. (2017). It's positive to be negative: Achilles tendon work loops during human locomotion. *PLOS ONE*, 12(7), e0179976. <https://doi.org/10.1371/journal.pone.0179976>
- Zelik, K. E., & Kuo, A. D. (2012). Mechanical work as an indirect measure of subjective costs influencing human movement. *PLOS ONE*, 7(2), e31143. <https://doi.org/10.1371/journal.pone.0031143>



Development and validation of the performance environment of practice scale

 **Steve Smith**  . University of Winchester. United Kingdom.
 **Ellie Gennings**. Bournemouth University. United Kingdom.
 **Nayoung Ahn**. Bournemouth University. United Kingdom.
 **Paul Sowden**. University of Winchester. United Kingdom.

ABSTRACT

The performance climate is affected by time, situation, and case, which requires frequent measurement within practice environments where athletes spend most of their time. This study aimed to develop and validate the Performance Environment of Practice Scale (PEPS). Three waves of data collection and analysis were conducted: content validity ($n = 30$), exploratory factor analysis ($n = 389$), and confirmatory factor analysis ($n = 198$). The scale development process provided initial evidence for the face validity and internal reliability of the 7-factor 26-item PEPS for use with athletes in a competitive practice environment. The PEPS provides a performance environment measure that carries lower response burden than previous measurement attempts. The PEPS can be used by coaches at regular intervals to measure the psychological performance climate of their practice environment. Further evaluation of the test-retest reliability and predictive validity of the PEPS is required, along with accuracy assessments against coach performance perceptions and the gathering of normative data.

Keywords: Physical activity psychology, Scale development, Validation, Psychological climate, Psychology, Sport.

Cite this article as:

Smith, S., Gennings, E., Ahn, N., & Sowden, P. (2026). Development and validation of the performance environment of practice scale. *Scientific Journal of Sport and Performance*, 5(3), 364-385. <https://doi.org/10.55860/MTRP5957>



Corresponding author. University of Winchester. United Kingdom.

E-mail: steve.smith@winchester.ac.uk

Submitted for publication November 17, 2025.

Accepted for publication January 09, 2026.

Published February 21, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/MTRP5957>

INTRODUCTION

Psychological climate is produced from individuals' perceptions of the environment in which they conduct tasks to meet outcomes (James & Sells, 2013). Perceptions are created by emotionally laden judgements (Lazarus, 1991) where an individual appraises a situation as either beneficial or harmful towards their goals (Pensgaard & Duda, 2002). The array of influencing factors that contribute towards psychological climate can be isolated and quantified (Durand-Bush & Salmela, 2002). The emotions experienced from situational appraisals have been found to provide a lasting impact on an individual and their behaviour (Roseman & Smith, 2001). First studied within industrial and organizational domains, psychological climate is recognized as a moderator variable between the structures and processes of an organization and employee outputs, making psychological climate a primary cause of job satisfaction and job performance (Furnham, 2012).

Psychological climate in the sporting domain has received considerable research attention (Arnold & Fletcher, 2012) because athletes operate within highly complex social and organizational environments that exert major influences on task performance (Hardy et al., 1996). As such, psychological climate in sport, which is based on the criteria of success and failure against contextually rich goals, can offer a measure of performance (Buch et al., 2017). Due to the performance focus, psychological climate research in sport (e.g., Douglas & Carless, 2006) has used the term performance environment, which will be used henceforth in this paper.

To better understand athletic performance environments, previous research has predominantly focused on competition. However, athletes spend most of their time within environments for practice, meaning the practice domain has an influence on competitive performance (Durand-Bush & Salmela, 2002; Ericsson, 2020; Macnamara et al., 2014). Deliberate practice activities have the goal of improving performance, are effortful and attention-demanding, are not necessarily enjoyable, and do not lead to immediate social or financial rewards (Ericsson et al., 1993). Deliberate practice has been suggested to have a linear relationship with the learning of sport-specific skills, physical fitness, and team strategies (Baker et al., 2003), which positions the psychological influencing factors of the practice environment as key components in athlete development and performance. Also, on-going long-term emotional exposure sustained through extensive practice activities has been found to influence performance (Baumeister, Vohs, DeWall, & Zhang, 2007), evidenced through coach-athlete conflict occurring far more often during practice than competition (Wachsmuth et al., 2018), which further strengthens the argument that practice can influence athletic performance.

Competition-dominant performance environment measures include those undertaken by Pain and Harwood (2007, 2008) who identified the psychological influencing factors within competitive soccer and developed the Performance Environment Survey (PES); a conceptually based instrument that captures information about physical, psychological, coaching, social, organizational, and environmental domains that influence team performance and preparation within a competitive environment. Pain et al. (2012) used the PES to measure the performance environment within a soccer team during a competitive season, which provided coaching staff with data to enhance performance outcomes. However, the 66-item PES (see Pain & Harwood, 2008) is focused only on perceptions within competitive environments (i.e., match day) and provides a response burden that can impact its functionality within a sport context (Rolstad et al., 2011). Despite the need to provide psychological measures that satisfy quality criteria, Horvath and Röthlin (2018) suggest there is a greater need for measurement tools to be shorter in sport performance settings to increase usability and accuracy.

The performance environment has also been studied more holistically within sporting organizations (Wagstaff, 2019). The High Performance Environment (HPE) model was theoretically constructed by Jones et al. (2009) to identify the optimal psychological climate of elite sporting groups. The model comprised of four factors to identify high performance, which included an organizational climate factor. Fletcher and Streeter (2016) applied the HPE model to an elite high performing swimming environment and reported the model to be capable of aiding practitioners in creating optimized performance environments for elite athletes. However, no further development of an HPE measurement tool has been conducted, which is most likely due to the complexity, volume and ever-evolving factors found within an entire organization. Therefore, measures of isolated performance environment domains, such as practice, would provide data that are more manageable and workable for real-world application (Arnold & Fletcher, 2012).

To better understand the psychological influences within the practice environment, motivational climate has been a focus of study for many decades (Lacerda et al., 2021). Motivational climate is the psychological environment that encompasses both competition and practice, which is created by coaches, peers and parents through situational factors that develop motivation based upon task/mastery and ego orientations (Harwood & Thrower, 2020; Smith et al., 2008). Based on Achievement Goal Theory (Nicholls, 1984), several measures of motivational climate have been developed to evaluate motivation towards the goals set within the environment (Lacerda et al., 2021). However, the performance environment of practice is impacted by affective facets that exist outside of the direct activities undertaken within competition and practice. For example, influencing factors include preparatory lifestyle behaviours, such as following recovery protocols between practice sessions (Smith et al., 2019), and appraisals being conducted on non-performance directed goals (i.e., threat from not conforming to the social identity of the group) (Smith et al., 2007). The psychological influencing factors between different domains (i.e., practice and competition) have also been found to differ (Fletcher et al., 2012). Therefore, motivational climate measures have limited capability to measure the performance environment of practice and further exploration of a valid domain-specific measure is required.

Smith et al. (2019) undertook the first known examination of the practice performance environment and reported differences against previously stated influencing factors from other sport domains (i.e., competition). For example, ability evaluations were made against teammates during practice that produced internalized ranking positions, which underpinned expectations about an athlete's own performance. Previous research by Mellalieu et al. (2009) identified internal and external performance expectations as providing competitive stress (negative) on athletes, but Smith et al. (2019) revealed a far more complex situation to exist in practice where contextual differences interacted with performance expectations to produce both positive and negative perceptions.

Further evaluation of the practice performance environment by Smith et al. (2020a; 2020b) explored the influencing factors cited by multiple stakeholders within the environment (e.g., players and coaches) and produced the Practice Environment Model (PEM). The PEM illustrates the cyclical (e.g., preparatory activity for the next session starts immediately after the last) and interrelatedness of the influencing factors in the practice environment (Smith et al., 2020b) and highlights the holistic nature of perception and influence. For example, influencing factors within the practice environment may be latent or without clear effect at the time of initial perception (e.g., negative emotions that provide resilience leading to positive performance) (Brown et al., 2015). Whereas measures of motivation climate, such as the perceived motivational climate in sport questionnaire-2 (PMCSQ-2) (Newton et al., 2000), provide a scale based on present situational perceptions focused solely upon task execution and does not consider the repetitive nature of practice due to its dual (competition and practice) focus.

A season long implementation of the PEM using an educational strategy with a national standard basketball team was conducted by Smith et al. (2021). The strategy was based on key influencing factors from the PEM that included high effort as a primary goal, pre-practice performance reflections, supportive communication, working towards shared team goals, enhanced practice preparation, and the understanding of what can be controlled within the environment. Findings provided evidence of enhanced performance following the educational strategy and highlighted the level of perceived control, expected performance, support behaviour, team-focused goals, internalized ability ranking against others within the group and practice preparation as key influencing factors of the psychological climate of practice. As the framework of factors provided by Smith and colleagues can be manipulated to enhance performance (Smith et al., 2021), the development of a valid and reliable measure of the performance environment of practice will aid those who create and maintain practice environments (i.e., coaches) to both evaluate and enhance athlete performance. To address the issue of not having a valid measure of the performance environment of practice, this study provided an initial validation of the Performance Environment of Practice Scale (PEPS).

Construction and scale development

The development of the PEPS was guided by MacKenzie et al. (2011) scale development framework. This paper details stages one to eight whereby the process of scale development (including defining the overarching construct, developing items, and specifying the measurement model) and scale validation (including content validity and factorial analysis) were undertaken.

Content validity is defined as 'the extent to which a specific set of items reflects a content domain' (DeVellis & Thorpe, 2022, p. 71) and was a central consideration throughout the development of the scale. Relevant literature was consulted to determine aspects that would construct the scale (Lamm et al., 2020). The performance environment of practice was theoretically underpinned by the practice environment research (Smith et al., 2019; 2020a; 2020b; 2021), which contained the key aspects and attributes of the construct (MacKenzie et al., 2011). Motivational climate and performance (competition-based) environment research that was referred to in the previous practice environment research, where psychological factors existed within the practice environment, were also reviewed (e.g., practice strategies grouped with competition (Harwood et al., 2015) and the requirement for practice activity between competitive play (Pain & Harwood, 2008)). However, no further aspects of the construct were identified outside of the contemporary practice environment research due to the specificity of the psychological influencing factors of practice. The first named author compiled a list of influencing factors that existed within the practice environment. All identified factors had a clear relationship with performance. To enhance robustness of scale construction and content validity by providing multiple sources, the second named author, who was an academic subject expert, acted as a critical friend to support the creation of initial items from the relevant influencing factors identified (Lamm et al., 2020).

In line with recommendations from MacKenzie et al. (2011), a detailed conception of the target construct was created that included specifying the phenomena to which the construct refers and the referents to which the construct applies. The performance environment of practice was defined as an athlete's perceptions of the psychological factors influencing performance in the sport practice environment. The construct is a representation of the emotional appraisals that identify both positive (beneficial) and negative (harmful) influences on performance (Lazarus & Folkman, 1984; Roseman & Smith, 2001) and are generated from both internal (self) and external (social) sources. The construct can be affected by time, situation, and case (Neil et al., 2007).

Following the review of previous practice environment literature, the authors identified seven reflective sub-

dimensions that captured the essential aspects of the construct. A definition of each sub-dimension can be seen in Table 1, and sub-dimension descriptions can be found in Appendix 1. Items were developed using procedures set out by Boateng et al. (2018) and MacKenzie et al. (2011), which included items being inductively developed using existing literature and deductively developed by two researchers until the saturation of ideas. Specifically, items had to start with the stem “*Within my Practice Environment I...*”, could not be double-barrelled and the critical friend had to agree that the item reflected its sub-dimension definition. A large volume of items were developed due to the expected extraction of items throughout the scale development process (Raykov & Marcoulides, 2011). An initial 89 item pool was created and organized into the seven sub-dimensions of control (9 items), expected performance (11 items), supportive behaviour (12 items), team-focused goals (14 items), ability ranking (22 items), preparation (11 items), and challenge (10 items).

Table 1. Definition of sub-dimensions.

Sub-Dimension	Definition
Control	The elements of practice performance that athletes attempt to directly influence through their own actions
Expected Performance	The impact of practice goals that define actions towards the outcome of performance
Support Behaviour	The assistance received from others within the practice environment to meet performance outcomes
Team-focused Goals	The extent that performance goals are adopted by all individuals within the practice environment
Ability Ranking	The self-creation of a hierarchy based on performance skills between oneself and teammates
Preparation	The lifestyle behaviours undertaken before entering the practice environment
Challenge	The activities of practice that build resilience from failure through reflective performance behaviours

The decision-making rules for formally specifying the measurement model proposed by Jarvis et al. (2003) were used to specify the model as reflective. Therefore, the model was suitable for factor analysis, which satisfied stage four of MacKenzie and colleagues (2011) scale development framework.

METHOD AND RESULTS

Following MacKenzie et al. (2011) stages one to eight, there were three waves of data collection reflecting assessment of content validity, and both exploratory and confirmatory factor analysis. Data were collected over a six-month period where each wave of data collection was analysed before the next wave was collected. This ensured that data collection/analysis informed the next wave of data collection/analysis. Institutional ethical approval was gained from the lead researcher's university ethics board. Data were analysed using the protocol outlined by MacKenzie et al. (2011).

WAVE ONE

Participants

The first wave of data collection was used to assess the developed items for content validity ($n = 30$). The sample (19 males, 11 females, mean age 21.27 years $\pm$ 1.31) were considered expert judges of content validity as they were the main population of interest for the scale whereby the sample met the following

inclusion criteria: (a) involved in a team sport practice environment for at least two years; (b) in a practice environment focused on performance development; (c) competing in governed leagues or tournaments; (d) undertaking deliberate practice activities regularly for a sustained period, such as an entire competitive season; (e) 16 years or over and possess sufficient intellectual ability to rate items against definitions of construct domains.

Protocol

For the first wave of data collection, potential participants were emailed an information sheet and consent form and given the opportunity to ask questions. To recruit participants, the research team promoted the study to sports teams and clubs that held an association, through staff, students, and professional activities, to the United Kingdom university through which the study was being undertaken. If participants met the recruitment criteria and were interested in partaking, they replied to the researcher with a signed consent form. Recommendations from MacKenzie et al. (2011) guided the assessment of content validity. Each participant was given a paper item rating sheet that contained a matrix with eight columns (see Appendix 2 for the task instructions and first two pages of the form). The first column and subsequent rows randomly listed the 89 items initially developed for the PEPS. A definition of each of the seven sub-dimensions headed columns two to eight (see Table 1). Three randomized versions of the item rating sheet were created to control for response-order effect (Mackinnon & Wang, 2020). Each participant was asked to familiarize themselves with the sub-dimension definitions. Participants were then asked to read each item and rate the appropriateness against all the sub-dimensions on a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*completely*). Due to the item rating sheet being 10 pages long with 89 items to individually rate, participants were encouraged to take regular breaks and reassured they could have up to two weeks to complete the task.

Data analysis

SPSS statistics (version 28; IBM, Armonk, NY) was used to analyse wave one data. First, data were assessed for normality via a Shapiro-Wilk test. Once normality assumptions were satisfied, a repeated measures analysis of variance (ANOVA) per item to determine whether items significantly captured their sub-dimension definition was computed. Pairwise comparisons, which indicate an item's mean rating against all sub-dimensions, were consulted to see if a significant difference was found. Items that were not significantly rated with their sub-dimension definition were removed. Significance was defined by an alpha level of .05.

Results

The normality assumptions were met, as assessed by Shapiro-Wilk ($p > .05$) and there were no missing data. Fifty-eight items were significantly rated by the sample to represent their intended sub-dimension's definition (p 's $< .05$) and were subsequently retained for wave two analysis (i.e., exploratory factor analysis [EFA]). Items that were not rated as significantly representing their intended sub-dimension were removed ($n = 31$).

WAVE TWO

Participants

The second wave of data were collected from 389 participants (227 males, 161 females, 1 preferred not to state gender, mean age 19.98 years $\pm$ 5.32), and all participants met the same criteria stated in wave one above.

Protocol

Data for wave two were collected using Jisc online survey and distributed to sports clubs, teams, and educational organizations that were associated to the university. Participant information, recruitment criteria and demographic data collection were on the first page of the survey with a required checkbox to confirm consent. Items were randomized using an online random generator before uploading them onto the survey platform. Participants were asked to rate each item based on their current perceptions of their practice environment on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). All responses were anonymous.

Data analysis

An EFA was carried out to examine the psychometric properties of the scale. Data were analysed using SPSS statistics (version 28; IBM, Armonk, NY), the R package psych (version 2.5.6) of R software (version 4.4.3), and R Studio (version 2024.04.2). Data were assessed for univariate outliers (z-scores > 3.29) and multivariate outliers, via calculation of Mahalanobis distances and comparison to a χ^2 distribution ($df = 23$, $p < .001$). Nineteen cases were identified with potential outliers. These cases were checked for data entry errors and any 'suspicious' patterns of responding (e.g. selecting all 1s). No such issues were identified. Based on this, all responses were retained as valid. To understand the influence of these potential outliers, an EFA was computed with and without outliers. The results yielded a similar factor structure in both cases, with one additional factor when outliers were excluded. This additional factor comprised of three items relating to "preparation." However, when the analysis was run with outliers included there was still the "preparation" factor evident, and the authors were satisfied that the solution adequately represented the underlying theory of practice performance environment. Consequently, as both versions of the analysis resulted in similar solutions, and because the intention was for respondents to use the full range of values on the 7-point Likert scales, a decision was made to retain the analysis with all cases included, including extreme responses, and this analysis is reported in full next.

Given the ordinal nature of the data, polychoric correlations were computed to assess underlying sub-dimensions. Although observed ordered categorical responses are often treated as continuous variables with multivariate normality, enabling the use of maximum likelihood (ML) method (Muthén & Kaplan, 1985), the complexity of the ordinal items and the models under examination led to using oblique EFA on the polychoric correlations among the latent variables. Hence, the EFA was run using an oblique technique with promax rotation, allowing for intercorrelated factors as opposed to the orthogonal rotation (Fabrigar et al., 1999). In factor analysis, principal axis factoring was employed based on the assumptions that the underlying factors remain interrelated.

Furthermore, with no missing data, multiple criteria were used to identify the most appropriate factor solution for the PEPS (Table 2; Pett et al., 2003; Harrington, 2009; Comrey & Lee, 2013; Tabachnick & Fidell, 2014). First, eigenvalues were looked at, along with the scree plot, to identify where the eigenvalues start to level off. Eigenvalues greater than 1.0 were used as the benchmark for factor retention (Kaiser, 1974). Data were checked for suitability of factor analysis by inspection of Kaiser Meyer Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity (Carpenter, 2018). A parallel analysis was also conducted to retain the number of factors, and ordinal alpha values (Zumbo et al., 2007) were checked for all factors to assess internal consistency. Finally, the cutoff for factor loadings was set to .40 or above on a single factor (Nunnally & Bernstein, 1994).

Table 2. Item extraction method.

Stage	Criteria for item removal
1	Correlation coefficient is >0.8
2	Mis-loading
3	Cross-loading occurs over >2 factors
4	Cross-loadings over 2 factors: remove items with higher cross-loadings first
5	Items with singular loadings $\leq .4$

Results

Using the wave two data set, an EFA was conducted to investigate the factor structure. The Kaiser-Meyer-Olkin measure of sampling adequacy was .80, exceeding the cut-off value of .70, and the Bartlett's test of sphericity was statistically significant ($\chi^2 [253] = 4375.73, p < .001$, indicating that the data were suitable for factor analysis (Kaiser, 1974). Based on low scores ($\lambda < 0.2$), the items from the subdimensions *Expected Performance* and *Control* were removed, along with 25 items that were misloaded (an item loading onto a factor that it was not theorized to) or cross-loaded (an item loading onto more than one factor). Furthermore, the Ability Ranking subscale was split into three factors due to showing a clear relationship between items and particular aspects of the subscale Ability Ranking, which were renamed Teammate Performance, Effort Changes, and Hierarchy Focus, with sub-dimension definitions displayed in Table 3.

Table 3. Split ability ranking sub-dimension definitions.

Sub-Dimension	Definition
Teammate performance	The perceived ability position within the team based on the success and failure of a teammate rather than own performance
Effort changes	Altered competitive focus against teammates of different ability level
Hierarchy focus	Self-perceived ability dependent on comparisons made against teammate performance

While the parallel analysis suggests that not all factors should be retained, factors six and seven are conceptually important and align with the theory of practice performance environments. Factors five, six and seven only included two items, but were retained due to their reliability scores, high factor loadings, and conceptual importance. A new item was developed for each factor, reflecting the sub-dimension definitions and the two remaining items (*effort changes*: “change my effort level when competing against different teammates”; team-focused goals: “fully commit myself to team goals”; preparation: “prepare myself to perform well”).

In the EFA, a 7-factor solution emerged with 22 items, explaining 59.7% of the total variance among the variables. Based on the screen plot, eigenvalues, parallel analysis, and model fit indices, we retained the final PEPS model with 7 factors and 22 items, which was used in the subsequent analysis. The results of the pattern matrix from promax rotation are presented in Table 4. All 22 items were loaded onto seven distinct factors, each representing a specific domain, without double loading: support behaviour (5 items), hierarchy focus (4 items), teammate performance (3 items), challenge (4 items), team-focused goals (2 items), effort changes (2 items), preparation (2 items). Ordinal alpha coefficients for the factors ranged from .73 to .92 (see Table 4), indicating that the overall items under each factor were all internally consistent and reliable.

Table 4. Rotated scale structure from EFA.

Code	Items (Within my Practice Environment I...)	Subdimensions						
		Support behaviour	Hierarchy focus	Teammate performance	Challenge	Team- focused goals	Effort changes	Preparation
SB2	Feel supported by coaches	0.97	-0.01	-0.05	-0.01	-0.06	-0.01	0.03
SB12	Have a good relationship with my coach/es	0.82	-0.12	-0.06	0.15	0.06	-0.03	-0.06
SB4	Feel my coach/es talk to me positively	0.75	-0.09	0.09	-0.07	0.01	0.05	0.13
SB8	Feel supported	0.78	0.02	0.02	0.15	-0.01	0.01	-0.03
SB10	Can rely on my coach to make me feel good even if I don't perform well	0.76	0.04	-0.02	-0.12	0.02	-0.03	0.05
AR4	Rank my ability against my teammates	-0.14	0.92	-0.02	0.02	0.01	0.04	-0.01
AR20	Compare my performance to others	0.03	0.91	-0.07	-0.18	0.12	-0.04	0.07
AR22	Judge my performance against teammate performance	-0.07	0.83	0.06	0.09	0.10	-0.09	0.03
AR15	Worry about what my teammates think of my ability	0.04	0.43	0.01	0.03	-0.19	0.05	0.05
AR7	Don't want teammates of similar ability to me to do well	-0.04	0.03	0.92	-0.05	0.10	-0.07	0.04
AR8	Want teammates to fail who are a similar level to me	-0.03	-0.09	0.91	0.09	-0.03	0.03	0.03
AR1	Don't want teammates to do better than me	0.02	0.25	0.62	-0.03	0.17	0.02	0.03
AR11	Get angry if a superior skilled teammate beats me in a drill	0.16	0.22	0.43	-0.03	-0.02	0.03	-0.18
CH7	See problems as challenges	-0.04	0.06	-0.12	0.77	0.06	-0.04	-0.16
CH2	Have overcome difficult situations	-0.12	-0.09	0.14	0.58	-0.05	0.01	0.15
CH5	Seek out challenge to improve	0.01	-0.13	0.06	0.69	-0.02	-0.02	0.02
CH6	Turn failure into future success	-0.12	0.21	-0.10	0.62	0.07	0.07	0.04
TG3	Put the team first and myself second	-0.03	0.04	0.03	0.07	0.64	0.01	-0.02
TG11	Hold the team's goals as more important than my individual goals	0.05	-0.03	0.02	-0.04	0.85	0.05	0.06
AR13	Drop my effort in drills against lesser skilled teammates	-0.01	-0.05	-0.06	0.06	-0.03	0.91	0.01
AR5	Change the way I compete against teammates depending on their ability level	-0.01	0.02	0.02	-0.07	0.08	0.72	-0.01
PR5	Perform injury prevention activities	0.05	0.10	-0.01	0.06	-0.01	-0.03	0.65
PR6	Perform recovery activities	0.08	0.01	0.02	-0.02	0.04	0.02	0.82
Eigenvalues		5.76	3.23	1.88	1.68	1.36	1.21	1.02
Variance Explained (%)		25.37	18.89	14.57	13.10	9.73	9.21	9.10
Ordinal α_o		0.91	0.83	0.77	0.77	0.77	0.72	0.74

WAVE THREE

Participants

The wave three data were collected for a confirmatory factor analysis (CFA; $n = 198$; 119 males, 79 females, mean age 22.20 years $\pm$ 8.24). All participants met the same criteria stated in the wave one above.

Protocol

As with the wave two data collection, a Jisc online survey was distributed to sports clubs, teams, and educational organizations that were associated to the university with the same participant information and consent collected. Items were randomized and uploaded to the survey platform. Participants were asked to rate each item based on their current perceptions of their practice environment on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). All responses were anonymous.

Data analysis

A CFA was conducted to examine the fit of the model identified in the EFA. Data were analysed using SPSS statistics (version 28; IBM, Armonk, NY), the lavaan package (version 0.6—19; Rosseel, 2012) of R software (version 4.4.3), and R Studio (version 2024.04.2). Previous research has indicated that psychological data are rarely normally distributed, resulting in overestimation of factor loadings (Chou & Bentler, 1995). Hence, weighted least squares with mean and variance adjusted (WLSMV) estimators, instead of the ML method, were used because it is appropriate for Likert-type scaled data and is robust to nonnormality, which is common in ordered categorical data (Myer et al., 2023). However, the ML estimation was also adopted for comparison with the WLSMV method. To evaluate the model fit, the Hu and Bentler's (1999) indices were employed: the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), the root mean square residual (SRMR), and the weighted root mean square residual (WRMR). The following cut-offs indicate an acceptable fit (CFI > .90; TLI > .90; RMSEA < .06; SRMR < .09; WRMR < 1.0), and significance was defined by an alpha level of .05 in this study.

Results

A CFA was first conducted to examine the 7-factor structure found in the previous EFA by employing the WLSMV estimation. Results of the CFA analysis revealed acceptable fit to the data ($\chi^2 [254] = 483.30$, $\chi^2/df = 1.90$, $p < .001$, CFI = .98, TLI = .97, RMSEA = .06, WRMR = 1.04). The correlations among the factors ranged from small (effort changes and preparation, $r = .17$) to moderate-to-large (challenge and preparation, $r = .68$). A second CFA was conducted using ML estimation for comparison. While the ML estimation model showed reasonably good to acceptable fit ($\chi^2 [254] = 483.30$, $\chi^2/df = 1.90$, $p < .001$, CFI = .98, TLI = .97, RMSEA = .07, SRMR = .07), the results were consistent with the WLSMV method. Two competing CFA analyses were evaluated, with factors expected to be related since the factor correlations were low to moderate (ranging from .17 to .68). The model fit indices were nearly identical ($|\Delta CFI| = .00$), not exceeding the threshold of .01 for meaningful difference (Cheung & Rensvold, 2002). Given the ordinal nature of the data, we chose to retain the WLSMV estimator.

As for construct reliability, both the ordinal alpha ($\alpha_o = .71-.87$) and the composite reliability (CR = .77-.88) were considered satisfactory (Nunnally & Bernstein, 1994). The communalities of the items ranged from .24 to .84. Although the items AR11 (get angry if a superior skilled teammate beats me in a drill) and CH2 (have overcome difficult situations) had low communalities ($h^2 < .3$), all items showed factors loadings above .40, ranging from .49 to .91 (Hair et al., 2010). Therefore, these items were kept given their significant factor loadings and theoretical importance. The average variance extracted (AVE) values for each construct were above the threshold of .5 (Fornell & Lacker, 1981), indicating convergent validity was acceptable (see Table

X). Further, discriminant validity was established, as the AVE values exceeded the corresponding squared correlations of all pairs of constructs (Fornell & Larcker, 1981). See Table 6 for the factor loadings and CR.

Table 6. Regression weight estimates standardized and unstandardized, along with internal reliability scores of the PEPS.

Items	Subdimension	Standardized regression weight	Unstandardized regression weight	Standard Error	Ordinal α_o
SB10	Support Behaviour	0.57	0.83	0.10	0.83
SB4		0.68	0.72	0.06	0.82
SB12		0.67	0.63	0.06	
SB2		0.89	1.00		
SB8		0.79	0.93	0.07	
ARA11	Hierarchy Focus	0.50	0.57	0.07	0.81
ARA1		0.85	0.92	0.06	
ARA8		0.83	0.94	0.06	
ARA7		0.89	1.00		
ARC15	Teammate Performance	0.60	0.74	0.08	0.75
ARC4		0.66	0.73	0.07	
ARC22		0.80	0.95	0.08	
ARC20		0.86	1.00		
CH6	Challenge	0.74	1.00		0.69
CH5		0.80	0.94	0.09	
CH7		0.67	0.90	0.10	
CH2		0.38	0.48	0.09	
TG3	Team-Focused Goals	0.69	0.88	0.13	0.85
TG11		0.80	1.00		
TG99		0.48	0.40	0.07	
ARB99	Effort Changes	0.81	0.941	0.076	0.79
ARB13		0.739	0.791	0.07	
ARB5		0.901	1		
PR6	Preparation	0.779	0.916	0.087	0.79
PR99		0.614	0.549	0.065	
PR5		0.872	1		

DISCUSSION

Initial validity evidence for the internal structure of the Performance Environment of Practice Scale (PEPS) has been provided for use with groups of individuals who regularly train within a practice environment to develop their performance and compete in sport. This study followed recommendations and guidance from scale development literature (MacKenzie et al., 2011; Boateng et al., 2018) while integrating theory (i.e., practice environment, performance environment and motivational climate research) and primary research to develop and provide the initial validation of the PEPS structure. Transparency allows readers to pass judgement on the methodological rigor and validity of a scale (Flake & Fried, 2020), therefore, this discussion details decision-making processes for clarity to present an accurate representation of the scale development

process. The data collection and analysis of the PEPS has been reported as a simple and linear process which does not reflect the complexities of scale development. During the validation of the PEPS, one theorized factor split, and two factors were removed. The following section aims to provide an honest discussion of the decision-making process while validating the PEPS and highlighting the tension between achieving statistical benchmarks and preserving underpinning theory.

Splitting of the ability rank subscale

Following the EFA, the Ability Ranking factor was split into three separate factors: Teammate Performance, Effort Changes and Hierarchy Focus. While this decision was data-led, reflecting on the underpinning research of the PEPS provides an explanation as to why the subscale was divided into three distinct factors. The ability ranking factor was shown by Smith et al. (2019) to hold a complex and dynamic relationship with performance expectations. The presence of strong comparisons within a practice environment, such as those found within an ego-orientated peer climate (Ingrell et al., 2016), can create performance expectations that are primarily influenced by perceived ability status within a group (Smith, 2003). Perceived performance outcome is, therefore, judged against comparison to others rather than other performance indicators, which provides the need for an isolated measure from other aspects of ability ranking and produced the Hierarchy Focus sub-dimension.

Effort is a particularly unique influencing factor due to its controllability (Douglas & Carless, 2006) and domain-wide impact within a performance climate (Hodge et al., 2014). Effort levels during intrateam competition in practice have been reported to differ depending on the perceived ability of the opposition (Smith et al., 2020a). Therefore, applied effort in practice can be an outcome of the status perceptions created (e.g., lower effort against lower ability opposition) and should hold its own distinctive sub-dimension factor, which was labelled Effort Changes.

The newly created Teammate Performance factor was rooted in the conflict between the success of the self against that of the team, which was exacerbated by intra-team competition (Mills et al., 2012). The impact of others' performance outcomes within the practice environment interacts with the direction of one's goals (individual vs team) and can influence practice performance (Smith et al., 2021). The perceptions of ability status within a team (i.e., Hierarchy Focus factor) and the influence of teammate performance have distinct differences (Koster & Aven, 2018), so should be measured separately.

Removal of two subscales: Expected performance and control

Two sub-dimensions were removed, Expected Performance and Control, due to poor internal reliability. This outcome may be attributed to ambiguity or a lack of explicit definition of these sub-dimensions, or the influence of various factors like situational anxiety, emotion, and confidence on performance expectations (Stamenković & Vesković, 2022). Performance expectations have also been found to be independent from previous performance (Feltz et al., 2008) but related to intra-team comparison (Smith et al., 2020a). The influence of ability status perceptions on performance expectations can be seen in the newly created Hierarchy Focus sub-dimension above, but as a stand-alone sub-dimension it seems Expected Performance could not be reliably measured and was removed.

The Control sub-dimension was similar to Expected Performance in such that when given further distinction its presence could be seen within the newly created sub-dimension Effort Changes. In fact, Smith et al. (2021) did not distinguish between effort and control when reporting findings. Effort has been described as more controllable than situational ability or talent (Douglas & Carless, 2006), with participants in the Smith et al.

(2021) study stating effort to be the only factor within the practice environment to be controllable. Therefore, the Control sub-dimension required further distinction and was removed.

Team-focused goals: Reliability

Scale development literature advocates for sub-dimensions to have at least three items per subscale for the dimension to be accurately represented in a scale (Carpenter, 2018). The resulting EFA included two items for the subscale Team-focused Goals. Based on MacKenzie et al. (2011) guidance of a minimum of three items per subscale, a third item was developed for the CFA which reflected the two remaining items. While the internal reliability score fell just short of the statistical benchmark ($\alpha = 0.69$) and removing the newly developed item that had a lower loading (0.47) would have rectified this ($\alpha = 0.74$), the researchers wanted to meet scale structure guidance over statistical benchmarks. The theoretical grounding that team-focused goals are a key element to the practice performance environment (e.g., Smith et al., 2020b), justify the retention of this factor and its items.

Scale limitations

While samples for the EFA and the CFA met the required practice involvement criteria, no further demographic data were collected, and the sample size for the CFA ($n = 198$) fell just short of Harrington's (2009) sample size recommendation ($n = 200$). Future research should test the PEPS with larger and more diverse sport populations where demographic data is collected. Populations such as individual and team sports, competition levels (e.g., elite and non-elite), and different cultures would improve the diversity of the data collected. Practice perceptions are highly unstable, and the cross-sectional study design does not allow for the determination of causal relationships (Morgado et al., 2017). Future longitudinal study can evaluate PEPS validity across different time points and compare against further measures of practice and competition performance (e.g., objective match results and subjective coach performance ratings). It is possible that researcher bias occurred during literature review interpretations to conceptualize the construct (Fisher et al., 2014). Further, practice environment research is a burgeoning field of study, and the PEPS may require updating.

Practical implications

A recent evaluation of questionnaire length conducted by Eisele et al. (2022) found a 60-item questionnaire led to increased perceived burden, lower compliance, increased careless responding, and decreased data quality and quantity when compared to a 30-item questionnaire. Galesic & Bosnjak (2009) also reported increased completion rates when participants believed the questionnaire to be shorter. From a practical perspective, the 26-item PEPS provides a performance environment measure with potentially lower response burden that can be used at higher frequencies with athletes by being quicker and less arduous to complete (Eisele et al., 2022). The lower response burden is created from a scale that has a specific domain focus (i.e., practice). Having a high frequency measure is particularly important due to practice environment perceptions having been conceptualized as unstable (Smith et al., 2019). The authors recommend that the measure may be suitable for daily use, or as and when required, and with competitive team sport athletes aged >16 years old. A version of the PEPS for practical use can be found in Appendix 3. The PEPS total score can be used to monitor the performance environment of practice in a group over time. Future research should aim to deliver PEPS normative data.

The PEPS is comparable to other short-form scales, such as the trait emotional intelligence questionnaire short-form that has 30 items (Laborde et al., 2017). The PEPS, therefore, could be used by coaches at regular intervals to measure the psychological performance climate of their practice environment. However, sports programs tend to favour very brief measures due to their ease of use and automation capacity (Duignan et

al., 2020). Longitudinal measures in sport are sought after (e.g., intra- and inter-season) but are often without high data collection frequency (e.g., Simms et al., 2021). Importantly, coaches have acknowledged the need for psychological measures but lack the understanding and time to invest in complicated and extensive measurement tools (Brink et al., 2018). Measures with items as few as five (e.g., Zsiso et al., 2020) and six (e.g., Andreassen et al., 2012) items have been developed and validated, and reduced item data collection methods should be the goal of future sport psychology research.

CONCLUSION

The current study undertook an initial development and validation of the PEPS by integrating theory and primary data to measure the performance environment of practice among competitive sport athletes. The factor analysis confirmed the structure of the model and that the scale has good internal reliability. Transparency in the reporting of the scale development and validation highlighted that changes to the scale structure were in accordance with underpinning theory. To continue building evidence of validity, future research should be conducted to evaluate the test-retest reliability and predictive validity of the PEPS. Future research should also assess PEPS accuracy against coach performance perceptions and whether the PEPS responds to implemented psychological climate changes.

AUTHOR CONTRIBUTIONS

Steve Smith conceptualised the study, led the scale development process, conducted data collection, contributed to data analysis and interpretation, and drafted the manuscript. Ellie Gennings led the statistical analyses (EFA and CFA), contributed to scale development, supported methodological design, assisted with data interpretation, and critically reviewed and revised the manuscript. Nayoung Ahn contributed to the statistical analyses (EFA and CFA), contributed to data interpretation and critically revised the manuscript for intellectual content. Paul Sowden provided theoretical and methodological oversight and supported data interpretation. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work.

SUPPORTING AGENCIES

No funding agencies were reported by the authors.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

- Andreassen, C. S., Torsheim, T., Brunborg, G. S., & Pallesen, S. (2012). Development of a Facebook addiction scale. *Psychological Reports*, 110(2), 501-517. <https://doi.org/10.2466/02.09.18.PR0.110.2.501-517>
- Arnold, R., & Fletcher, D. (2012). A research synthesis and taxonomic classification of the organizational stressors encountered by sport performers. *Journal of Sport and Exercise Psychology*, 34(3), 397-429. <https://doi.org/10.1123/jsep.34.3.397>

- Baker, J., Horton, S., Robertson-Wilson, J., & Wall, M. (2003). Nurturing sport expertise: factors influencing the development of elite athlete. *Journal of Sports Science & Medicine*, 2(1), 1-9. [https://doi.org/10.1016/S1440-2440\(01\)80002-5](https://doi.org/10.1016/S1440-2440(01)80002-5)
- Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behaviour: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, 11(2), 167-203. <https://doi.org/10.1177/1088868307301033>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Brink, M. S., Kuyvenhoven, J. P., Toering, T., Jordet, G., & Frencken, W. G. (2018). What do football coaches want from sport science?. *Kinesiology*, 50(1), 150-154.
- Brown, H. E., Lafferty, M. E., & Triggs, C. (2015). In the face of adversity: Resiliency in winter sport athletes. *Science & Sports*, 30(5), e105-e117. <https://doi.org/10.1016/j.scispo.2014.09.006>
- Buch, R., Nerstad, C. G., & Säfvenbom, R. (2017). The interactive roles of mastery climate and performance climate in predicting intrinsic motivation. *Scandinavian Journal of Medicine & Science in Sports*, 27(2), 245-253. <https://doi.org/10.1111/sms.12634>
- Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, 12(1), 25-44. <https://doi.org/10.1080/19312458.2017.1396583>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Chou, C.-P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37-55). Sage Publications, Inc.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press. <https://doi.org/10.4324/9781315827506>
- DeVellis, R.F. & Thorpe, C.T. (2022). *Scale development: Theory and applications*. 5th ed.) Sage publications.
- Douglas, K. & Carless, D. (2006). *Performance environment research: Research report*. London: UK Sport Performance Environment Research.
- Duignan, C., Doherty, C., Caulfield, B., & Blake, C. (2020). Single-item self-report measures of team-sport athlete wellbeing and their relationship with training load: A systematic review. *Journal of Athletic Training*, 55(9), 944-953. <https://doi.org/10.4085/1062-6050-0528.19>
- Durand-Bush, N., & Salmela, J. H. (2002). The development and maintenance of expert athletic performance: Perceptions of world and Olympic champions. *Journal of Applied Sport Psychology*, 14(3), 154-171. <https://doi.org/10.1080/10413200290103473>
- Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*, 29(2), 136-151. <https://doi.org/10.1177/1073191120957102>
- Ericsson, K. A. (2020). Towards a science of the acquisition of expert performance in sports: Clarifying the differences between deliberate practice and other types of practice. *Journal of Sports Sciences*, 38(2), 159-176. <https://doi.org/10.1080/02640414.2019.1688618>
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406. <https://doi.org/10.1037/0033-295X.100.3.363>

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). Self-efficacy in Sport. *Human Kinetics*. <https://doi.org/10.5040/9781718206625>
- Fisher, R., Maritz, A., & Lobo, A. (2014). Evaluating entrepreneurs' perception of success: Development of a measurement scale. *International Journal of Entrepreneurial Behavior & Research*, 20(5), 478-492. <https://doi.org/10.1108/IJEBR-10-2013-0157>
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456-465. <https://doi.org/10.1177/2515245920952393>
- Fletcher, D., Hanton, S., Mellalieu, S. D., & Neil, R. (2012). A conceptual framework of organizational stressors in sport performers. *Scandinavian Journal of Medicine & Science in Sports*, 22(4), 545-557. <https://doi.org/10.1111/j.1600-0838.2010.01242.x>
- Fletcher, D., & Streeter, A. (2016). A case study analysis of a high performance environment in elite swimming. *Journal of Change Management*, 16(2), 123-141. <https://doi.org/10.1080/14697017.2015.1128470>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- Furnham, A. (2012). Culture at work. In A. Furnham (Ed), *The psychology of behaviour at work: The individual in the organization*. Psychology Press. <https://doi.org/10.4324/9780203506974>
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349-360. <https://doi.org/10.1093/poq/nfp031>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice-Hall.
- Hardy, L., Jones, J. G., & Gould, D. (1996). *Understanding psychological preparation for sport: Theory and practice of elite performers*. John Wiley & Sons Inc.
- Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195339888.001.0001>
- Harwood, C. G., Keegan, R. J., Smith, J. M., & Raine, A. S. (2015). A systematic review of the intrapersonal correlates of motivational climate perceptions in sport and physical activity. *Psychology of Sport and Exercise*, 18, 9-25. <https://doi.org/10.1016/j.psychsport.2014.11.005>
- Harwood, C. G., & Thrower, S. N. (2020). Motivational climate in youth sport groups. In M.W. Bruner, M.A. Eys & L.J. Martin (Eds). *The power of groups in youth sport* (pp. 145-163). Academic Press. <https://doi.org/10.1016/B978-0-12-816336-8.00009-3>
- Hodge, K., Henry, G., & Smith, W. (2014). A case study of excellence in elite sport: Motivational climate in a world champion team. *The Sport Psychologist*, 28(1), 60-74. <https://doi.org/10.1123/tsp.2013-0037>
- Hoe, S. L. (2008). Issues and procedures in adopting structural equation modelling technique. *Journal of Quantitative Methods*, 3(1), 76.
- Horvath, S., & Röthlin, P. (2018). How to improve athletes' return of investment: Shortening questionnaires in the applied sport psychology setting. *Journal of Applied Sport Psychology*, 30(2), 241-248. <https://doi.org/10.1080/10413200.2017.1382020>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Humble, S. (2020). *Quantitative analysis of questionnaires: Techniques to explore structures and relationships*. Routledge. <https://doi.org/10.4324/9780429400469>

- Ingrell, J., Johnson, U., & Ivarsson, A. (2016). Relationships between ego-oriented peer climate, perceived competence and worry about sport performance: A longitudinal study of student-athletes. *Sport Science Review*, 25(3-4), 225-242. <https://doi.org/10.1515/ssr-2016-0012>
- James, L. R., & Sells, S. B. (2013). Psychological climate: Theoretical perspectives and empirical research. In D. Magnusson (Ed.), *Toward a psychology of situations: An interactional perspective* (pp. 285-306). Psychology Press.
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218. <https://doi.org/10.1086/376806>
- Jones, G., Gittins, M., & Hardy, L. (2009). Creating an environment where high performance is inevitable and sustainable: The high performance environment model. *Annual Review of High Performance Coaching and Consulting*, 1, 139-149. <https://doi.org/10.1260/ijssc.4.suppl-2.671q532j757771rl>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, 34(1), 111-117. <https://doi.org/10.1177/001316447403400115>
- Kline, P. (2000). *Handbook of psychological testing*. Psychology Press. <https://doi.org/10.4324/9781315812274>
- Koster, J., & Aven, B. (2018). The effects of individual status and group performance on network ties among teammates in the National Basketball Association. *PloS ONE*, 13(4): e0196013. <https://doi.org/10.1371/journal.pone.0196013>
- Laborde, S., Guillén, F., & Watson, M. (2017). Trait emotional intelligence questionnaire full-form and short-form versions: Links with sport participation frequency and duration and type of sport practiced. *Personality and Individual Differences*, 108, 5-9. <https://doi.org/10.1016/j.paid.2016.11.061>
- Lacerda, A., Filgueiras, A., Campos, M., Keegan, R., & Landeira-Fernández, J. (2021). Motivational Climate Measures in Sport: A Systematic Review. *The Spanish Journal of Psychology*, 24, E27. <https://doi.org/10.1017/SJP.2021.13>
- Lamm, K. W., Lamm, A. J., & Edgar, D. (2020). Scale development and validation: Methodology and recommendations. *Journal of International Agricultural and Extension Education*, 27(2), 24-35. <https://doi.org/10.4148/2831-5960.1115>
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819-834. <https://doi.org/10.1037/0003-066X.46.8.819>
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly* 35(2), 293-334. <https://doi.org/10.2307/23044045>
- Mackinnon, S. P., & Wang, M. (2020). Response-Order Effects for Self-report Questionnaires: Exploring the role of Overclaiming Accuracy and Bias. *Journal of Articles in Support of the Null Hypothesis*, 16(2).
- Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, 25(8), 1608-1618. <https://doi.org/10.1177/0956797614535810>
- Mellalieu, S. D., Neil, R., Hanton, S., & Fletcher, D. (2009). Competition stress in sport performers: Stressors experienced in the competition environment. *Journal of Sports Sciences*, 27(7), 729-744. <https://doi.org/10.1080/02640410902889834>
- Mills, A., Butt, J., Maynard, I., & Harwood, C. (2012). Identifying factors perceived to influence the development of elite youth football academy players. *Journal of Sports Sciences*, 30(15), 1593-1604. <https://doi.org/10.1080/02640414.2012.710753>

- Morgado, F. F., Meireles, J. F., Neves, C. M., Amaral, A., & Ferreira, M. E. (2017). Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica* 30(3). <https://doi.org/10.1186/s41155-016-0057-1>
- Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *The British Journal of Mathematical and Statistical Psychology*, 38(2), 171-189. <https://doi.org/10.1111/j.2044-8317.1985.tb00832.x>
- Myers, N. D., Pacewicz, C. E., Hill, C. R., & Chun, H. (2023). Factor analysis with ordered categorical indicators and measurement of self-efficacy in physical activity contexts: A substantive-methodological synergy. *Measurement in Physical Education & Exercise Science*, 4(27), 332-351. <https://doi.org/10.1080/1091367X.2023.2186789>
- Neil, R., Fletcher, D., Hanton, S., & Mellalieu, S. D. (2007). (Re)conceptualizing competition stress in sport performers. *Sport and Exercise Psychology Review*, 3, 23-29. <https://doi.org/10.53841/bpssepr.2007.3.2.23>
- Newton, M., Duda, J. L., & Yin, Z. (2000). Examination of the psychometric properties of the Perceived Motivational Climate in Sport Questionnaire-2 in a sample of female athletes. *Journal of Sports Sciences*, 18(4), 275-290. <https://doi.org/10.1080/026404100365018>
- Ng, M. M., Firth, J., Minen, M., & Torous, J. (2019). User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatric Services*, 70(7), 538-544. <https://doi.org/10.1176/appi.ps.201800519>
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328-346. <https://doi.org/10.1037/0033-295X.91.3.328>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402. <https://doi.org/10.3758/BF03200807>
- Pain, M. A., & Harwood, C. (2007). The performance environment of the England youth soccer teams. *Journal of Sports Sciences*, 25(12), 1307-1324. <https://doi.org/10.1080/02640410601059622>
- Pain, M. A., & Harwood, C. G. (2008). The performance environment of the England youth soccer teams: A quantitative investigation. *Journal of Sports Sciences*, 26(11), 1157-1169. <https://doi.org/10.1080/02640410802101835>
- Pain, M. A., Harwood, C., & Mullen, R. (2012). Improving the performance environment of a soccer team during a competitive season: An exploratory action research study. *The Sport Psychologist*, 26(3), 390-411. <https://doi.org/10.1123/tsp.26.3.390>
- Pensgaard, A. M., & Duda, J. L. (2002). "If we work hard, we can do it" a tale from an Olympic (gold) medalist. *Journal of Applied Sport Psychology*, 14(3), 219-236. <https://doi.org/10.1080/10413200290103518>
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). Making sense of factor analysis: The use of factor analysis for instrument development in health care research. Sage. <https://doi.org/10.4135/9781412984898>
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Routledge. <https://doi.org/10.4324/9780203841624>
- Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health*, 14(8), 1101-1108. <https://doi.org/10.1016/j.jval.2011.06.003>
- Roseman, I. J., & Smith, C. A. (2001). Appraisal theory. In K.R. Scherer, A. Schorr & T. Johnstone (Eds). *Appraisal processes in emotion: Theory, methods, research* (pp. 3-19). Oxford University Press. <https://doi.org/10.1093/oso/9780195130072.003.0001>

- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Simms, M., Arnold, R., Turner, J. E., & Hays, K. (2021). A repeated-measures examination of organizational stressors, perceived psychological and physical health, and perceived performance in semi-elite athletes. *Journal of Sports Sciences*, 39(1), 64-77. <https://doi.org/10.1080/02640414.2020.1804801>
- Smith, A. L. (2003). Peer relationships in physical activity contexts: A road less travelled in youth sport and exercise psychology research. *Psychology of Sport and Exercise*, 4(1), 25-39. [https://doi.org/10.1016/S1469-0292\(02\)00015-8](https://doi.org/10.1016/S1469-0292(02)00015-8)
- Smith, S. M., Brown, H., & Cotterill, S. T. (2021). Implementing and Evaluating the Practice Environment Model Using Action Research. *International Sport Coaching Journal*, 9(1), 100-110. <https://doi.org/10.1123/iscj.2020-0117>
- Smith, S. M., Cotterill, S. T., & Brown, H. (2019). A case study of factors influencing performance in the practice environment. *Case Studies in Sport and Exercise Psychology*, 3(1), 33-40. <https://doi.org/10.1123/cssep.2019-0004>
- Smith, S. M., Cotterill, S. T., & Brown, H. (2020a). An interpretative phenomenological analysis of performance influencing factors within the practice environment. *Journal of Physical Education and Sport*, 20(4), 1646-1657. <https://doi.org/10.7752/jpes.2020.04224>
- Smith, S. M., Cotterill, S. T., & Brown, H. (2020b). An Interpretative Phenomenological Analysis of Coach Perceptions in the Practice Environment. *The Sport Psychologist*, 34(4), 257-267. <https://doi.org/10.1123/tsp.2020-0021>
- Smith, R. E., Cumming, S. P., & Smoll, F. L. (2008). Development and validation of the motivational climate scale for youth sports. *Journal of Applied Sport Psychology*, 20(1), 116-136. <https://doi.org/10.1080/10413200701790558>
- Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology*, 99(1), 99-114. <https://doi.org/10.1037/0022-0663.99.1.99>
- Stamenković, D., & Vesković, A. (2022). The association between intensity and direction of competitive anxiety with pre-performance expectations and competition performance. *Facta Universitatis. Series: Physical Education and Sport*, 175-189.
- Tabachnick, B. G. & Fidell, L. S., (2014). *Using multivariate statistics* (New international Ed.) Pearson.
- Wachsmuth, S., Jowett, S., & Harwood, C. G. (2018). On understanding the nature of interpersonal conflict between coaches and athletes. *Journal of Sports Sciences*, 36(15), 1-8. <https://doi.org/10.1080/02640414.2018.1428882>
- Wagstaff, C. R. (2019). Taking stock of organizational psychology in sport. *Journal of Applied Sport Psychology*, 31(1), 1-6. <https://doi.org/10.1080/10413200.2018.1539785>
- Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, S. A. (2020). Development of the short version of the spielberger state-trait anxiety inventory. *Psychiatry Research*, 291, 113223. <https://doi.org/10.1016/j.psychres.2020.113223>
- Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, 6(1), 21-29. <https://doi.org/10.22237/jmasm/1177992180>



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

APPENDIX 1

Sub-Dimension	Description
Control	Athletes who do not aim to directly influence all performance outcomes may perceive less negativity. A focus on the elements that an athlete can have direct influence over in the practice environment (e.g., effort) can lead to perceived positivity and is not affected by performance or outcome
Expected Performance	Athletes who are aware of and reflect on their physical or mental fatigue create more accurate short-term performance goals. Unrealistic short-term goals not adjusted for the current athlete state can cause negativity through realized performance that is lower than the usual standard. Performance goals that are informed by the current state will buffer lower than standard performance causing a decrease in negativity
Support Behaviour	Athletes who feel aided by others within the practice environment will experience greater positivity. Positive communication (e.g., encouragement after failure) leads to greater positivity. Critical communication from teammates and coaches will increase negativity
Team-focused Goals	Athletes who have an individual orientation will experience negativity through a lack of togetherness. Athletes who have a team orientation will benefit from the success of teammates, perceive intra-team competition as challenging rather than threatening, and be able to access support from others in the environment
Ability Ranking	Athletes who create internal skill rankings of all players can cause maladaptive intra-team competition behaviours. Responses to success and failure outcomes against teammates will differ depending on the perceived position of that player in the team
Preparation	Athletes that make life-style decisions based on enhancing performance in practice will experience positivity within the practice environment. A failure to ensure adequate actions are undertaken to recover physically and psychologically from fatigue can cause negative effects.
Challenge	Athletes require both performance success and failure. Following those moments of performance failure, the athlete can reflect and build resilience towards tasks in the future that provide difficulty. Athletes who seek to develop and improve their performance through accepting failure as a process will experience positivity.

APPENDIX 2

Item Rating Task to Assess Content Validity of the *Perceived Performance Climate in Team Practice Scale (PPCTPS)*

Perceived performance climate is defined as the psychological factors perceived by an athlete to influence performance in the team sport practice environment. Performance perceptions are gained from an athlete's identification of both the positive (beneficial) and negative (harmful) influences on performance. The perceived factors of influence are generated from both internal (self) and external (social) sources. Perceptions are unstable and can be affected by time (e.g., different practice sessions), situation (e.g., different practice activities), and case (e.g., different practice environment).

When completing the PPCTPS, each item is anchored by '*Within my current practice environment I...*'

How to Complete the Item Rating

Participants are required to rate the extent to which each item captures each aspect of the domain using a five-point Likert scale ranging from 1 (not at all) to 5 (completely). The top of all columns contains the domain definition with items listed on the left-hand side. An example of a completed item rating task is shown below. In the nutrition example below a banana has been given a rating of 1 (not at all) for 'junk food' but a 5 (completely) for 'healthy snack'. It is recommended you take regular breaks when completing this task over the two-week period.

Example matrix.

Item	<i>Junk food</i> is food that has very little nutritional value.	<i>Dairy products</i> are foods made from animal products.	Foods with sources of fibre are required for healthy gut.	A <i>healthy snack</i> food offers good nutrition.	A high sugar content food raises blood sugar levels.	A high fat content food can raise cholesterol.
Banana	1	1	4	5	4	1
Cake	5	2	3	1	5	5
Fizzy drink	5	1	1	1	5	1

Please indicate on the form (by circling the item) if you believe there to be a spelling or grammar error in an item

Each question item below follows '*Within my current practice environment I...*'

Item	The elements of practice performance that an athlete attempts to control .	The extent to which an athlete thinks about how they expect to perform .	The supportive behaviour received from others.	The extent that performance goals are team-focused goals .	The extent that athletes will ability-rank themselves against teammates.	An environment that creates challenge to build resilience and improve performance.	The life-style behaviours undertaken in preparing to enter the practice environment.
Get annoyed if a teammate makes a mistake							
Prepare by getting adequate sleep							
Cope well with adversity							
Make my teammates feel good							
Work with the team to achieve high performance							
Have team goals that we all work towards							
Criticise my teammates if they make mistakes							
Cheer on my teammates if they do well							
Try to control things I cannot							
Judge my performance against teammate performance							
Ensure I am mentally ready to practise							
Want the team to do well							

APPENDIX 3

Practice Performance Environment Scale

Directions

Thinking about the practice and training activities you undertake for your sport, read the following statements carefully and circle the number that best represents how you feel right now.

Each question begins with: Within my Practice Environment I...	Strongly Disagree					Strongly Agree					Direction
1 Feel supported by coaches	1	2	3	4	5	6	7				Positive
2 Don't want teammates to do better than me	1	2	3	4	5	6	7				Negative
3 Want teammates to fail who are a similar level to me	1	2	3	4	5	6	7				Negative
4 Change the way I compete against teammates depending on their ability level	1	2	3	4	5	6	7				Negative
5 Worry about what my teammates think of my ability	1	2	3	4	5	6	7				Negative
6 Get angry if a superior skilled teammate beats me in a drill	1	2	3	4	5	6	7				Negative
7 Turn failure into future success	1	2	3	4	5	6	7				Positive
8 Perform injury prevention activities	1	2	3	4	5	6	7				Positive
9 Feel supported	1	2	3	4	5	6	7				Positive
10 Drop my effort in drills against lesser skilled teammates	1	2	3	4	5	6	7				Negative
11 Rank my ability against my teammates	1	2	3	4	5	6	7				Negative
12 Have a good relationship with my coach/es	1	2	3	4	5	6	7				Positive
13 Hold the teams goals as more important than my individual goals	1	2	3	4	5	6	7				Positive
14 Prepare myself to perform well	1	2	3	4	5	6	7				Positive
15 Put the team first and myself second	1	2	3	4	5	6	7				Positive
16 Have overcome difficult situations	1	2	3	4	5	6	7				Positive
17 Can rely on my coach to make me feel good even if I don't perform well	1	2	3	4	5	6	7				Positive
18 Change my effort level when competing against different teammates	1	2	3	4	5	6	7				Negative
19 Judge my performance against teammate performance	1	2	3	4	5	6	7				Negative
20 Compare my performance to others	1	2	3	4	5	6	7				Negative
21 Don't want teammates of similar ability to me to do well	1	2	3	4	5	6	7				Negative
22 Perform recovery activities	1	2	3	4	5	6	7				Positive
23 Fully commit myself to team goals	1	2	3	4	5	6	7				Positive
24 Seek out challenge to improve	1	2	3	4	5	6	7				Positive
25 Feel my coach/es talk to me positively	1	2	3	4	5	6	7				Positive
26 See problems as challenges	1	2	3	4	5	6	7				Positive

Scoring

Total Practice Performance Environment Score: Reverse negative direction scores. Add up all question scores and divide by 26.

Specific Sub-Dimension Scoring: Reverse negative direction scores (Teammate Performance, Effort Changes, Hierarchy Focus).

Support Behaviour: $Q1+Q9+Q12+Q17+Q25 / 5$. *Teammate Performance:* $Q2+Q3+Q6+Q21 / 4$. *Effort Changes:* $Q4+Q10+Q18 / 3$.

Hierarchy Focus: $Q5+Q11+Q19+Q20 / 4$. *Challenge:* $Q7+Q16+Q24+Q26 / 4$. *Preparation:* $Q8+Q14+Q22 / 3$. *Team-Focused Goals:* $Q13+Q15+Q23 / 3$.



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

The effect of mental fitness on reducing persecutory thinking among athletes in selected track events

 **Liqa Abdulzahra Obaid**  . College of Physical Education and Sport Sciences for Women. University of Baghdad. Baghdad, Iraq.

ABSTRACT

The present study aimed to examine the effect of mental fitness on reducing persecutory thinking among athletes in selected track events, as well as to develop and validate measurement scales for both variables within a sport-specific context. A descriptive correlational research design was adopted. The research population consisted of male track and field athletes aged 15–17 years from sports clubs in Baghdad during the 2024–2025 competitive season, with a main application sample of 95 athletes. Two scales were constructed for the purposes of the study: a Mental Fitness Scale and a Persecutory Thinking Scale, each comprising 20 items and demonstrating satisfactory validity and reliability indices. Descriptive findings indicated that the athletes exhibited a high level of mental fitness and a low level of persecutory thinking. Inferential analyses revealed a strong, statistically significant inverse relationship between mental fitness and persecutory thinking. Simple linear regression analysis confirmed that mental fitness was a significant negative predictor of persecutory thinking, accounting for a substantial proportion of variance in maladaptive cognitive perceptions related to mistrust and perceived injustice. These results suggest that athletes with higher levels of cognitive regulation, attentional control, and adaptive thinking are less likely to develop negative interpretations of social interactions within the sports club environment. The study underscores the importance of mental fitness as a protective psychological factor in competitive sport and highlights the value of incorporating structured psychological skills training programs to enhance athletes' mental fitness and psychological well-being.

Keywords: Physical activity psychology, Mental fitness, Persecutory thinking, Track and field athletes, Sport psychology, Cognitive regulation.

Cite this article as:

Obaid, L. A. (2026). The effect of mental fitness on reducing persecutory thinking among athletes in selected track events. *Scientific Journal of Sport and Performance*, 5(3), 386–394. <https://doi.org/10.55860/HUKA5365>



Corresponding author. College of Physical Education and Sport Sciences for Women. University of Baghdad. Baghdad, Iraq.

E-mail: liqaa.abdulzahra@copew.uobaghdad.edu.iq

Submitted for publication February 03, 2026.

Accepted for publication March 04, 2026.

Published March 18, 2026.

Scientific Journal of Sport and Performance. ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/HUKA5365>

INTRODUCTION

Psychological dimensions like these have been drawn increasing attention in sports sciences because, as they play a stabilizing role of athletes' performance, their resilience and psychological health. In addition to physical and technical-tactical readiness, athletes' cognitive dimension, in the way that athletes deal with competitive stress or how they successfully relate to others in sport scenarios, also greatly affects (Currie et al., 2021; Rice et al., 2016). In this regard, mental fitness has been considered more and more as a central construct including athletes' abilities to control their thoughts and themselves under pressure (stress) and to stay cognitively flexible in stressful situations; hence, during such conditions they would also easily take good decisions under stress (Gucciardi et al., 2017; Hackfort & Klöppel, 2020).

Psychological functioning including attentional/cognitive control, emotion regulation, reflection and problem-solving is related to mental well-being. Higher mental fitness of athletes is related to higher levels of resilience, increased capability to focus and better coping abilities for evaluation feedback and competition-induced stressors (Siekańska et al., 2021; Bicalho et al., 2020). These cues in both the training and competitive settings, not only make it easier to perform on an athletic arena but also enable better psychological adaptation and social integration. The empirical evidence in applied sport psychology suggests that structured activities based on mental imagery and cognitive training may improve specific motor performance such as the accuracy of decisions made by young athletes (Abdulkareem & Ali Hassan, 2025). This is consistent with the findings of previous study showing that formalized mental training exercises can promote psychological hardiness and adaptive personality traits, such as those used in coping with competitive pressure and psychological demands, among adolescent athletes (Hassan & Abdulkareem, 2025).

On the other hand, dysfunctional cognitive patterns can harm athletes' psychological functioning. Persecutory ideation, an example of such a pattern, is characterized by longstanding beliefs that they are being unfairly victimized, singled out or deliberately injured by others. In current cognitive models, paranoid thinking is imagined as a continuum of thought from mild suspiciousness to chronic persecutory belief (Freeman, 2016; Horita, 2021). Persecutory thinking in sports may be strengthened by social comparison, unfairness perceptions, selection demands, and coach–athlete conflict emerging as negative affect or social withdrawal (Combs et al., 2013).

There is empirical evidence to suggest that cognitive vulnerabilities such as reduced cognitive flexibility and poor reappraisal are involved in the origin and maintenance of persecutory beliefs (Freeman, 2016). Sport psychological solutions to these vulnerabilities involve the enhancement of mental fitness such that improved cognitive control and adaptive thinking support balanced interpretations of social interactions and performance-related feedback (Gupta et al., 2022). Nonetheless, the research on such areas as mental fitness and maladaptive cognitive processes has been increasing, individual associations between these constructs have not been studied in-depth within sport populations.

This deficiency is even more evident when examining track and field athletes, who are subject to great competitive pressure, individual responsibility and performance control. These situational factors could contribute to enhanced vulnerability for maladaptive thinking patterns but may also emphasize the need of mental fitness for keeping psychological balance (Reardon et al., 2023). Furthermore, there is a marked dearth of reliable and valid psychometrics that are available specifically designed to assess mental fitness and persecutory thinking in the context of sport, particularly in athletes from non-Western contexts (Opree et al., 2018).

The construction of sound measurement instruments is a requirement for evidence-based testing and intervention in the field of sport psychology. Strong psychometrical tools allow for the detection of cognitive risk factors, psychological monitoring and personalized mental training programs specifically tailored to improve resilience and mitigate maladaptive thinking (Gupta et al., 2022; Siekańska et al., 2021).

Accordingly, the present study aimed to (1) develop and validate scales measuring mental fitness and persecutory thinking among athletes in selected track events, (2) determine the levels of these constructs within the study sample, and (3) examine the relationship and predictive effect of mental fitness on persecutory thinking. It was hypothesized that mental fitness would be significantly and negatively associated with persecutory thinking, such that higher levels of mental fitness would contribute to lower levels of persecutory cognitive tendencies among track athletes.

METHODS

Research design

The design used in the present research was descriptive research with correlational analysis aiming at explaining a relationship and predicting power between mental fitness and persecutory thinking among track and field athletes. When there was no physical modification, this design has been deemed suitable for the study of psychological variables in the absence of experimental manipulation.

Participants

The current study was conducted on all male track athletes participating in running events (100, 200, 400 and 800 m races) and from different sports clubs in Baghdad for the season of 2024–2025. The final sample comprised 206 athletes between ages 15 and 17 years who had trained for a period of time of between 5 and 6 years. A census sample design was used, with the entire available population.

The sample was divided into three sub-samples according to the objectives of the study:

- Pilot sample: 6 athletes (2.9%) selected randomly to test the clarity and feasibility of the measurement tools.
- Scale construction sample: 105 athletes (51.0%) selected randomly to establish the psychometric properties of the two scales.
- Main application sample: 95 athletes (46.1%) selected randomly for the final application and hypothesis testing.

Ethical approval

The study was conducted in compliance with the ethical guidelines of scientific research and received approval from the Scientific and Ethical Committee of the College of Physical Education and Sports Sciences, University of Baghdad. The involvement was voluntary and all athletes provided written consent before data collection. All respondents were guaranteed that their confidentiality was strictly protected and data were solely used for research purposes.

Instruments

Mental fitness scale

The development of the Mental Fitness Scale was needed because no scales for track and field were available. The questionnaire consisted of 20 items that were created to assess cognitive regulation, attentional control and decision-making as well as adaptive thinking in both training and competition situations. Responses were scored on a three-point scale (3 = Always, 2 = Sometimes, 1 = Does not apply),

and higher scores indicate greater mental fitness. The scores could range from 20-60, assuming a hypothetical Mean of 40, as shown in Table 1.

Table 1. Mental fitness scale for track and field athletes.

No.	Item statements
1	I avoid rushing to express my opinion in situations I encounter at the sports club.
2	I focus my thinking on factors that contribute to success in training and competition tasks at the sports club.
3	I am able to continuously train my mind to develop perception in response to various situations I face at the sports club.
4	I can train my mind to understand viewpoints that differ from my own at the sports club.
5	I am able to express my point of view with humility at the sports club.
6	I find it easy to recall thoughts that support coping with challenging situations at the sports club.
7	I direct my senses toward the most important stimuli related to situations I encounter at the sports club.
8	I show readiness to listen attentively to different opinions at the sports club.
9	I am aware of properly performing each task assigned to me at the sports club.
10	I find it easy to deal with solutions to my problems at the sports club.
11	I activate my thinking to compensate for forgetting the coach's instructions at the sports club.
12	I listen carefully to all instructions provided by the coach at the sports club.
13	I respect the opinions of my fellow athletes at the sports club and do not interrupt them.
14	I continuously develop my thinking to respond effectively to various situations at the sports club.
15	I am able to define my future goals within the Table 1 Table 1e sports club.
16	I make sure not to rush when making decisions at the sports club.
17	I mentally prepare myself to receive information related to various situations at the sports club.
18	I organize information about different situations in my memory at the sports club.
19	I remain alert to the tasks required by the current situation at the sports club.
20	I pay attention to evaluating my decisions at the sports club.

Persecutory thinking scale

The Persecutory Thinking Scale was also developed to measure sense of injustice, distrust and hostility in the sport club setting. A 20-item scale was used and rated according to the three-point Likert scale. Increasing scores indicated greater persecutory thinking, with a possible total score of 20–60 (hypothetical mean = 40) as shown in Table 2.

Table 2. Persecutory thinking scale for track and field athletes.

No.	Item statements
1	I constantly feel that coaches evaluate my training performance unfairly at the sports club.
2	I feel that fate alone is what brought me to this sports club.
3	I believe that everyone at the sports club is hypocritical.
4	I perceive sincerity as completely absent within the sports club.
5	I believe that noble values are gradually disappearing within the sports club.
6	I feel that my fellow athletes at the sports club were imposed on me by circumstances.
7	I believe that the sports club exploits me without justification.
8	I feel that my fellow athletes hold negative assumptions about me at the sports club.
9	I feel that my fellow athletes attempt to control my decisions at the sports club.
10	I feel that fairness is lacking among athletes at the sports club.
11	I feel that my fellow athletes mock or ridicule me at the sports club.
12	I believe that my fellow athletes seek to hinder my success at the sports club.
13	I feel rejected by my fellow athletes at the sports club.
14	I feel a lack of trust toward my fellow athletes at the sports club.
15	I feel that the smiles of my fellow athletes at the sports club are insincere.
16	I believe that the sports community within the club is entirely deceitful.
17	I believe that my coach deliberately undermines my value at the sports club.
18	I believe that the sanctions imposed on me by the club are unjust.
19	I believe that justice is merely a word with no real meaning at the sports club.
20	I feel that my achievements are not respected at the sports club.

Scale development and validation

Item discrimination was tested via the extreme group's strategy comparing upper and lower 27% of the scores in participants. Analysis revealed that for all items, of the strategy and the persecutory thinking scale, there was a significant difference between groups with t-values ranging from 2.82 to 15.10 ($p < .001$). The internal consistency was checked using the item-total correlation and coefficients of ranging from .49 to .82 for the Mental Fitness Scale, .49 to .90 for the Persecutory Thinking Scale, which suggests adequate homogeneity of items for both of scales. Reliability The Cronbach's alpha coefficients of the reliability estimates were. 858 for the Mental Fitness Scale. Enclitic and 0.866 for the Persecutory Thinking Scale, with high internal reliability. Inspection of normality on the score distribution yielded satisfactory levels of skewness (0.10 for Mental Fitness and 0.27 for Persecutory Thinking), thus confirming that the data satisfied assumptions underlying subsequent parametric testing.

Data collection procedures

Upon completion of the validation, the final versions of both scales were administered to the main sample ($n = 95$) at their respective club centres between December 15 and January 9. In both groups, participation was voluntary and athletes were explained the purpose of the study before data collection. All participants were given an informed consent form before collecting the data.

Statistical analysis

Statistical calculations were performed by the use of Statistical Package for the Social Sciences (SPSS) program. Descriptive statistics (means, SDs and percentages) were calculated. Inferential analyses comprised independent-sample t-tests, one-sample t-tests, Pearson correlation coefficients and simple linear regression analysis regarding the association and predictive role of mental fitness with persecutory thinking. The threshold level of significance was at $p \leq .05$.

RESULTS

As in Table 3, we can see that the athletes are highly mentally fit because their scores were higher than the hypothetical mean. In contrast, persecutory thinking was low, with scores below the sample mean. This trend seems to mean that the athletes have developed overall effective cognitive and adaptive mental regulations, having a little feeling of injustice, suspicion or hostility in sports club environment.

Table 3. Comparison of the observed mean with the hypothetical mean for the study variables.

Scale	Total Score	Hypothetical Mean	Mean	SD	Mean Difference	t	Sig.	Significance
Mental fitness	60	40	52.99	3.956	12.989	32.005	<.001	Significant
Persecutory thinking	60	40	33.79	1.637	-6.211	36.986	<.001	Significant

Note. Measurement unit = score. Differences were tested using a one-sample t-test ($df = 94$). Statistical significance was set at $p \leq .05$.

Table 4. Simple linear regression results between mental fitness and persecutory thinking.

Predictor	Outcome	Sample Size (n)	Pearson's r (R)	R ²	Contribution (%)	Standard Error of Estimate
Mental fitness	Persecutory thinking	95	.901	.811	80.9	.715

Table 4 The relationship between mental fitness and persecutory thinking among the track and field athletes The Results showed that there was a significant correlation between those two. The sign of this relationship is such that greater mental fitness is related to lower levels of persecutory thinking. The present results

emphasize the role of state mental fitness as a fundamental psychological factor associated with healthier cognitive appraisals and less negative attributions to social interactions in sports.

As shown in Table 5, the result validates that the simple linear regression model developed in this study is suitable and statistically reasonable. The large goodness-of-fit statistics suggest that the model adequately fits the data on the mental fitness-persecutory thinking relations. This is supportive of the employment of mental fitness as a predictor variable for explaining mental health differences in persecutory thinking among athletes.

Table 5. Goodness-of-fit test for the simple linear regression model.

Predictor	Outcome	Source of Variance	Sum of Squares	df	Mean Square	F	Sig.	Significance
Mental fitness	Persecutory thinking	Regression	204.308	1	204.308	400.169	<.001	Significant
		Error	47.482	93	0.511			

Note. The F-value was statistically significant at $p \leq .05$, indicating good model fit.

The regression calculated values in Table 6 suggest that mental fitness has a strong negative impact on persecutory thinking. This suggests that improvements in mental fitness play a meaningful role in the reductions of maladaptive cognitive processes that are associated with such perceived injustice and mistrust. The results highlight the fact that mental fitness plays a protective role in achieving psychological equilibrium and cognitive stability under competitive sports conditions.

Table 6. Regression coefficients for the effect of mental fitness on persecutory thinking.

Outcome variable	Predictors	β (Beta)	SE	t	Sig.	Significance
Persecutory thinking	Constant	53.538	.990	54.083	<.001	Significant
	Mental fitness	-0.373	.019	20.004	<.001	Significant

Note. Regression coefficients were tested at $p \leq .05$.

DISCUSSION

The current study showed that athletes who specialized in some track events had high mental fitness and low persecutory thinking, and the predictive power of mental fitness on persecutory thinking is significantly negative. These results suggest that high cognitive regulation, attentional control and positive thinking can serve a protective psychological factor against maladaptive social-cognitive interpretations in sport.

The athletes' superior mental fitness might be attributed to the nature of track and field training, which involves concentration, self-control, goal-directed thinking and constant monitoring of performance. Such cognitive demands are found to even promote executive functioning and cognitive flexibility, which are thought to lower the risk of experiencing distorted interpretations of interpersonal events (Klein et al., 2018). If an athlete can control his attention and assess situations without bias, then he is less likely to perceive ambiguous input from coaches or teammates as hostile or unfair.

Consistent with cognitive models, which predict that individual's ability to regulate and interpret emotion will lead to a reduction in paranoid thinking, the low level of persecutory thinking observed in this study resonates well with such models. Cognitive flexibility facilitates thinking about alternative causes to social events, thus reducing the tendency of assigning negative intentions to others (Bentall et al., 2012). This could account for the fact that athletes high on mental fitness experienced less perception of injustice, mistrust or hostility in their sport club.

The inverse association between mental fitness and paranoid thinking contributes to theoretical models in cognitive psychology and sport psychology which take a focus on the contribution of mental skills to social cognition. Social perceptual disorganization, excessive suspiciousness and thought disturbance have been related to cognitive dysfunction (Freeman et al., 2002). Therefore, the present results suggest that mental fitness is associated with not only performance but also social coping and adaptive response to a competitive environment.

Furthermore, with respect to persecutory ideation, mental fitness also showed substantial amount of variance; therefore, has predictive validity as well. This is also consistent with previous research that has shown psychological skills training involving changing one's cognitive bias (e.g., reappraisal), goal-setting, and attentional focus can reduce maladaptive cognitive biases and enhance emotional control (Gross, 2015; Nicholls et al., 2016). Athletes who are cognitively prepared for ambiguity and emergence will be better equipped to manage the requirements of performance pressure, social complexity in competitive sport.

From an environmental perspective, the club-level climate has a significant impact on athletes' cognitive evaluations. Research on the coach-athlete relationship indicates that perceived fairness, trust, and good communication result in less negative attributional style among athletes (Jowett & Shanmugam, 2016). As a consequence, the low level of persecutory thinking observed in the present sample could possibly be accounted for either by supportive club environments or/and by some kind of attenuation related to mental fitness on relational stressors.

This study's findings also have important practical implications. Considering that mental fitness is a modifiable psychological variable, interventions to improve attentional control, cognitive flexibility and emotion regulation could probably lead toward reduction of persecutory thinking among athletes. Cognitive-bias modification and mindfulness-based training have effectively decreased maladaptive interpretation biases and increased mental well-being (see Cristea et al., 2015). The inclusion of such methods in standard training regimes may benefit athletic performance, and mental health.

However, the study has certain restrictions. The correlational nature of the design means that causal inferences cannot be made, we require longitudinal or experimental research to establish whether increasing mental fitness indeed has a lasting impact on persecutory ideation. Moreover, the sample was limited to male adolescent athletes within a particular culture, which may limit the generalizability of results. Subsequent studies should be conducted in various age ranges, female athletes and sports components as well as with intervention designs.

Practical implications

Coaches and sport psychologists may benefit from implementing structured mental-skills training (attentional control, cognitive reappraisal, goal-directedness, and brief mindfulness) throughout regular training cycles while applying relationship-related practices that communicate a sense of justice and respect. Such systemic and individual -level approaches are likely to enhance not only performance-relevant cognitive processes, but also reduce maladaptive social-cognitive tendencies among athletes

CONCLUSIONS

The results of the current study suggest that the athletes in selected track events have a high mental fitness as well as low level of persecutory ideation. This finding is indicative of the capacity for efficient cognitive

processing and adopted skilful thinking habits among the athletes, which in turn seem to be associated with healthy psychological functioning within a sports club context.

Results also provide further evidence for a strong inverse association between mental fitness and persecutory ideation, indicating that the former could be a prominent predictor of reduction in maladaptive cognitive appraisals. Those athletes with greater levels of Mental Fitness reported that appraisals of their social accountant ability and of the evaluative situation were less likely to be perceived as threatening or unfair. The present study generally reiterates the significance of mental fitness for prevention at highly competitive levels. Increased sports psychological fitness might then not only improve athletic mastery but contribute to good mental health and positive social cognition in the athletes.

Recommendations

From these results, coaches and sport psychologists should have an interest in the mental fitness of their athletes, imparting them formal psychological skills training programs. Such interventions should focus on improving attentional control, cognitive flexibility and adaptive decision making.

Taking method for the prevention of cognitive disfunctions in athletes by looking at psychological monitoring of mental and persecutory thinking also, intermittent periodic checking on mental health and persecutory thinking can be a significant tool to establish cognitive health as well as early detection of their maladaptive theories. Thus, establishing supportive and fair training environments could decrease negative appraisals along with foster of psychological need satisfaction in athletes.

SUPPORTING AGENCIES

No funding agencies were reported by the author.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

REFERENCES

- Abdulkareem, O. W., & Ali Hassan, M. F. (2025). The impact of mental games on improving shooting accuracy among young basketball players in Iraqi clubs. *Scientific Journal of Sport and Performance*, 4(3), 342-351. <https://doi.org/10.55860/OHNP7224>
- Bentall, R. P., Wickham, S., Shevlin, M., & Varese, F. (2012). Do specific early-life adversities lead to specific symptoms of psychosis? A study from the 2007 the Adult Psychiatric Morbidity Survey. *Schizophrenia bulletin*, 38(4), 734-740. <https://doi.org/10.1093/schbul/sbs049>
- Bicalho, C. C. F., Melo, G. F., & Noce, F. (2020). Resiliencia de los atletas: una revisión sistemática basada en un análisis de redes de citas. *Cuadernos de Psicología del Deporte*, 20(3), 26-40. <https://doi.org/10.6018/cpd.391581>
- Combs, D. R., Finn, J. A., Wohlfahrt, W., Penn, D. L., & Basso, M. R. (2013). Social cognition and social functioning in nonclinical paranoia. *Cognitive neuropsychiatry*, 18(6), 531-548. <https://doi.org/10.1080/13546805.2013.766595>
- Cristea, I. A., Mogoase, C., David, D., & Cuijpers, P. (2015). Practitioner review: Cognitive bias modification for mental health problems in children and adolescents: A meta-analysis. *Journal of Child Psychology and Psychiatry*, 56(7), 723-734. <https://doi.org/10.1111/jcpp.12383>

- Currie, A., Blauwet, C., Bindra, A., Budgett, R., Campriani, N., Hainline, B., ... & Goutteborge, V. (2021). Athlete mental health: future directions. *British Journal of Sports Medicine*, 55(22), 1243-1244. <https://doi.org/10.1136/bjsports-2021-104443>
- Freeman, D. (2016). Persecutory delusions: a cognitive perspective on understanding and treatment. *The Lancet Psychiatry*, 3(7), 685-692. [https://doi.org/10.1016/S2215-0366\(16\)00066-3](https://doi.org/10.1016/S2215-0366(16)00066-3)
- Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, 41(4), 331-347. <https://doi.org/10.1348/014466502760387461>
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological inquiry*, 26(1), 1-26. <https://doi.org/10.1080/1047840X.2014.940781>
- Gucciardi, D. F., Hanton, S., & Fleming, S. (2017). Are mental toughness and mental health contradictory concepts in elite sport? A narrative review of theory and evidence. *Journal of science and medicine in sport*, 20(3), 307-311. <https://doi.org/10.1016/j.jsams.2016.08.006>
- Gupta, S., & McCarthy, P. J. (2022). The sporting resilience model: A systematic review of resilience in sport performers. *Frontiers in psychology*, 13, 1003053. <https://doi.org/10.3389/fpsyg.2022.1003053>
- Hackfort, D., & Klöppel, Y. P. (2020). Mental fitness. In *The Routledge International Encyclopedia of Sport and Exercise Psychology* (pp. 249-267). Routledge. <https://doi.org/10.4324/9781315187228>
- Hassan, M. F. A., & Abdulkareem, O. W. (2025). The Effect of Mental Training on Psychological Hardiness and Selected Personality Traits among Adolescent Male Volleyball Players. *International Journal of Exercise Science*, 18(4), 1186. <https://doi.org/10.70252/MQUH7716>
- Horita, Y. (2021). Conjecturing harmful intent and preemptive strike in paranoia. *Frontiers in Psychology*, 12, 726081. <https://doi.org/10.3389/fpsyg.2021.726081>
- Jowett, S., & Shanmugam, V. (2016). Relational coaching in sport: Its psychological underpinnings and practical effectiveness. In *Routledge international handbook of sport psychology* (pp. 471-484). Routledge.
- Klein, H. S., Kelsven, S., & Pinkham, A. E. (2018). Increased social cognitive bias in subclinical paranoia. *Schizophrenia research: cognition*, 12, 74-76. <https://doi.org/10.1016/j.scog.2018.05.002>
- Nicholls, A. R., Morley, D., & Perry, J. L. (2016). Mentally tough athletes are more aware of unsupportive coaching behaviours: Perceptions of coach behaviour, motivational climate, and mental toughness in sport. *International Journal of Sports Science & Coaching*, 11(2), 172-181. <https://doi.org/10.1177/1747954116636714>
- Oprea, S. J., Buijzen, M., & Van Reijmersdal, E. A. (2018). Development and validation of the psychological well-being scale for children (PWB-c). *Societies*, 8(1), 18. <https://doi.org/10.3390/soc8010018>
- Siekańska, M., Bondar, R. Z., di Fronso, S., Blecharz, J., & Bertollo, M. (2021). Integrating technology in psychological skills training for performance optimization in elite athletes: A systematic review. *Psychology of Sport and Exercise*, 57, 102008. <https://doi.org/10.1016/j.psychsport.2021.102008>



The relationship between the special judo fitness test and field testing

-  **Ryan Cedeno**  . College of Health Professions & Medical Sciences. Department of Health Sciences and Clinical Practice. Barry University. Miami Shores, Florida, United States of America.
-  **Tal Amasay**. College of Health Professions & Medical Sciences. Department of Health Sciences and Clinical Practice. Barry University. Miami Shores, Florida, United States of America.
-  **Zacharias Papadakis**. College of Health Professions & Medical Sciences. Department of Health Sciences and Clinical Practice. Barry University. Miami Shores, Florida, United States of America.
- Claire Egret**. College of Health Professions & Medical Sciences. Department of Health Sciences and Clinical Practice. Barry University. Miami Shores, Florida, United States of America.
-  **Amariliss Owens**. College of Health Professions & Medical Sciences. Department of Health Sciences and Clinical Practice. Barry University. Miami Shores, Florida, United States of America.

ABSTRACT

Judo, an Olympic sport, uses the Special Judo Fitness Test (SJFT) to predict an athlete's competition readiness. However, the SJFT's high anaerobic technical demands may increase pre-competition injury risk. Purpose: To identify field tests that could predict performance on the SJFT, providing alternative ways to evaluate an athlete's performance readiness. Methods: Ten judo athletes (age: 33.8 ± 8.3 years; height: 171.5 ± 3.7 cm; weight: 79.5 ± 9.7 kg) performed eight field tests, including pro-agility, six-meter timed hop, modified handgrip strength, maximum push-ups, modified Gi flexed-arm hang, upper quarter Y-balance, maximum sit-ups, and 300-yard shuttle run tests. Pearson correlations and backward stepwise regression analyses were used to determine the relationship between the field tests and the SJFT score and index, $p = .05$. Results: The SJFT score was associated with the pro-agility ($r = -.70$) and the six-meter hop ($r = -.73$) and the SJFT index was associated with the pro-agility ($r = .73$) and the 300-yard shuttle run ($r = .73$), $p < .05$. A backward stepwise regression showed that the six-meter hop predicted the SJFT score, $R^2 = .52$, $p = .018$, while the 300-yard shuttle run predicted the SJFT index, $R^2 = .53$, $p = .017$. Conclusion: The six-meter hop and 300-yard shuttle tests can serve as measures for identifying an athlete's readiness for SJFT performance or, alternatively, their preparedness for competition.

Keywords: Performance analysis, Agility, Endurance, Anaerobic, Aerobic, Sport performance.

Cite this article as:

Cedeno, R., Amasay, T., Papadakis, Z., Egret, C., & Owens, A. (2026). The relationship between the special judo fitness test and field testing. *Scientific Journal of Sport and Performance*, 5(3), 395-409. <https://doi.org/10.55860/DBNQ2927>



Corresponding author. Department of Health Promotion & Clinical Practice. College of Health & Wellness Barry University Miami Shores, Florida, 33161 United States of America.

E-mail: rycedeno@barry.edu

Submitted for publication January 14, 2026.

Accepted for publication February 25, 2026.

Published March 18, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/DBNQ2927>

INTRODUCTION

Judo involves unique movement patterns that include various throwing and ground techniques in dynamic situations where athletes' grips and positions constantly change (Gardasevic & Stankovic, 2019; Sertić & Segedi, 2012). Success in judo competitions requires athletes to have advanced technical and tactical skills and a high level of physical fitness. Athletes optimize their techniques to conserve energy and exploit their opponent's weaknesses (Sterkowicz et al., 1997). A standard judo match is scheduled for four minutes, but if it's a draw, an unlimited golden score period begins (International Judo Federation, 2023). High-level matches typically last about three minutes, featuring short bursts of activity (20–30 seconds) with brief rest periods (5–10 seconds) (Castarlenas & Planas, 1997; Van Malderen et al., 2006). Judo depends mainly on anaerobic metabolism to provide the necessary strength, speed, and power for executing techniques, while aerobic metabolism supports intermittent efforts and recovery during pauses in the match (Franchini et al., 2009).

Coaches and sports science professionals routinely utilize numerous physical fitness evaluations to identify talented athletes across various age groups. However, many of these tests are conducted in laboratory settings, and some experts argue that they do not closely resemble the competitive performance and sport-specific movement patterns of judo athletes (Drid et al., 2012; Hesari et al., 2014). Therefore, Sterkowicz and Franchini (1995) introduced the special judo fitness test (SJFT) to mimic a judo match and monitor athletes for talent and progress.

The SJFT consists of one 15-second bout and two 30-second bouts of throwing, separated by 10-second rest intervals. During the test, the athlete, positioned three meters away from two assistants spaced six meters apart, performs as many ippon seoi-nage (shoulder throw) techniques as possible. The total number of throws from each bout is recorded to determine the SJFT score. Heart rate measurements are taken immediately post-test and one minute later to calculate the SJFT index, which provides insight into cardiovascular recovery (Sterkowicz & Franchini, 1995). This test reflects judo's energy demands, with approximately 42% of energy derived from anaerobic sources (Franchini et al., 2011). Elite judo athletes typically execute more throws and achieve a lower SJFT index than novices, making the test a reliable performance indicator (Sterkowicz & Franchini, 2001).

According to Arazi et al. (2017), there has been limited focus on examining the correlation between a full range of required physical fitness assessments for a successful judo contest and an athlete's performance in the SJFT. Their study included 50 senior male judo athletes with at least two years of intermediate to advanced judo experience. The study aimed to clarify the relationship between anthropometric and physical fitness assessments and the SJFT index. Notably, there were negative correlations with hand grip strength ($r = -.39$), vertical jump ($r = -.34$), and 35-meter sprints with 10-second rest intervals ($r = -.38$). Positive correlations were found with 45-meter sprint speed ($r = .37$) and agility as measured by the 4 x 9-meter shuttle run ($r = .35$) with $p < .05$. Similarly, a study by Ceylan et al. (2022) found a relationship between physical fitness assessments and the SJFT index in a sample of 19 elite senior judo athletes (11 female). They found correlations between the SJFT index and 20-meter sprint speed ($r = .74$), agility T-test ($r = .77$), and vertical jump ($r = -.74$) with $p < .001$. However, there was no relationship between grip strength and the SJFT index. These findings highlight the complex relationship between various physical assessments and judo-specific fitness as assessed by the SJFT.

A limited number of studies have focused on examining the correlation between a dynamic and sport-specific set of physical fitness assessments required for a successful judo contest and performance in the SJFT.

Most of the studies primarily investigated the correlation between the SJFT and factors such as aerobic power and anaerobic capacity, emphasizing energy systems in laboratory settings (Hesari et al., 2014; Franchini et al., 2011; Franchini et al., 2005; Detanico et al., 2012). To the best of our knowledge, no studies involving senior judo athletes have correlated field testing with the SJFT score, and only two studies have correlated field testing with the SJFT index.

Therefore, given the limited evidence regarding the relationship between field tests and the SJFT score and index, as well as the potential to reduce the risk of injury associated with frequent use of the SJFT, this study aims to identify alternative judo-specific field tests that correlate with the SJFT score and index. It is hypothesized that these field tests will serve as significant predictors of both the SJFT score and index.

MATERIALS AND METHODS

Participants

Ten judo athletes (age: 33.8 ± 8.3 years; height: 171.5 ± 3.7 cm; weight: 79.5 ± 9.7 kg; total years of practice: 13.3 ± 9.0) participated in this study. Following the provision of informed consent, participants completed a questionnaire to assess their eligibility. The inclusion criteria for participants were as follows: individuals with no existing medical conditions and no reported injuries in the previous six months. Participants were required to hold a minimum rank of orange belt in judo and have at least two years of judo experience. Prior to the actual data collection session, the height, weight, and body fat percentage of participants were recorded. The study received approval from the Institutional Review Board Committee at Barry University. All procedures adhered to the ethical standards outlined in the Declaration of Helsinki.

Measures of judo testing

The SJFT was conducted as described by Sterkowicz and Franchini Sterkowicz and Franchini (1995) (Figure 1). Athletes performed one 15-second and two 30-second bouts of ippon seoi-nage throws, with 10-second breaks in between. Heart rates were measured immediately after testing and one-minute post-testing. The number of correct technical throws from the three bouts was summed to determine the SJFT score. This score was then used to calculate the SJFT index, using the following formula: $[\text{Final HR (bpm)} + \text{HR 1 minute after the test (bpm)}] / [\text{sum of total throws}]$.

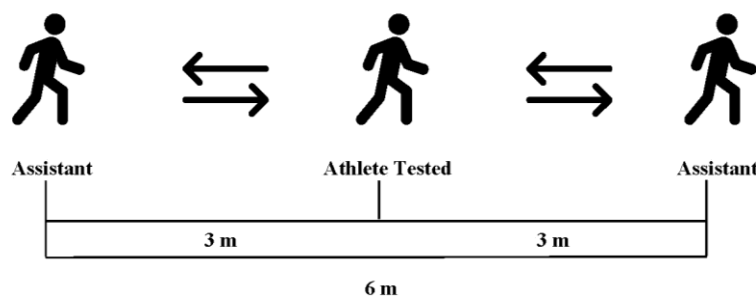


Figure 1. The special judo fitness test.

Measures of field testing

Athletes performed the pro agility test according to specific guidelines (Tomchuk, 2011). Two total trials were performed in opposite starting directions with a 2-minute rest between trials and the fastest trial was used for analysis. In the six-meter timed hop test, participants hopped forward as swiftly as possible using the same leg. Each leg underwent three trials, with the fastest overall time recorded for analysis (Tomchuk, 2011).

Handgrip strength was assessed according to Tomchuk, modifying the protocol with participants replicating the ippon seoi-nage (Shoulder Throw) technique while exerting pressure on the dynamometer. Three attempts were made with each hand, and the highest value for each hand was used for analysis. This method aims to measure handgrip strength in a more dynamic and sport-specific movement. The push-up test measured the maximum number of push-ups within one minute, maintaining a plank position, lowering their chest to touch a balance pad, and pushing back up (Tomchuk, 2011). The modified Gi flexed arm-hang test was conducted according to (Franchini et al., 2011). One trial was performed unless the participant opted for a second attempt, in which case a 2-minute break was given. In the sit-up test, the participants performed as many sit-ups as possible within one minute (Tomchuk, 2011). The upper quarter Y-balance test (UQYBT) was performed according to (Westrick et al., 2012) in a randomized sequence. The process was repeated two additional times with each arm. The composite reach score for each arm was used for analysis. Finally, athletes performed the 300-yard shuttle run test, and the final score used for analysis was the average of the two trials (Tomchuk, 2011).

Procedures

All athletes were assessed in a single session. The testing sequence adhered to the National Strength and Conditioning Association's (NSCA) Testing Guidelines, which recommend conducting non-fatiguing tests first, followed by agility tests, power and maximal strength assessments, muscular endurance evaluations, and concluding with anaerobic and aerobic capacity tests (McGuigan, 2016). A three-minute recovery time was allotted between each test to prevent fatigue. Testing was conducted at the judo athlete's respective dojo.

Before testing, participants completed a judo-specific warm-up on the judo mat, incorporating various movements designed to prepare them for performance. The warm-up began with general movements to increase blood flow and flexibility, progressing to more judo-specific exercises. To conclude the warm-up and enhance their readiness for the SJFT, athletes practiced ukemi (break falls). Following this, they engaged in five minutes of partner-assisted moving uchikomi (repetition training), with a focus on the ippon seoi-nage technique. This progression ensured that the athletes were thoroughly prepared for the demands of the SJFT.

Analysis

A power analysis was conducted using G*Power (version 3.1.9.7) (Faul et al., 2020). We selected the F-test family with a linear multiple regression fixed model (R^2 deviation from zero) and specified an effect size of .35, an alpha level of .05, a power of .80, and eight predictors. This analysis indicated a requirement for a sample size of 52 participants to adequately power a regression model incorporating eight predictors.

In consideration of the exploratory nature of this study, Pearson product-moment correlation analyses were conducted to identify significant relationships between the independent and dependent variables. Independent variables that demonstrated significant correlations with the SJFT score or index were chosen for inclusion in separate regression models. Following this initial correlation analysis, a subsequent power analysis was performed using the aforementioned parameters. This analysis only included variables that significantly correlated with the SJFT score or index as predictors in the models, directly addressing the research question.

Descriptive statistics were presented as means (M) and standard deviations (SD) for data that were normally distributed, or as medians and interquartile ranges (IQR) for data that were not normally distributed. As previously mentioned, two distinct Pearson correlation analyses were conducted to examine the relationships between the independent variables and the SJFT score and index, respectively. Variables that demonstrated

significant correlation with the SJFT score or index were included in two backward stepwise regression analyses; one aimed at predicting the SJFT score, and the other at predicting the SJFT index. All statistical analyses were performed with IBM SPSS Statistics, version 30.0 (IBM Corp., 2024), maintaining a significance level of $p < .05$.

RESULTS

As previously noted, the initial power analysis determined a required sample size of 52 participants. However, following a subsequent power analysis, this requirement was adjusted to 31 participants. Given the exploratory nature of the study, which required participants of similar height and weight, a convenience sample of 10 participants was employed. Descriptive statistics and the normal distribution of the data were verified, with the results presented in Table 1.

Table 1. Descriptive statistics of participants' field test data including CI, SD, and Shapiro-Wilk test results.

Variable	Mean	95% CI Lower	95% CI Upper	SD	Shapiro-Wilk W	Shapiro-Wilk p
Age	33.8	27.9	39.7	8.27	.95	.64
Height (cm)	171.5	168.8	174.2	3.72	.95	.65
Weight (kg)	79.5	72.6	86.4	9.70	.99	1.00
Body Fat (%)	16.6	13.1	20.0	4.80	.86	.08
Training Years	13.3	6.9	19.7	8.98	.94	.53
SJFT Index	15.4	13.0	17.8	3.36	.85	.06
Pro-Agility (s)	5.7	5.2	6.1	0.67	.95	.64
6-meter Hop Test (s)	2.1	2.0	2.3	0.26	.98	.95
Hand Grip Strength	48.2	43.8	52.6	6.09	.95	.71
Push-ups (s)	47.4	37.3	57.5	14.20	.93	.48
GI Hang (s)	29.4	16.8	42.0	17.7	.93	.48
UQYBT Composite Score RH	1.02	0.9	1.1	0.11	.93	.48
UQYBT Composite Score LH	1.02	0.9	1.1	0.11	.93	.48
Sit-ups (s)	46.4	39.8	53.0	9.19	.94	.56
300-yard Shuttle (s)	76.3	71.7	80.8	6.36	.94	.56

Note. CI = Confidence Interval, UQYBT = Upper Quarter Y-Balance Test, SL = Superolateral, M = Medial, IF = Inferolateral, RH = Right Hand, LH = Left Hand.

Correlations

A Pearson product-moment correlation was employed to identify relationships between the SJFT score, SJFT index, and field testing. Positive correlations were observed between the SJFT score and the pro agility and six-meter hop tests, as shown in Table 2. Negative correlations were found between the SJFT index and the pro agility and 300-yard shuttle tests, as presented in Table 3. These correlated variables were subsequently utilized to predict the SJFT score and index.

Table 2. Correlation between field tests and the SJFT score.

Field test	r	p-Value
Pro-agility	-.70	.025*
6m Hop Test	-.73	.018*
Hand Grip Strength	-.06	.864
Maximum Push-ups	.07	.839
Maximum GI Hang	.22	.543
Maximum Sit-ups	.49	.154
300-Shuttle	-.50	.143
UQYBT Composite Score Right Hand	.36	.302
UQYBT Composite Score Left Hand	.14	.700

Note. UQYBT = Upper Quarter Y-Balance Test. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3. Correlation between field tests and the SJFT Index.

Field test	r	p-Value
Pro-agility	.73	.017*
6m Hop Test	.63	.053
Hand Grip Strength	.25	.488
Maximum Push-ups	-.27	.444
Maximum GI Hang	.001	.998
Maximum Sit-ups	-.58	.078
300-Shuttle	.73	.017*
UQYBT Composite Score Right Hand	-.40	.249
UQYBT Composite Score Left Hand	-.36	.309

Note. UQYBT = Upper Quarter Y-Balance Test. * $p < .05$, ** $p < .01$, *** $p < .001$.

Backward stepwise regression - SJFT score

Negative correlations were identified between the SJFT score and both the pro agility and six-meter hop tests. The pro agility and six-meter hop tests also did not significantly correlate with each other ($r = .63$, $p = .052$) (Table 5). In the initial model, which encompassed both the pro agility and six-meter hop tests, the overall model was significant ($p < .005$) (Table 4). However, when assessing the individual contribution of the predictors, neither was significant in the overall model (Table 5). Further analysis showed that model two, which included only the six-meter hop test, as an individual predictor, was also significant ($p < .005$) (Table 5). Therefore, the SJFT score can be predicted using the six-meter hop test ($R^2 = .52$, $p = .018$). The regression equation is: $SJFT \text{ score} = 43.128 - 10.106 \times (\text{Six-meter Hop})$.

Table 4. The special judo fitness test score Anova.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	68.379	2	34.190	5.682	.034 ^a
	Residual	42.121	7	6.017		
	Total	110.500	9			
2	Regression	57.606	1	57.606	8.713	.018 ^b
	Residual	52.894	8	6.612		
	Total	110.500	9			

Note. a. Predictor: six-meter hop, pro-agility; b. Predictor: six-meter hop.

Table 5. The special judo fitness test score coefficients.

Model		Unstd. B	Coef. Std. Error	Std. Coef. Beta	t	Sig.	Coll Tolerance	Stats VIF
1	(Constant)	47.590	7.783		6.114	<.001		
	Pro-Agility	-2.129	1.591	-0.401	-1.338	.223	.606	1.651
	Six Meter Hop	-6.581	4.196	-0.470	-1.568	.161	.606	1.651
2	(Constant)	43.128	7.372		5.850	<.001		
	Six Meter Hop	-10.106	3.424	-0.722	-2.952	.018	1.000	1.000

Note. Unstd: Unstandardized; Coef: Coefficient; Std: Standard; Coll: Collinearity; Stats: Statistics.

Backward stepwise regression - SJFT index

Positive correlations were observed between the SJFT index and the pro agility and 300-yard shuttle tests. The pro agility and 300-yard shuttle tests were correlated with each other ($r = .775$, $p = .008$) (Table 7). In the first model, which included both tests, the overall model was significant ($p < .042$) (Table 6). However, neither the pro agility nor the 300-yard shuttle run tests were significant individual predictors when included together in the model (Table 7). Further examination revealed that model two, which included only the 300-yard shuttle run test, was a significant individual predictor ($p < .017$) (Table 7). Therefore, the SJFT index can be predicted using the 300-yard shuttle run test ($R^2 = .53$, $p = .017$). The regression equation is: $SJFT \text{ index} = -14.034 + 0.386 \times (300\text{-yard shuttle run})$.

Table 6. The special judo fitness test index Anova.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	61.13	2	30.565	5.156	.042 ^a
	Residual	41.494	7	5.928		
	Total	102.624	9			
2	Regression	54.298	1	54.298	8.988	.017 ^b
	Residual	48.326	8	6.041		
	Total	102.624	9			

Note. a. Predictor: 300-yard shuttle run, pro-agility; b. Predictor: 300-yard shuttle run.

Table 7. The special judo fitness test index coefficients.

Model		Unstand. B	Coef Std. Error	Stand Coef Beta	t	Sig.	Coll Tolerance	Statis VIF
1	(Constant)	-12.990	9.817		-1.323	.227		
	Pro-Agility	2.088	1.945	0.408	1.074	.319	.400	2.503
	300-Yard Shuttle Run	0.218	0.202	0.411	1.081	.316	.400	2.503
2	(Constant)	-14.034	9.862		-1.423	.193		
	300-Yard Shuttle Run	0.386	0.129	-0.727	2.998	.017	1.000	1.000

Note. Unstd: Unstandardized; Coef: Coefficient; Std: Standard; Coll: Collinearity; Stats: Statistics.

DISCUSSION

Previous studies examining the relationship between the SJFT and field tests have focused solely on the SJFT index rather than the SJFT score and SJFT index. The primary objective of this study was to identify the relationships between the SJFT score and SJFT index with various field tests used to assess athletic performance in judo athletes. Our hypothesis, which anticipated correlations between the SJFT score and index with specific field tests, was partially supported by our findings. The SJFT score exhibited moderate to high correlations with the pro-agility and six-meter hop tests. At the same time, the SJFT index demonstrated moderate to high correlations with the pro-agility and 300-yard shuttle run tests.

Looking at the relation between the SJFT score and the field test we identified significant correlations with the pro-agility test ($r = -.70$) and the six-meter hop test ($r = -.73$). To our knowledge, no other study investigated the relation between the SJFT score and field tests in senior judo athletes, the majority of studies used laboratory testing (Franchini et al., 2005; Detanico et al., 2012). Franchini et al. (2005), reported that elite athletes exhibited significantly greater peak and mean power ($p < .05$) during the upper-body Wingate test on a Monark cycle ergometer, alongside superior SJFT scores, in a controlled laboratory setting. Similarly, Detanico et al. (2012) identified correlations between the SJFT score and both peak velocity and anaerobic threshold velocity during a treadmill incremental test ($r = .70$ and $r = .60$, respectively; $p < .01$).

Our study performed a linear regression analysis to predict the SJFT score using correlated field test scores. The linear regression analysis showed that the combination of the pro agility and six-meter hop tests could predict the SJFT score ($R^2 = .62$ with the overall model to be statistically significant ($p = .034$). However, individually, the pro agility test ($p = .233$) and the six-meter hop test ($p = .161$) were not significant predictors when combined in the model. Further analysis using backward stepwise regression demonstrated that the six-meter hop test alone could predict the SJFT score with 52% accuracy ($p = .018$).

These findings suggest that the six-meter hop test may be an effective tool for predicting SJFT score, given its moderate to high correlation values relative to laboratory-based assessments. Since the six-meter hop test requires maximal effort (McArdle et al., 2016) and is considered a maximum test that emphasizes

anaerobic metabolism and lower leg power (Tomchuk, 2011). Then, due to similar physiological metabolic pathways involved it is logically assumed that the six-meter hop test can be an alternative testing tool to estimate the SJFT score. Moreover, the ippon seoi-nage throw in the SJFT necessitates anaerobic power (Franchini et al., 2011), further underscoring the relevance of the six-meter hop test as a predictor of the SJFT score.

The SJFT index demonstrated significant correlations with the pro-agility test ($r = .73$) and the 300-yard shuttle run ($r = .73$). These findings align with those of Arazi et al. (2017), who reported a correlation between the SJFT index and the 4x9 m agility test ($r = .35$, $p = .01$), and (Ceylan et al., 2022) who observed a correlation between T-test agility times and the SJFT index ($r = .77$, $p < .001$). Together, these results reinforce the association between agility measures and SJFT performance.

The 300-yard shuttle run, an established test for anaerobic power and capacity (Tomchuck, 2011), which demonstrated a strong correlation with the SJFT index ($r = .73$, $p = .017$). This finding is consistent with the work of (Arazi et al., 2017) who reported a relationship between the SJFT index and the 6x35 meter sprints anaerobic power ($r = -.38$). Additionally, Ceylan et al. (2022) observed correlations between SJFT index scores and various anaerobic power measures, including mean power output on a cycle ergometer in a laboratory setting ($r = -.72$, $p < .00$). However, they did not identify similar correlations with field tests.

Following, the regression analysis results indicated that the combination of the pro-agility and 300-yard shuttle run tests demonstrated a strong predictive capability for the SJFT index ($R^2 = .60$). The overall regression model was statistically significant ($p = .042$). However, neither the pro-agility ($p = .319$) nor the 300-yard shuttle run tests ($p = .316$) were significant predictors when included together in the model. Further analysis using backward stepwise regression revealed that the 300-yard shuttle run test alone could predict the SJFT index with 53% accuracy and was individually statistically significant ($p = .017$), indicating better predictive capability.

The short-term energy system (glycolysis) is activated during maximal efforts lasting up to three minutes (McArdle et al., 2016). The Special Judo Fitness Test (SJFT) primarily targets this energy pathway, with 42% of the energy derived from anaerobic sources (Franchini et al., 2011). The SJFT, which lasts one minute and 15 seconds, requires athletes to perform at maximum effort, thereby assessing their anaerobic capacity. Comparatively, in the 300-yard shuttle run test, another anaerobic capacity assessment (Tomchuk, 2011), athletes had an average completion time of one minute and 16 seconds, respectively, in the study. This similarity in duration between the SJFT and the 300-yard shuttle run test suggests a comparable anaerobic demand. Our study found a stronger correlation than that reported by Arazi et al. (2017). This finding implies that the 300-yard shuttle run test may be a more suitable field test for evaluating anaerobic capacity in judo athletes. This alignment in test durations and energy system engagement underscores the relevance of the 300-yard shuttle run as a possible reliable measure for this specific athletic population.

Practical applications

The findings indicate that the six-meter hop test scores enable the prediction of the SJFT score with 52% accuracy. Similarly, the 300-yard shuttle run test scores allow coaches to predict approximately 53% of an athlete's SJFT index. Our study suggests that instead of frequently administering the SJFT, potentially exposing athletes to unnecessary injury risk, coaches can rely on field tests such as the six-meter hop tests to predict the SJFT score or the 300-yard shuttle run to estimate the SJFT index. These tests may provide an effective means of establishing baselines and monitoring sport-specific performance due to their simplified protocols. This practical approach helps coaches and athletes better understand and track the physical

demands and performance capabilities relevant to judo, ensuring that athletes are optimally prepared for the SJFT and competition.

Limitations

The sample size included only 10 participants, which is below the 52 recommended by the power analysis. This small sample probably reduced the generalizability of the findings and increased the risk of Type II errors. Smaller sample sizes have an increase of 14% to 21% of type II errors and partial correlations have a 43% to 85% increase in type II errors (Knudson & Lindsey, 2014). All participants were adult judo athletes, so the findings may not be applicable to younger athletes, recreational practitioners, or those from other martial arts disciplines. Potentially affecting the external validity of the study (Furler et al., 2012). Due to the lack of a control or comparison group and the fact that it is difficult to attribute observed effects solely to the interventions or conditions of the study, we may have an increase in bias (Hannan, 2008).

In addition, fatigue is a state of physical exhaustion caused by exertion, characterized by an inability to maintain the same level of exercise intensity, leading to a decline in performance (Evans & Lambert, 2007). Testing within a single session may introduce fatigue, thereby impacting performance across different tests. However, based on the NSCA guidelines, non-fatiguing tests should be done first, followed by agility, power, strength, endurance, and aerobic and anaerobic capacity tests (McGuigan, 2016). Thus, using this sequence fatigue was likely not an issue.

Furthermore, due to the pilot-in-nature correlation study design, results from this effort cannot establish causality between the examined variables (Altman & Krzywinski, 2015). But the strength of this endeavour is that we presented concepts (e.g., field tests) that can be further examined to analyse cause and effects between different field tests and their predictive power to substitute the standard SFJT test.

Future research

In general, using larger sample sizes is likely to enhance the accuracy and reliability of exercise science research results (Knudson & Lindsey, 2014). As such, future research should investigate the same research questions by utilizing larger sample sizes, probably more than 31, and adjust for attrition as well. Additionally, incorporating control groups and experimental designs can help establish causal relationships between training methods and performance outcomes. This approach distributes participants' unique characteristics across testing, preventing selective bias. In addition, random assignment provides an unbiased estimate of error effects and ensures these effects are statistically independent (Kirk, 2009).

Furthermore, since fatigue can be a factor to influence both performance and injury (Meeusen, 2013), studies also need to consider the athlete's training cycle. According to Jones et al. (2017), intensified training periods, such as the preseason, times of heightened competition, and the return of injured players to full training, have been found to both affect performance and elevate the risk of injury. Therefore, a comprehensive analysis of an athlete's performance and injury risk must account for the variations in training loads and the timing within the training cycle. This approach ensures that the findings are reflective of real-world conditions.

On top of that, expanding the demographic diversity of participants, including variations in age, gender, and skill level, can provide a more comprehensive understanding of judo performance. Along this line, Krstulović et al. (2006) and Thomas et al. (1989) have emphasized the importance of creating physiological profiles for judo athletes. Including a broader range of tests that address tactical and psychological aspects of judo could offer a more holistic view of an athlete's performance. Furthermore, integrating qualitative assessments, such as athlete interviews, might provide insights into factors influencing performance that quantitative measures

alone might miss (Poizat et al., 2013). These subjective indicators can potentially lead to a decline in performance. However, in other instances, these indicators can be relevant and sufficient for making decisions in dynamic, uncertain situations with significant time constraints (Poizat et al., 2013).

CONCLUSION

This study aimed to identify alternative judo-specific field tests that show a relationship with the SJFT. We found that the SJFT score correlated with pro agility and six-meter hop tests, whereas the SJFT index was associated with the pro-agility and the 300-yard shuttle run tests. Moreover, our study suggests that the pro-agility test can predict the SJFT score and the 300-yard shuttle run test can predict the SJFT index, reducing the need to perform the SJFT and thus minimizing the risk to athletes from being thrown and injured. These findings align with previous research, reinforcing the use of specific field tests to improve athletes' performance in the SJFT, which can be implemented in any dojo.

AUTHOR CONTRIBUTIONS

The authors stated that all contributors participated equally in the preparation of this document.

SUPPORTING AGENCIES

No funding agencies were reported by the authors.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

- Altman, N., & Krzywinski, M. (2015). Points of significance: Association, correlation and causation. *Nat. Methods*, 12(10), 885-887. <https://doi.org/10.1038/nmeth.3587>
- Arazi, H., Noori, M., & Izadi, M. (2017). Correlation of anthropometric and bio-motor attributes with Special Judo Fitness Test in senior male judokas. *Ido Mov. Cult. J. Martial Arts Anthropol.*, 17(4), 19-24.
- Castarlenas, J. L., & Planas, A. (1997). Study of the temporal structure of judo combat. *J. Phys. Educ. Sports*, 47, 32-39.
- Ceylan, B., Šimenko, J., & Balci, Ş. S. (2022). Which performance tests best define the Special Judo Fitness Test classification in elite judo athletes? *J. Funct. Morphol. Kinesiol.*, 7(4), 101. <https://doi.org/10.3390/jfmk7040101>
- Detanico, D., Dal Pupo, J., Franchini, E., & dos Santos, S. G. (2012). Relationship of aerobic and neuromuscular indexes with specific actions in judo. *Sci. Sports*, 27(1), 16-22. <https://doi.org/10.1016/j.scispo.2011.06.001>
- Drid, P., Trivić, T., & Tabakov, S. (2012). Special judo fitness test-a review. *Serb. J. Sports Sci.*, 6(4), 185-190.
- Evans, W. J., & Lambert, C. P. (2007). Physiological basis of fatigue. *Am. J. Phys. Med. Rehabil.*, 86(1), S29-S46. <https://doi.org/10.1097/PHM.0b013e31802ba53c>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2020). GPower for Windows (Version 3.1.9.7) [Computer software].

- Furler, J., Magin, P., Pirotta, M., & van Driel, M. (2012). Participant demographics reported in "Table 1" of randomised controlled trials: A case of "inverse evidence"? *Int. J. Equity Health*, 11, 1-4. <https://doi.org/10.1186/1475-9276-11-1>
- Franchini, E., de Moraes Bertuzzi, R. C., Takito, M. Y., & Kiss, M. A. P. D. (2009). Effects of recovery type after a judo match on blood lactate and performance in specific and non-specific judo tasks. *Eur. J. Appl. Physiol*, 107(4), 377-383. <https://doi.org/10.1007/s00421-009-1134-2>
- Franchini, E., Miarka, B., Matheus, L., & Del Vecchio, F. (2011). Endurance in judogi grip strength tests: Comparison between elite and non-elite judo players. *Arch. Budo*, 7(1), 1-4.
- Franchini, E., Sterkowicz, S., Szmatlan-Gabrys, U., Gabrys, T., & Garnys, M. (2011). Energy system contributions to the Special Judo Fitness Test. *Int. J. Sports Physiol. Perform.*, 6(3), 334-343. <https://doi.org/10.1123/ijsp.6.3.334>
- Franchini, E., Takito, M. Y., Kiss, M. A. P. D., & Sterkowicz, S. (2005). Physical fitness and anthropometrical differences between elite and non-elite judo players. *Biol. Sport*, 22(4), 315-328.
- Gardasevic, N., & Stankovic, N. (2019). The most frequently used judo techniques in accordance with current sport rules. *Acta Kinesiol.*, 13(1), 87-92.
- Hannan, E. L. (2008). Randomized clinical trials and observational studies: Guidelines for assessing respective strengths and limitations. *JACC Cardiovasc. Interv.*, 1(3), 211-217. <https://doi.org/10.1016/j.jcin.2008.02.002>
- Hesari, A. F., Mirzaei, B., Ortakand, S. M., Rabienejad, A., & Nikolaïdis, P. T. (2014). Relationship between aerobic and anaerobic power, and Special Judo Fitness Test (SJFT) in elite Iranian male judokas. *Apunts Med. Esport*, 49(181), 25-29. <https://doi.org/10.1016/i.apunts.2013.07.005>
- IBM Corp. (2024). IBM SPSS Statistics for Windows (Version 30.0) [Computer software].
- Jones, C. M., Griffiths, P. C., & Mellalieu, S. D. (2017). Training load and fatigue marker associations with injury and illness: A systematic review of longitudinal studies. *Sports Med.*, 47, 943-974. <https://doi.org/10.1007/s40279-016-0619-5>
- International Judo Federation. (2023). Rules. Retrieved from [Accessed 2026, 06 March]: https://78884ca60822a34fb0e6-082b8fd5551e97bc65e327988b444396.ssl.cf3.rackcdn.com/up/2024/04/IJF_Sport_and_Organisation_Rul-1712049108.pdf
- Kirk, R. E. (2009). Experimental design. In *Sage Handbook of Quantitative Methods in Psychology* (pp. 23-45). Sage. <https://doi.org/10.4135/9780857020994.n2>
- Knudson, D. V., & Lindsey, C. (2014). Type I and Type II errors in correlations of various sample sizes. *Compr. Psychol*, 3, 03-CP. <https://doi.org/10.2466/03.CP.3.1>
- Krstulović, S., Žuvela, F., & Katić, R. (2006). Biomotor systems in elite junior judoists. *Coll. Antropol.*, 30(4), 845-851.
- McArdle, W. D., Katch, F. I., & Katch, V. L. (2016). *Essentials of exercise physiology* (5th ed.). Wolters Kluwer.
- McGuigan, M. (2016). Principles of test selection and administration. In G. G. Haff & N. T. Triplett (Eds.), *Essentials of strength training and conditioning* (4th ed., pp. 249-258). Human Kinetics.
- Meeusen, R., Duclos, M., Foster, C., Fry, A., Gleeson, M., Nieman, D., Raglin, J., Rietjens, G., Steinacker, J., & Urhausen, A. (2013). Prevention, diagnosis, and treatment of the overtraining syndrome: Joint consensus statement of the European College of Sport Science (ECSS) and the American College of Sports Medicine (ACSM). *Eur. J. Sport Sci.*, 13(1), 1-24. <https://doi.org/10.1080/17461391.2012.730061>
- Poizat, G., Sève, C., & Saury, J. (2013). Qualitative aspects in performance analysis. In J. L. Seifert, J. Komar, & P. Robineau (Eds.), *Routledge handbook of sports performance analysis* (pp. 309-320). Routledge.

- Sertić, H., & Segedi, I. (2012). Structure of importance of techniques of throws in different age groups in men judo. *J. Combat Sports Martial Arts*, 3(1), 59-62. <https://doi.org/10.5604/20815735.1047649>
- Sterkowicz, S., & Franchini, E. (1995). The Special Judo Fitness Test. *Antropomotoryka*, 12(13), 29-44.
- Sterkowicz, S., & Franchini, E. (2001). Specific fitness of elite and novice judoists. *J. Hum. Kinet.*, 6(1), 81-98.
- Sterkowicz, S., Kiejda, I., & Blach, W. (1997). Characteristics of judo fighting methods in the Olympic Games 1988-1996. *Training*, 1(33), 27-43.
- Thomas, S. G., Cox, M. H., LeGal, Y. M., Verde, T. J., & Smith, H. K. (1989). Physiological profiles of the Canadian National Judo Team. *Can. J. Sport Sci.*, 14(3), 142-147.
- Tomchuk, D. (2011). *Companion guide to measurement and evaluation for kinesiology*. Jones & Bartlett Learning.
- Van Malderen, K., Jacobs, C., Ramon, K., Zinzen, E., Deriemaeker, P., & Clarys, P. (2006). Time and technique analysis of a judo fight: A comparison between males and females. In *Book of Abstracts of the 11th Annual Congress of the European College of Sport Science* (p. 101). Lausanne, Switzerland.
- Westrick, R. B., Miller, J. M., Carow, S. D., & Gerber, J. P. (2012). Exploration of the Y-balance test for assessment of upper quarter closed kinetic chain performance. *Int. J. Sports Phys. Ther.*, 7(2), 139-147.



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

Player role-specific workload profiles and the internal-external load relationship in semi-professional basketball: A two-season longitudinal study

 **Hoani K. Smith**  . Department of Tourism, Sport and Society. Lincoln University. Christchurch, New Zealand.
 **Stephen P. Bird**. School of Health, Psychological and Medical Sciences. University of Southern Queensland. Queensland, Australia.
 **Peter D. Olsen**. Department of Applied Sciences and Social Practice. Ara Institute of Canterbury. Christchurch, New Zealand.
 **Thomas Kavanagh**. Department of Tourism, Sport and Society. Lincoln University. Christchurch, New Zealand.
 **Michael J. Hamlin**. Department of Tourism, Sport and Society. Lincoln University. Christchurch, New Zealand.

ABSTRACT

Understanding how internal and external loads interact is critical for optimising basketball performance. This two-season study quantified internal (session-RPE, wellness) and external (IMU-derived accelerations, decelerations, distance, sprints, heart rate, and jumps) demands across 36 games, 292 trainings, and 74 skill sessions in 18 players (20.6 ± 1.3 years, 191.9 ± 6.5 cm, 87.3 ± 7.2 kg). Players were classified as starters (>24 min/game-1), rotation (10 – 24 min-1), or bench (< 10 min-1). Starters accumulated greater game loads ($ES = 1.65$) and high intensity sprints ($ES = 1.32$) than rotation and bench players ($p < .05$). Trainings produced higher total volume but lower intensity, while games elicited more accelerations and decelerations (Zone 2 accelerations = 142 ± 38 and 27 ± 10 ; Zone 3 = 75 ± 44 and 67 ± 20 , respectively). Game sRPE correlated with accelerations (Zone 2: $r = .70$; Zone 3: $r = .54$) and decelerations (Zone 2: $r = .68$; Zone 3: $r = .52$), indicating that sRPE reflects high-intensity demand. Integrating internal and external measures clarifies workload. Practically, bench and rotation players may benefit from high-intensity exposure, whereas moderating volume around MD + 2 may support recovery. Findings show role-specific patterns guiding evidence-based load management.

Keywords: Performance analysis, Load monitoring, Inertial measurement, Session-RPE, Training microcycle, Acceleration–deceleration, Team sport.

Cite this article as:

Smith, H. K., Bird, S. P., Olsen, P. D., Kavanagh, T., & Hamlin, M. J. (2026). Player role-specific workload profiles and the internal-external load relationship in semi-professional basketball: A two-season longitudinal study. *Scientific Journal of Sport and Performance*, 5(3), 410–423. <https://doi.org/10.55860/QDNS6807>



Corresponding author. Department of Tourism, Sport and Society, Lincoln University. PO Box 85084. Lincoln 7647. New Zealand.

E-mail: hoani.smith2@lincoln.ac.nz

Submitted for publication January 23, 2026.

Accepted for publication February 25, 2026.

Published March 19, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/QDNS6807>

INTRODUCTION

Basketball is a dynamic, high-intensity team sport characterised by intermittent bursts of physical activity, requiring athletes to perform repeated high-speed movements, accelerations, decelerations, and changes of direction (Ferreira et al., 2021). To meet these demands, players must be physically and psychologically prepared, with effective load monitoring essential for optimising performance while supporting readiness and recovery across congested competitive schedules (Edwards et al., 2018; Petway, Freitas, Calleja-González, Medina Leal, & Alcaraz, 2020). Monitoring load allows coaches to assess demands, track responses, and adjust workloads to support readiness and recovery (Gamonaes, Hernández-Beltrán, Escudero-Tena, & Ibáñez, 2023).

Player monitoring typically involves assessing two primary dimensions: 1) external load; and 2) internal load (Foster, 1998). External load represents the mechanical work measured by metrics such as total distance, sprints, accelerations, and decelerations using Global Positioning System (GPS) or inertial movement units (IMU), summarising the volume and intensity of training and competition (Halsen, 2014). These external load metrics provide insights into the volume and intensity of demands in both training and games (Stojanović et al., 2018). In basketball, external load variables such as accelerations ($>2 \text{ m.s}^{-2}$), decelerations ($>2 \text{ m.s}^{-2}$), and high-speed running ($>18 \text{ km.h}^{-1}$) are considered key indicators of workload, particularly during high-intensity match scenarios (Vázquez-Guerrero, Suarez-Arrones, Casamichana Gómez, & Rodas, 2018). Conversely, internal load reflects a player's psychobiological response to external stress. It is typically assessed through heart rate monitoring or subjective scales like the session Rating of Perceived Exertion (sRPE) (Haddad, Stylianides, Djaoui, Dellal, & Chamari, 2017). Monitoring internal load is essential for understanding how athletes cope with training demands and identified individual variability in fatigue, recovery, and readiness (Petway et al., 2020).

Understanding the relationship between external and internal load is crucial for delivering appropriate physical training to induce the desired adaptations (Fox, O'Grady & Scanlan, 2020). Integrating both measures allows practitioners to contextualise mechanical demands against individual psychophysiological responses, facilitating informed decisions regarding training prescription and recovery strategies (Fox et al., 2020). By tailoring loads to individual needs, coaches can balance stress and recovery to enhance outcomes. This is particularly relevant in basketball, where training and games differ markedly in physical and cognitive demands. Despite its practical value, limited research has examined these relationships in basketball across multiple contexts and extended timeframes (Fox, Power & Scanlan, 2020; Scanlan et al., 2014; Svilar et al., 2018).

Scanlan et al., (2014) examined relationships between accelerometer-derived workload and internal measures, such as summated-heart-rate-zones (SHRZ), training impulse (TRIMP), and session-rating of perceived exertion (sRPE), during pre-season training (Scanlan et al., 2014). They found significant correlations, with accumulated load showing a stronger correlation with SHRZ ($r = .61$), compared to moderate correlations between external workload and sRPE ($r = .49$) or TRIMPs ($r = .38$). Accumulated load, typically calculated as player load, provides a composite measure of total movement intensity and volume (Li et al., 2024). Player load is computed by summing the rate of change in acceleration across three movement planes, giving a widely used proxy for overall demand (Li et al., 2024). SHRZ appeared to better reflect physiological intensity, while sRPE captured subjective effort. Interestingly, accelerations correlated only moderately with sRPE ($r = .49$), whereas decelerations showed a stronger association ($r = .61$), suggesting movement types contribute unevenly to internal responses.

Svilar et al., (2018) extended this work by exploring accelerometer-derived external loads, IMU variables, and sRPE, in professional basketball players (Svilar & Jukic, 2018). They reported strong correlations between accumulated external workload and sRPE, ($r = .84$), especially for decelerations, ($r = .83$), and changes of direction ($r = .84$), while number of jumps showed lower correlations ($r = .49 - .55$). In contrast, IMU-derived accelerations ($r = .65$) and high-intensity accelerations ($r = .53$) were less strongly associated with sRPE (Svilar & Jukic, 2018). Together, these findings indicate that specific movement types, particularly decelerations and changes of direction, exert stronger influences on perceived exertion than accelerations or jumps.

While previous research highlights the importance of selecting appropriate monitoring variables, most studies have focused on short-term training periods or isolated session types, limiting their applicability to long-term load management. In particular, few studies have simultaneously examined internal and external load relationships across training and competitive games, over multiple seasons, and between player roles. Given the substantial differences in exposure experienced by starters, rotation, and bench players, a more comprehensive, longitudinal approach is required to inform applied load management strategies.

Therefore, the primary aim of this study was to quantify and compare internal and external load demands across two competitive seasons in semi-professional basketball players. A secondary aim was to examine differences between session types (team training, skill training, and games) and player roles (starter, rotation, bench). It was hypothesised that (1) starters would accumulate greater total and high intensity loads than rotation and bench players, and (2), games would elicit higher internal and external loads compared with training and skill sessions.

MATERIALS AND METHODS

Participants

Eighteen semi-professional, male basketball players volunteered to participate in this study (Table 1). Players belonged to the same team, a premier basketball club in the Canterbury region of New Zealand. Ten players participated in both seasons, while eight contributed data for a single season. All participants gave written informed consent (Declaration of Helsinki), with approval from an institutional human ethics committee (HEC2022-25).

Table 1. Physical characteristics of year 1 and year 2 players.

Characteristic	Year 1, 2022 (n = 14)	Year 2, 2023 (n = 10)
Age (yr)	20.3 $\pm$ 1.3	20.8 $\pm$ 1.3
Height (cm)	191.9 $\pm$ 6.5	191.8 $\pm$ 6.1
Body Mass (kg)	88.9 $\pm$ 7.5	85.6 $\pm$ 6.9
Sum of 8 Skinfolts (mm)	78.9 $\pm$ 27.6	66.6 $\pm$ 18.1
Years Playing Basketball (yrs)	7.4 $\pm$ 2.2 years	8.5 $\pm$ 2.5 years

Note. Data are mean $\pm$ SD. Sum of 8 skinfolts included triceps, subscapular, biceps, iliac crest, supraspinale, abdominal, front thigh and medial calf.

Design

A longitudinal, observational design was employed where internal and external training loads were collected during all training sessions and games over the combined 2022 and 2023 seasons. A total of 36 competitive games were played across two 18-week in-season phases. Overall, 292 IMU samples from team trainings, 74 for skill trainings, and 229 samples from games were included for analysis. Additionally, 252 days of

internal load data covering the same training period (Health and Sport Technologies, Ltd., trading as Metrifit, Millgrange, Greenore, Co., Louth, Ireland) were analysed. No experimental manipulation of training load was undertaken.

METHODS

All practices and games were included and assigned to the following activity categories: Team training, competitive games, strength training, and skill training (i.e. drills with limited contact focused on skill development e.g., spot shooting, ball handling). Team training sessions were on Monday and Thursday nights, strength training occurred on Monday, Tuesday, and Thursday mornings, and individual skill work was scheduled for Wednesday afternoons. Games were held on Saturday afternoons (Figure 1). Participants were separated by role, only in the game context, with starters playing more than 24 minutes, rotation players completing 10-24 minutes, and bench players playing fewer than 10 minutes per game (Palmer, Wundersitz, Bini, & Kingsley, 2021).

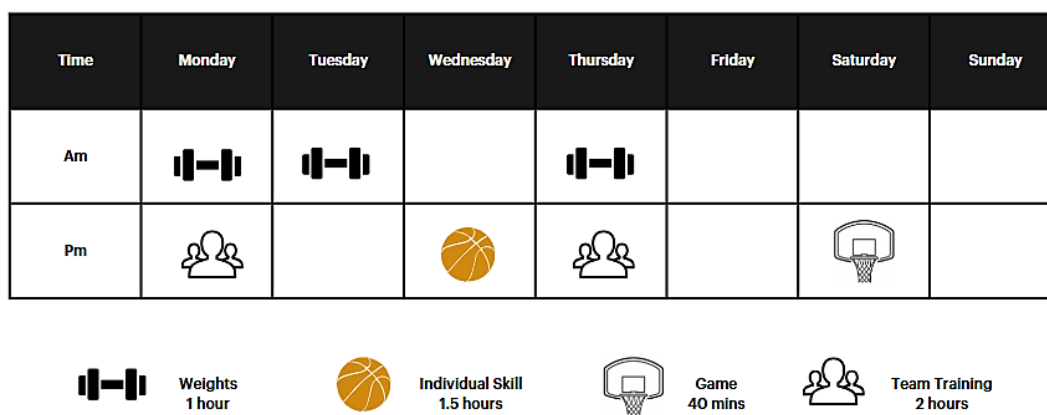


Figure 1. Weekly practice schedule for male basketball players.

Table 2. General goals and specific training contents across the structured microcycle in basketball.

Session	Goals	Contents
MD +1	Recovery	Passive recovery, or very light individual work to facilitate recovery
MD +2	Strength Training	Ideal day for heavier loads and more challenging resistance work.
	Team Training	To accumulate the highest weekly basketball-specific load and optimise the team's playing model. Simulated 5-on-5 competition.
MD -4	Strength Training	Moderate to high-intensity strength training. Tapering down from MD+2 but still ensuring maintenance of strength gains.
MD -3	Individual Training	To develop player's basic skills
		Structural Individual skills
MD -2	Strength Training	Low to moderate volume with high intensity strength / speed work. Emphasis is on maintenance, avoiding excessive fatigue while keeping muscle activation high.
	Team Training	Tactical drills, focusing on upcoming match strategies, working on team patterns, and how to adapt to the opponent.
MD	Game Day Performance	

Note. MD+1 is match day plus 1. MD+2, match day plus 2. MD-4, match day minus 4. MD-3, match day minus three. MD-2, match day minus two. MD, match day.

Training periodisation

Across both competitive seasons, training was organised into macrocycles reflecting the preparatory, competitive, and taper phases. The preparatory phase-built capacity, the competitive phase balanced training and recovery, and tapering reduced load while maintaining intensity. These phases provided the framework for the weekly microcycle. The structured microcycle formed the foundation of each training week, aimed at optimising player performance. While literature outlines general frameworks, the current study's microcycle (i.e., MD+2, MD-4, etc) was adapted to team scheduling, player availability, and coaching constraints- (Vretaros, 2024). This approach integrated physical, technical, and tactical elements across sessions, as outlined in Table 2.

External workload

External load was monitored using accelerometer, gyroscope and magnetometer sensors included in the VXSport Omni devices (VXSport, New Zealand) to capture all on-court activities. The device was worn in a fitted garment between the scapulae, sampling at 100 Hz. Each player used the same device throughout the study for consistency. After all training sessions and games, data was downloaded using the manufacturer software package and exported for integration (VXSport, New Zealand). These devices have been found to be reliable for measuring external load in basketball, particularly for tracking movement patterns such as speed and directional changes (Smith, Bird, Olsen, Kavanagh, & Hamlin, 2024).

Variables analysed included total duration (min), duration of high-intensity heart rate (min: sec, over 180 bpm), total distance covered (m), distance rate covered ($\text{m} \cdot \text{min}^{-1}$), high-intensity distance rate (m , over 15 $\text{km} \cdot \text{h}^{-1}$), maximal speed ($\text{km} \cdot \text{h}^{-1}$), sprints total (number), high-intensity sprints (number, over 18 $\text{km} \cdot \text{h}^{-1}$), HR average (bpm), HR max total (bpm), acceleration (zone 1, 0 – 2 $\text{m} \cdot \text{s}^{-2}$, zone 2, 2.1 – 4 $\text{m} \cdot \text{s}^{-2}$, zone 3, 4.1 – 10 $\text{m} \cdot \text{s}^{-2}$), deceleration (zone 1, 0 – 2 $\text{m} \cdot \text{s}^{-2}$, zone 2, 2.1 – 4 $\text{m} \cdot \text{s}^{-2}$, zone 3, 4.1 – 10 $\text{m} \cdot \text{s}^{-2}$), max jump height (cm), jump count (number), body impacts (number), distance speed very high (m , >20 $\text{km} \cdot \text{h}^{-1}$), distance speed (m) (zone 1, 0 – 6 $\text{km} \cdot \text{h}^{-1}$, zone 2, 6 – 12 $\text{km} \cdot \text{h}^{-1}$, zone 3, 12.1 – 18 $\text{km} \cdot \text{h}^{-1}$, zone 4, 18.1 – 24 $\text{km} \cdot \text{h}^{-1}$, zone 5, >24.1 $\text{km} \cdot \text{h}^{-1}$), accelerations total (number) and decelerations total (number). To ensure accuracy in measuring distance indoors, each player completed a one-time calibration session using the VXSport Omni system, as described by Smith, Bird, Olsen & Hamlin (2024).

Table 3. Daily wellness internal load monitoring likert scale.

Measure	1	2	3	4	5
Mood	Very stressed	Quite stressed	Slightly stressed	Little stress	No stress
Sleep Quality	Poor	Below average	Normal	Good	Very good
Energy Levels	Extremely low	Very low	Low	Normal	High/Excellent
Muscle Soreness	Extremely sore	Very sore	Quite sore	Mild soreness	No soreness
Academic Pressure	Academic pressure high	Academic pressure building	Heavy academic day	Normal academic pressure	No academic pressure
Diet Yesterday	Poor	Fair	Good	Very good	Excellent
Health	So sick I have to stay in bed	Feeling well below par	Just ok	Feeling healthy/normal	Never better

Internal workload

Each day, player well-being was evaluated upon waking via the administration of questionnaires on a phone app (Health and Sport Technologies Ltd., trading as Mtrifit, Milgrange, Greenore, Co. Louth, Ireland). The questions used a five-point Likert scale to record athletes' subjective ratings (Table 3). This approach has been validated in previous research, showing that athlete self-reported measures using an app, help

effectively track training loads, wellness, and injury risk (Hamlin, Wilkes, Elliot, Lizamore & Kathiravel, 2019). Participants also recorded daily session-RPE using a modified 10-point scale factor (Foster et al., 2001), a reliable indicator of exercise intensity (Eston & Williams, 1988; Gabbett & Domrow, 2007; Hamlin & Hellemans, 2007; Impellizzeri, Rampinini, Coutts, Sassi, & Marcora, 2004).

Statistical analysis

Changes in the mean of the variables and standard deviations representing the between-and within-subject variability were estimated using a mixed modelling procedure (Proc Mixed) in the Statistical Analysis System (Version 9.4, SAS Institute, Cary, NC, United States). Chances that the true effects were substantial were estimated when a value for the smallest worthwhile effect was entered into the calculation. We chose 0.20 standardised units (representing change in mean divided by the between-subject SD at baseline) as the smallest worthwhile change (Cohen, 1988). To make inferences about the true (population) uncertainties in estimate of change were presented as 95% confidence intervals and as likelihoods that the effect was increased, decreased, or was trivial. Additionally, *p*-values were reported, with statistical significance set at ≤ 0.05 . Effect size (ES) were defined as <0.20 , trivial; $0.20-0.59$, small; $0.60-1.19$, moderate; $1.20-1.99$, large; and >2.0 , very large (Hopkins, Marshall, Batterham & Hanin, 2009).

For external load variables, accelerometer-derived data (i.e. total distance, sprints, accelerations and decelerations) were recorded only on training/game days using VXSport Omni devices (VX Log VX105, Firmware V7.0.21.0, December 20, 2023, VXSport, New Zealand). These variables were summed to weekly totals, allowing the analysis of both short-term (weekly) and long-term (seasonal) workload fluctuations. Internal load data were collected daily using subjective ratings (RPE, sleep quality, energy levels, and muscle soreness). For analysis, the mean of each wellness variable was calculated per week, while RPE values were multiplied by session duration to calculate session-RPE, which was then summed weekly to quantify overall training load. This process produced weekly external and internal load values that were used to examine correlations between the two domains.

Session duration was recorded for training (start to end, including rest) and games (on-court minutes only). RPE was reported after every session and multiplied by session duration to calculate session-RPE. These values were summed daily and across the week to determine weekly training load. Weekly training loads were also categorised by player role (starter, rotation, or bench player), and separate analyses were conducted to assess differences between groups. Associations between selected external and internal variables were calculated using Pearson correlation coefficients (PROC CORR) in SAS.

RESULTS

External game load

The external training loads for the players over two seasons are shown in Table 4. Starters recorded the highest total game load (242.1 ± 40.4 , au), covering the most distance (4944.7 ± 1033.3 m), and performing the most total sprints (177.6 ± 53.7). Starters also had the most high-intensity sprints (78.2 ± 37.5), the longest time spent at a high-intensity heart rate (17.5 ± 14.4 min), and highest distance covered per minute (61.3 ± 9.3 m.min⁻¹). Additionally, starters showed the highest number of Zone 2 accelerations (142.2 ± 37.7) and Zone 3 decelerations (67.4 ± 19.5), along with the highest mean (151.8 ± 18.5 bpm) and maximum heart rates (192.9 ± 19.0 bpm).

Table 4. Average game load data per game over two seasons combined, separated into starter, rotation, and bench players.

Variable	Starter	Rotation	Bench	All
Minutes Played (min)	31.6 ± 4.4 ^{a, b}	20.5 ± 4.5 ^c	7.0 ± 4.0	21.5 ± 11.3
Game Duration (min)	81.0 ± 13.1	79.0 ± 10.9	82.3 ± 13.6	80.9 ± 12.8
Game RPE	7.6 ± 0.5 ^{a, b}	7.0 ± 1.0 ^c	5.9 ± 1.5	7.0 ± 1.3
Total Game Load (au)	242.1 ± 40.4 ^{a, b}	146.1 ± 42.5 ^c	45.7 ± 30.3	159.7 ± 91.6
Duration HR Hi-Int (min: sec)	17.5 ± 14.4 ^b	16.6 ± 17.5 ^c	8.7 ± 10.1	15.2 ± 14.8
Game Load per min (au)	7.6 ± 0.5 ^{a, b}	7.0 ± 1.0 ^c	6.0 ± 1.4	7.0 ± 1.2
Distance Total (m)	4944.7 ± 1033.3 ^{a, b}	3384.8 ± 900.5 ^c	1437.5 ± 779.2	3486.5 ± 1755.2
Distance Rate (m.min ⁻¹)	61.3 ± 9.3 ^{a, b}	43.7 ± 9.2 ^c	18.0 ± 10.8	43.6 ± 20.9
Distance Speed Hi-Int (m)	723.6 ± 343.4 ^{a, b}	488.4 ± 228.7 ^c	222.7 ± 136.1	516.2 ± 341.4
Speed Max (km.h ⁻¹)	26.2 ± 3.7 ^{a, b}	24.3 ± 4.2	23.6 ± 4.9	24.9 ± 4.4
Sprints Total (num)	177.6 ± 53.7 ^{a, b}	120.2 ± 34.8 ^c	48.8 ± 27.8	124.0 ± 69.4
Sprints Hi-Int (num)	78.2 ± 37.5 ^{a, b}	49.7 ± 25.2 ^c	21.5 ± 12.7	54.2 ± 37.7
HR Average Total (bpm)	151.8 ± 18.5 ^{a, b}	140.5 ± 19.3 ^c	125.4 ± 20.4	141.5 ± 22.1
HR Max Total (bpm)	192.9 ± 19.0 ^b	188.4 ± 24.7	181.3 ± 26.1	188.5 ± 23.1
Acceleration Zone 1 (num)	19.1 ± 11.8 ^b	16.0 ± 9.4 ^c	6.9 ± 4.9	14.7 ± 10.9
Acceleration Zone 2 (num)	142.2 ± 37.7 ^{a, b}	102.3 ± 26.2 ^c	40.5 ± 23.1	101.08 ± 53.2
Acceleration Zone 3 (num)	74.9 ± 43.7 ^{a, b}	45.8 ± 26.9 ^c	20.7 ± 11.7	52.0 ± 40.2
Deceleration Zone 1 (num)	74.9 ± 37.8 ^{a, b}	47.0 ± 22.8 ^c	20.6 ± 12.3	51.7 ± 36.7
Deceleration Zone 2 (num)	26.9 ± 9.7 ^{a, b}	20.7 ± 7.5 ^c	9.9 ± 5.2	20.3 ± 10.9
Deceleration Zone 3 (num)	67.4 ± 19.5 ^{a, b}	50.2 ± 15.8 ^c	19.8 ± 11.9	48.8 ± 26.0
Acceleration Zone 1 (min)	43.5 ± 27.2 ^{a, b}	25.8 ± 15.8 ^c	11.8 ± 7.5	29.8 ± 24.4
Acceleration Zone 2 (min)	0.6 ± 0.4 ^b	0.8 ± 0.4	1.2 ± 1.5	0.8 ± 0.9
Acceleration Zone 3 (min)	4.5 ± 1.0 ^b	5.0 ± 0.9	6.5 ± 4.0	5.2 ± 2.5
Deceleration Zone 1 (min)	2.3 ± 1.3	2.2 ± 1.2	3.1 ± 1.4	2.5 ± 1.3
Deceleration Zone 2 (min)	0.9 ± 0.3 ^b	1.0 ± 0.3	1.4 ± 1.0	1.0 ± 0.6
Deceleration Zone 3 (min)	2.1 ± 0.5 ^b	2.4 ± 1.2	3.0 ± 1.9	2.5 ± 1.2
Max Jump Height (cm)	55.3 ± 11.1 ^{a, b}	47.6 ± 10.6	44.3 ± 16.4	50.4 ± 13.4
Jump Count (num)	35.3 ± 31.0 ^{a, b}	21.0 ± 11.6 ^c	9.7 ± 6.9	24.9 ± 25.0
Body Impacts (num)	1.6 ± 1.2	1.4 ± 0.5	1.1 ± 0.3	1.5 ± 1.0
Distance Speed Very-High (m)	82.6 ± 58.9 ^{a, b}	52.0 ± 40.8 ^c	27.7 ± 23.9	58.3 ± 51.8
Distance Zone 1 0-6 km.h ⁻¹ (m)	1655.5 ± 368.5 ^{a, b}	1154.5 ± 346.5 ^c	517.4 ± 268.8	1177.8 ± 589.3
Distance Zone 2 6.1-12 km.h ⁻¹ (m)	2040.5 ± 431.5 ^{a, b}	1438.4 ± 338.3 ^c	585.7 ± 344.5	1443.3 ± 729.3
Distance Zone 3 12.1-18 km.h ⁻¹ (m)	1066.0 ± 351.7 ^{a, b}	723.1 ± 246.1 ^c	284.1 ± 183.3	744.3 ± 436.6
Distance Zone 4 18.1-24 km.h ⁻¹ (m)	163.7 ± 109.0 ^{a, b}	95.4 ± 67.4 ^c	46.1 ± 39.2	111.4 ± 94.2
Distance Zone 5 24.1+ km.h ⁻¹ (m)	15.6 ± 13.3 ^{a, b}	7.7 ± 7.0	8.3 ± 8.3	12.1 ± 11.7
Accelerations Total (num)	237.4 ± 70.0 ^{a, b}	163.3 ± 46.1 ^c	64.7 ± 37.4	165.8 ± 92.4
Decelerations Total (num)	138.2 ± 40.3 ^{a, b}	95.9 ± 28.5 ^c	38.7 ± 22.6	97.3 ± 53.6

Note. Data are mean ± SD, au, arbitrary unit; RPE, rate of perceived exertion; Min: sec, minutes: seconds; Acceleration Zone 1, 0 – 2 m.s-2; Acceleration Zone 2, 2.1 – 4 m.s-2; Acceleration Zone 3, 4.1 – 10 m.s-2; Deceleration Zone 1, 0 – 2 m.s-2; Deceleration Zone 2, 2.1 – 4 m.s-2; Deceleration Zone 3, 4.1 – 10 m.s-2; ^a significant difference between starter and rotation ($p < .05$); ^b significant difference between starter and bench ($p < .05$); ^c significant difference between rotation and bench ($p < .05$).

Team training and skill training external load

Team training sessions showed significantly higher physical demands than skill training, including greater total load (645.9 ± 152.5 au), total distance (4941.1 ± 1230.5 m, $p < .05$), high-intensity distance (544.8 ± 333.9 m, $p < .05$), sprints (157.0 ± 55.1), and high-intensity sprints (56.7 ± 32.7, $p < .05$). Heart-rate responses were also highest in team sessions (max 185.6 ± 22.5 bpm; mean 141.6 ± 18.3 bpm). Team training showed significantly more Zone 3 accelerations (67.3 ± 38.8, $p < .05$) and Zone 3 decelerations (58.1

± 22.1 , $p < .05$) compared to skill training. Skill training recorded higher jump-related outputs, including maximal jump height (59.1 ± 10.7 cm) and total jump count (103.8 ± 57.3).

External game load versus team training load

When comparing external load, team training was significantly higher for all roles (645.9 ± 152.6 au, vs. starters 242.1 ± 40.4 au; rotation 146.1 ± 42.5 au; bench 45.7 ± 30.3 au; $p < .05$) (Tables 5 and 4, respectively). Distance rate tended to be higher in team training (52.5 ± 10.2 m.min⁻¹) than games (43.6 m.min⁻¹), though it was not statistically significant (Table 5). Despite the overall greater number of acceleration and deceleration zones during team training, games showed higher per-minute averages, particularly for starters and rotation players in zones 2 and 3 (Tables 5 and 4, respectively).

Table 5. Average external training load data from two seasons of team and skill training sessions.

Variable	Team trainings	Skill trainings	All games
Training Duration (min)	93.2 $\pm$ 16.3 *	78.7 $\pm$ 10.3	21.5 $\pm$ 11.3
Training RPE	6.9 $\pm$ 0.8	6.8 $\pm$ 1.2	80.9 $\pm$ 12.8
Total Training Load (au)	645.9 $\pm$ 152.6 *	532.7 $\pm$ 123.8	7.0 $\pm$ 1.3
Duration HR Hi-Int (min: sec)	12.2 $\pm$ 13.8 *	9.6 $\pm$ 10.9	159.7 $\pm$ 91.6
Distance Total (m)	4941.1 $\pm$ 1230.5 *	3582.2 $\pm$ 1170.5	3486.5 $\pm$ 1755.2
Distance Rate (m.min ⁻¹)	52.5 $\pm$ 10.2	46.0 $\pm$ 12.4	43.6 $\pm$ 20.9
Distance Speed Hi-Int (m)	544.8 $\pm$ 333.9 *	149.5 $\pm$ 131.0	516.2 $\pm$ 341.4
Speed Max (km.h ⁻¹)	24.9 $\pm$ 3.8 *	21.6 $\pm$ 4.2	24.9 $\pm$ 4.4
Sprints Total (num)	157.0 $\pm$ 55.1 *	89.2 $\pm$ 47.5	124.0 $\pm$ 69.4
Sprints Hi-Int (num)	56.7 $\pm$ 32.7 *	18.9 $\pm$ 15.6	54.2 $\pm$ 37.7
HR Average Total (bpm)	141.6 $\pm$ 18.3	136.1 $\pm$ 18.5	141.5 $\pm$ 22.1
HR Max Total (bpm)	185.8 $\pm$ 22.5 *	174.9 $\pm$ 20.0	188.5 $\pm$ 23.1
Acceleration Zone 1 (num)	18.3 $\pm$ 15.7 *	11.0 $\pm$ 6.6	14.7 $\pm$ 10.9
Acceleration Zone 2 (num)	134.1 $\pm$ 42.9 *	83.6 $\pm$ 38.3	101.08 $\pm$ 53.2
Acceleration Zone 3 (num)	67.3 $\pm$ 38.8 *	40.9 $\pm$ 29.8	52.0 $\pm$ 40.2
Deceleration Zone 1 (num)	59.4 $\pm$ 32.2 *	23.6 $\pm$ 17.1	51.7 $\pm$ 36.7
Deceleration Zone 2 (num)	27.1 $\pm$ 11.3 *	19.3 $\pm$ 8.4	20.3 $\pm$ 10.9
Deceleration Zone 3 (num)	58.1 $\pm$ 22.1 *	27.8 $\pm$ 13.4	48.8 $\pm$ 26.0
Acceleration Zone 1 (min)	34.5 $\pm$ 21.8 *	14.0 $\pm$ 11.1	29.8 $\pm$ 24.4
Acceleration Zone 2 (min)	0.2 $\pm$ 0.2	-	0.8 $\pm$ 0.9
Acceleration Zone 3 (min)	0.2 $\pm$ 0.2	-	5.2 $\pm$ 2.5
Deceleration Zone 1 (min)	0.7 $\pm$ 0.4	-	2.5 $\pm$ 1.3
Deceleration Zone 2 (min)	0.3 $\pm$ 0.1	-	1.0 $\pm$ 0.6
Deceleration Zone 3 (min)	0.6 $\pm$ 0.2	-	2.5 $\pm$ 1.2
Max Jump Height (cm)	54.0 $\pm$ 11.3	59.1 $\pm$ 10.7	50.4 $\pm$ 13.4
Jump Count (num)	60.3 $\pm$ 39.0 *	103.8 $\pm$ 57.3	24.9 $\pm$ 25.0
Body Impacts (num)	1.2 $\pm$ 1.1 *	2.3 $\pm$ 2.7	1.5 $\pm$ 1.0
Distance Speed Very-High (m)	68.6 $\pm$ 61.6	20.4 $\pm$ 22.9	58.3 $\pm$ 51.8
Distance Zone 1 0-6 km.h ⁻¹ (m)	2057.3 $\pm$ 545.1	2049.5 $\pm$ 722.2	1177.8 $\pm$ 589.3
Distance Zone 2 6.1-12 km.h ⁻¹ (m)	1909.7 $\pm$ 538.5 *	1213.7 $\pm$ 481.3	1443.3 $\pm$ 729.3
Distance Zone 3 12.1-18 km.h ⁻¹ (m)	805.0 $\pm$ 372.3 *	267.9 $\pm$ 209.1	744.3 $\pm$ 436.6
Distance Zone 4 18.1-24 km.h ⁻¹ (m)	118.6 $\pm$ 98.6 *	26.4 $\pm$ 22.7	111.4 $\pm$ 94.2
Distance Zone 5 24.1+ km.h ⁻¹ (m)	12.5 $\pm$ 12.6	11.9 $\pm$ 12.1	12.1 $\pm$ 11.7
Accelerations Total (num)	218.8 $\pm$ 73.5 *	135.2 $\pm$ 64.5	165.8 $\pm$ 92.4
Decelerations Total (num)	119.1 $\pm$ 41.6 *	59.4 $\pm$ 27.2	97.3 $\pm$ 53.6

Note. Data are mean $\pm$ SD, au, arbitrary unit; RPE, rate of perceived exertion; Min: sec, minutes: seconds; Acceleration Zone 1, 0 – 2 m.s⁻²; Acceleration Zone 2, 2.1 – 4 m.s⁻²; Acceleration Zone 3, 4.1 – 10 m.s⁻²; Deceleration Zone 1, 0 – 2 m.s⁻²; Deceleration Zone 2, 2.1 – 4 m.s⁻²; Deceleration Zone 3, 4.1 – 10 m.s⁻²; *significant difference between team trainings and skill trainings ($p < .05$).

Table 6. Selected demanding external variables for starters, rotation and bench players.

External variable	Role	MD+2		MD-3		MD-2		MD
		Mean $\pm$ SD	% Diff	Means $\pm$ SD	% Diff	Means $\pm$ SD	% Diff	Means $\pm$ SD
Duration HR Hi-Int (min: sec) > 180 bpm	Starter	3.3 $\pm$ 1.82	- 82%	18.2 $\pm$ 15.9	+ 1%	18.1 $\pm$ 16.1	+ 1%	18.0 $\pm$ 14.4
	Rotation	22.7 $\pm$ 23.8	+ 59%	-	-	3.2 $\pm$ 2.5	- 78%	14.3 $\pm$ 14.5
	Bench	-	-	4.8 $\pm$ 5.9	- 47%	4.7 $\pm$ 4.3	- 48%	9.1 $\pm$ 10.4
	All	11.7 $\pm$ 15.5	- 22%	12.1 $\pm$ 15.3	- 19%	12.7 $\pm$ 12.7	- 15%	15.0 $\pm$ 14.0
Distance Total (m)	Starter	4622.7 $\pm$ 1125.6	- 8%	4510.0 $\pm$ 1197.8	- 11%	4737.5 $\pm$ 809.5	- 6%	5042.9 $\pm$ 1003.8 ^{1,2}
	Rotation	3571.2 $\pm$ 1301.8	+ 4%	3382.5 $\pm$ 958.9	- 1%	3508.2 $\pm$ 833.5	+ 2%	3432.9 $\pm$ 919.3 ³
	Bench	-	-	3816.0 $\pm$ 3160.8	+ 164%	2825.7 $\pm$ 3080.0	+ 96%	1444.4 $\pm$ 794.4
	All	4822.4 $\pm$ 1388.9 ^{a, b}	+ 39%	3666.7 $\pm$ 1212.7 ^c	+ 5%	4871.7 $\pm$ 1216.1 ^e	+ 40%	3479.2 $\pm$ 1805
Distance Speed Hi-Int (m) > 15km.h ⁻¹	Starter	732.2 $\pm$ 394 $\pm$ 7	- 1%	639.8 $\pm$ 471 $\pm$ 6	- 14%	603.7 $\pm$ 154.4	- 19%	742.1 $\pm$ 355.5 ^{1,2}
	Rotation	334.5 $\pm$ 207.8	- 34%	573.0 $\pm$ 398 $\pm$ 4	+ 14%	435.6 $\pm$ 107.0	- 14%	503.7 $\pm$ 211.2 ³
	Bench	-	-	460.0 $\pm$ 203.6	+ 107%	324.3 $\pm$ 274.5	+ 46%	222.1 $\pm$ 137.4
	All	528.3 $\pm$ 356.0 ^a	+ 2%	222.5 $\pm$ 263.6 ^{c, d}	- 57%	548.0 $\pm$ 309.1	+ 6%	519.2 $\pm$ 340.8
Sprints Hi-Int (num) > 18 km.h ⁻¹	Starter	71.0 $\pm$ 26.7	- 11%	68.8 $\pm$ 47.1	- 14%	71.5 $\pm$ 16.3	- 11%	79.9 $\pm$ 37.3 ^{1,2}
	Rotation	36.8 $\pm$ 21.0	- 28%	61.2 $\pm$ 44.9	+ 20%	45.2 $\pm$ 13.9	- 12%	51.1 $\pm$ 23.3 ³
	Bench	-	-	47.0 $\pm$ 26.9	+ 118%	30.7 $\pm$ 30.2	+ 42%	21.6 $\pm$ 12.8
	All	55.2 $\pm$ 34.1 ^a	+ 1%	26.4 $\pm$ 28.3 ^{c, d}	- 51%	57.1 $\pm$ 31.3	+ 5%	54.4 $\pm$ 37.8
Acceleration Zone 1 (num) 0 – 2 m.s ⁻²	Starter	16.2 $\pm$ 5.9	- 17%	16.0 $\pm$ 9.1	- 18%	14.2 $\pm$ 2.1	- 28%	19.6 $\pm$ 12.3 ²
	Rotation	21.2 $\pm$ 7.6	+ 33%	6.7 $\pm$ 1.5	- 58%	16.0 $\pm$ 14.7	0%	16.0 $\pm$ 9.0 ³
	Bench	-	-	15.0 $\pm$ 14.1	+ 114%	8.0 $\pm$ 7.0	+ 14%	7.0 $\pm$ 5.0
	All	17.9 $\pm$ 19.9	+ 22%	11.5 $\pm$ 7.1 ^c	- 22%	18.3 $\pm$ 10.5	+ 24%	14.7 $\pm$ 11.1
Acceleration Zone 2 (num) 2.1 – 4 m.s ⁻²	Starter	136.2 $\pm$ 37.7	- 7%	109.0 $\pm$ 59.0	- 25%	139.0 $\pm$ 27.1	- 5%	145.8 $\pm$ 33.7 ^{1,2}
	Rotation	105.8 $\pm$ 45.9	+ 2%	113.0 $\pm$ 34.6	+ 8%	105.4 $\pm$ 37.3	+ 1%	104.2 $\pm$ 25.0 ³
	Bench	-	-	132.5 $\pm$ 108.2	+ 229%	76.3 $\pm$ 83.4	+ 89%	40.3 $\pm$ 23.3
	All	132.3 $\pm$ 47.1 ^{a, b}	+ 31%	87.3 $\pm$ 40.9 ^c	- 14%	132.1 $\pm$ 41.2 ^e	+ 30%	101.3 $\pm$ 53.9
Acceleration Zone 3 (num) 4.1 – 10 m.s ⁻²	Starter	79.5 $\pm$ 25.1	+ 3%	53.0 $\pm$ 39.8	- 31%	68.7 $\pm$ 26.9	- 11%	77.3 $\pm$ 44.6 ^{1,2}
	Rotation	37.0 $\pm$ 22.1	- 21%	62.2 $\pm$ 27.9	+ 32%	60.2 $\pm$ 48.3	+ 28%	47.0 $\pm$ 27.2 ³
	Bench	-	-	62.5 $\pm$ 50.2	+ 202%	37.0 $\pm$ 43.3	+ 79%	20.7 $\pm$ 11.9
	All	67.3 $\pm$ 40.6 ^a	+ 28%	42.8 $\pm$ 30.7 ^c	- 19%	65.2 $\pm$ 37.4	+ 24%	52.7 $\pm$ 41.3
Deceleration Zone 1 (num) 1 – 0 – 2 m.s ⁻²	Starter	23.7 $\pm$ 7.4	- 15%	20.2 $\pm$ 8.1	- 27%	28.2 $\pm$ 3.0	+ 1%	27.8 $\pm$ 9.7 ^{1,2}
	Rotation	24.3 $\pm$ 5.4	+ 15%	13.0 $\pm$ 4.9	- 38%	19.4 $\pm$ 12.6	- 8%	21.1 $\pm$ 7.4 ³
	Bench	-	-	22.5 $\pm$ 20.5	+ 137%	13.3 $\pm$ 15.5	+ 40%	9.5 $\pm$ 5.2
	All	25.9 $\pm$ 11.2 ^{a, b}	+ 26%	19.2 $\pm$ 8.6 ^c	- 7%	27.3 $\pm$ 11.3 ^e	+ 33%	20.6 $\pm$ 11.2
Deceleration Zone 2 (num) 2.1 – 4 m.s ⁻²	Starter	56.5 $\pm$ 11.5	- 19%	54.3 $\pm$ 34.8	- 22%	53.5 $\pm$ 10.3	- 23%	69.4 $\pm$ 17.1 ^{1,9}
	Rotation	46.0 $\pm$ 15.0	- 9%	58.7 $\pm$ 14.0	+ 16%	44.4 $\pm$ 15.0	- 13%	50.6 $\pm$ 16.1 ³
	Bench	-	-	46.5 $\pm$ 33.2	+ 135%	36.3 $\pm$ 39.7	+ 83%	19.8 $\pm$ 12.0
	All	56.3 $\pm$ 24.1 ^a	+ 15%	31.9 $\pm$ 19.2 ^{c, d}	- 35%	58.3 $\pm$ 20.3	+ 19%	48.9 $\pm$ 26.4
Deceleration Zone 3 (num) 4.1 – 10 m.s ⁻²	Starter	44.2 $\pm$ 21.4	- 1%	31.1 $\pm$ 27.1	- 30%	46.0 $\pm$ 17.5	+ 3%	44.7 $\pm$ 27.4 ^{1,9}
	Rotation	19.7 $\pm$ 10.8	- 26%	36.0 $\pm$ 20.0	+ 35%	21.8 $\pm$ 13.8	- 18%	26.6 $\pm$ 15.5 ³
	Bench	-	-	35.0 $\pm$ 25.5	+ 197%	17.3 $\pm$ 16.6	+ 47%	11.8 $\pm$ 7.6
	All	34.4 $\pm$ 22.8 ^a	+ 14%	16.9 $\pm$ 15.3 ^{c, d}	- 44%	33.8 $\pm$ 21.1	+ 12%	30.2 $\pm$ 24.7

Note. Data are mean $\pm$ SD; % Diff shows the percentage difference between training and game values; MD+2, match day plus two; MD-3, match day minus three; MD-2, match day minus two; MD, match day; ^a significant difference between all MD+2 and MD-3; ^b significant difference between all MD+2 and MD; ^c significant difference between all MD-3 and MD-2; ^d significant difference between all MD-3 and MD; ^e significant difference between all MD-2 and MD; ¹ significant difference between starter and rotation on MD; ² significant difference between starter and bench on MD; ³ significant difference between rotation and bench on MD; Where n $\leq$ 3 data was not presented.

External game load versus team trainings most demanding variables

The total distance was significantly lower on match days (MD) compared to other days (4822.4 ± 1388.9 m on MD+2 and 3479 ± 1805 m on MD; $p < .05$, ES = 0.32, Table 6). In contrast, high-intensity distance covered was greater on MD, with significant differences between MD+2 and MD-3 (528.3 ± 356.0 vs 222.5 ± 263.6 , $p < .05$, ES = 0.81). Sprint efforts were also higher on match days, with starters completing 79.9 ± 37.3 sprints compared to 26.4 ± 28.3 on MD-3 ($p < .05$, ES = 0.70). Acceleration demands in zone 2 showed small to moderate differences, between MD+2 and MD (132.3 ± 47.1 vs. 101.3 ± 53.9 , $p < .05$, ES = 0.26). Deceleration demands were similarly varied, with the highest values seen on MD+2 (56.3 ± 24.1) and lower values on MD (31.9 ± 19.2 , $p < .05$, ES = 0.55). Significant differences were also found between starters and bench players on match day, with starters covering more distance (5042.9 ± 1003 vs. 1444.4 ± 794.4 , $p < .05$, ES = 1.1).

Internal workload

The average daily internal load was similar across player roles, with no significant differences observed between Starters, Rotation, and Bench Players (Table 7). Small variations were observed in health status, diet, and mood, however, overall patterns remained consistent across groups.

Table 7. Internal daily load averages measured over two seasons.

Variable	Starter	Rotation	Bench	All
Mood state	3.5 ± 0.7	3.7 ± 0.6	3.6 ± 0.5	3.6 ± 0.6
Sleep quality	3.4 ± 0.9	3.4 ± 0.9	3.4 ± 0.8	3.5 ± 0.8
Energy levels	3.5 ± 0.7	3.6 ± 0.5	3.6 ± 0.6	3.6 ± 0.6
Muscle readiness	3.5 ± 0.6	3.7 ± 0.7	3.8 ± 0.6	3.6 ± 0.6
Academic pressure	3.8 ± 0.8	3.8 ± 0.8	3.8 ± 1.0	3.8 ± 0.8
Diet yesterday	3.4 ± 0.7	3.3 ± 0.7	3.4 ± 0.9	3.4 ± 0.8
Sleep duration (hours)	7.7 ± 1.2	7.7 ± 1.1	8.2 ± 1.3	8.0 ± 1.1
Health	3.5 ± 0.7	3.6 ± 0.8	3.6 ± 0.7	3.6 ± 0.7

Note. Data are mean $\pm$ SD.

Internal versus external associations

Large correlations were observed between game sRPE and accelerations in ($r = .70$, $p < .001$), and zone 3 ($r = .54$, $p < .001$), with smaller associations observed for Zone 1 accelerations ($r = .31$, $p < .001$). Moderate to large correlations were also identified between game sRPE and decelerations (Zone 1: $r = .49$; Zone 2: $r = .68$; Zone 3: $r = .52$; all $p < .001$). Game sRPE demonstrated a large correlation with total game load ($r = .70$, $p < .001$). No meaningful associations were observed between wellness variables and acceleration or deceleration metrics.

DISCUSSION

This study investigated internal and external load demands in semi-professional male basketball players across two competitive seasons, including games, team training, and skill sessions. The primary finding was that RPE strongly correlated with high-intensity external metrics (accelerations, decelerations, and overall game load) confirming its value as an indicator of peak physical demands. Role related differences were also evident, with starters consistently exposed to higher total and high-intensity efforts than rotation and bench players. Analysis of session type revealed clear contrasts: team trainings accumulated the greatest total volume, games imposed the highest per-minute intensities, and skill sessions emphasised jumping outputs. Collectively, these results highlight the importance of monitoring both internal and external variables to guide role-specific and context specific load management strategies.

A key finding was the strong relationship between RPE and high-intensity load metrics. Zone 3 decelerations ($r = .52$) and Zone 3 accelerations ($r = .5$) demonstrated stronger correlations with RPE than lower-intensity zones, suggesting that perceived exertion is closely linked to the most physically demanding actions of the game. These findings align with Scanlan et al. (2014), who reported moderate correlations between sRPE and accelerometer derived loads in basketball players ($r = .49$), and Svilar & Jukic (2018) who found significant correlations between player load and sRPE ($r = .84$) (Scanlan et al., 2014; Svilar & Jukic, 2018). Collectively, this evidence supports the use of sRPE as a practical indicator of high-intensity workload in applied basketball settings.

Although training sessions produced greater overall volume at lower intensities, games consistently imposed greater demands at higher intensities across measures such as RPE, time spent in high-intensity heart rate zones, accelerations and decelerations, and high intensity sprints. This pattern was mirrored in decelerations, particularly in zone 3 ($4.1 - 10 \text{ m.s}^{-2}$), where games placed the highest demands on players. This was also observed in a systematic review by Stojanovic, (2018) where players performed more frequent accelerations and decelerations during games, with 2.2 to 4.3 accelerations per min in games versus 1.2 to 2.5 per min in trainings (Stojanović et al., 2018). These findings emphasise the importance of appropriately exposing players to high-intensity demands during training. However, skill training, while lower in overall intensity, had the most jumps relative to both games and team training sessions, likely due to the specific nature of skill work, where repetitive actions such as shooting, lay-ups, and rebounding drills, are emphasised more frequently. This suggests that skill sessions may contribute uniquely to jump exposure within the weekly training structure.

While per-minute ratios provide useful context, examining peak-demand external variables is also essential, as these highlight the accelerations, decelerations, and high-speed efforts that games demand most. Incorporating these findings into training design helps ensure sessions replicate the highest match-play intensities across starter, rotation, and bench roles. As shown in Table 6, starters covered significantly more distance at high intensities compared to rotational and bench players. The greatest difference was observed in zone 2 decelerations, where starters exhibited a considerably higher number of decelerations than both rotational and bench players. These findings echo García et al. (2022), who reported that while training occasionally surpassed match play in high-speed running ($>18 \text{ km.h}^{-1}$) it did not replicate the broader intensity demands of competition (García et al., 2022). Similarly, García et al. (2022) reported that starters experienced higher values for both external and internal loads than rotation players, reinforcing that role-based differences extend beyond external measures. The present findings extend this work by demonstrating that these role-based differences persist across multiple seasons and session contexts.

This study has several limitations. The sample size prevented positional analysis (e.g., guards vs. forwards, front court vs. backcourt), which may have revealed role-specific differences. Strength training data were limited to a general overview, without details on repetitions, sets, and resistance loads, restricting insight into concurrent strength adaptations. Additionally, IMUs could not filter out low-intensity movements during stoppages (e.g., fouls), likely inflating low-intensity workload values and underrepresenting actual on-court demands. Future studies should incorporate positional analyses, detailed strength-training data, and advanced IMU filtering to refine accuracy.

CONCLUSION

The findings from this study provide valuable insights into the internal and external workload demands of semi-professional basketball players across two competitive seasons. Clear role-specific differences were

observed, with starters consistently exposed to higher total and high-intensity loads than rotation and bench players. Games imposed the greatest high intensity demands, particularly accelerations, decelerations, and time spent in elevated heart rate zones, while team trainings accumulated larger overall volumes and skill sessions placed unique emphasis on jumps. Taken together, these results show that internal and external load measures do not always align, with external metrics highlighting role and session-based differences, while internal measures such as RPE confirmed that starters also perceived greater exertion than the other roles. This demonstrates that combining internal and external metrics provides the clearest picture of how demands differ across games, team trainings, and skill sessions. These findings have practical implications for how training exposure is distributed across player roles. Differences observed across the weekly microcycle further highlight the importance of considering both volume and intensity when planning training around competition. By incorporating both IMU-derived metrics and subjective measures such as RPE, this study expands on previous research, offering a comprehensive approach to monitoring basketball workloads. Future research should build on these findings by including positional analysis and refining load management strategies tailored to individual player roles. Such work will support more informed and individualised approaches to performance monitoring in basketball.

AUTHOR CONTRIBUTIONS

The study was designed by HS and MH. HS, MS and PO were responsible for data collection. HS conducted the data analysis, with MH contributing to data interpretation. HS drafted the manuscript. MH, SB, PO, and TK critically reviewed the manuscript and contributed important intellectual content. All authors have read and approved the final version of the manuscript.

SUPPORTING AGENCIES

No funding agencies were reported by the authors.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

- Cohen, J. (1988). *Statistical Power Analysis for Behavioural Sciences*. New York: Routledge.
- Edwards, T., Spiteri, T., Piggott, B., Bonhotal, J., Haff, G. G., & Joyce, C. (2018). Monitoring and Managing Fatigue in Basketball. *Sports*, 6(1), 19. <https://doi.org/10.3390/sports6010019>
- Eston, R. G., & Williams, J. G. (1988). Reliability of ratings of perceived effort regulation of exercise intensity. *British Journal of Sports Medicine*, 22(4), 153-155. <https://doi.org/10.1136/bjbm.22.4.153>
- Ferreira, M., Camões, M., Lima, R. F., Silva, R., Castro, H. d. O., Mendes, B., . . . Clemente, F. M. (2021). Variations of workload and well-being measures across a professional basketball season. *Revista Brasileira de Cineantropometria & Desempenho Humano*, 23. <https://doi.org/10.1590/1980-0037.2021v23e75863>
- Foster, C. (1998). Monitoring training in athletes with reference to overtraining syndrome. *Medicine and science in sports and exercise*, 30(7), 1164-1168. <https://doi.org/10.1097/00005768-199807000-00023>

- Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., . . . Dodge, C. (2001). A new approach to monitoring exercise training. *Journal of strength and conditioning research*, 15(1), 109-115. <https://doi.org/10.1519/00124278-200102000-00019>
- Fox, J. L., O'Grady, C. J., & Scanlan, A. T. (2020). The Relationships Between External and Internal Workloads During Basketball Training and Games. *International Journal of Sports Physiology and Performance*, 15(8), 1081-1086. <https://doi.org/10.1123/ijspp.2019-0722>
- Gabbett, T. J., & Domrow, N. (2007). Relationships between training load, injury, and fitness in sub-elite collision sport athletes. *Journal of Sports Sciences*, 25(13), 1507-1519. <https://doi.org/10.1080/02640410701215066>
- Gamonales, J. M., Hernández-Beltrán, V., Escudero-Tena, A., & Ibáñez, S. J. (2023). Analysis of the External and Internal Load in Professional Basketball Players. *Sports*, 11(10), 195. <https://doi.org/10.3390/sports11100195>
- García, F., Schelling, X., Castellano, J., Martín-García, A., Pla, F., & Vázquez-Guerrero, J. (2022). Comparison of the most demanding scenarios during different in-season training sessions and official matches in professional basketball players. *Biology of Sport*, 39(2), 237-244. <https://doi.org/10.5114/biolsport.2022.104064>
- Haddad, M., Stylianides, G., Djaoui, L., Dellal, A., & Chamari, K. (2017). Session-RPE Method for Training Load Monitoring: Validity, Ecological Usefulness, and Influencing Factors. *Frontiers in Neuroscience*, 11. <https://doi.org/10.3389/fnins.2017.00612>
- Halsen, S. L. (2014). Monitoring Training Load to Understand Fatigue in Athletes. *Sports Medicine*, 44(S2), 139-147. <https://doi.org/10.1007/s40279-014-0253-z>
- Hamlin, M., & Hellemans, J. (2007). Effect of intermittent normobaric hypoxic exposure at rest on haematological, physiological, and performance parameters in multi-sport athletes. *Journal of Sports Sciences*, 25(4), 431-441. <https://doi.org/10.1080/02640410600718129>
- Hamlin, M., Wilkes, D., Elliot, C., Lizamore, C., & Kathiravel, Y. (2019). Monitoring Training Loads and Perceived Stress in Young Elite University Athletes. *Frontiers in Physiology*, 10, 34. <https://doi.org/10.3389/fphys.2019.00034>
- Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and science in sports and exercise*, 41(1), 3-13. <https://doi.org/10.1249/MSS.0b013e31818cb278>
- Impellizzeri, F. M., Rampinini, E., Coutts, A. J., Sassi, A., & Marcora, S. M. (2004). Use of RPE-based training load in soccer. *Medicine & Science in Sports & Exercise*, 36(6), 1042-1047. <https://doi.org/10.1249/01.MSS.0000128199.23901.2F>
- Li, X., Shang, L., Qin, S., & Yu, H. (2024). The impact of internal and external loads on player performance in Chinese basketball association. *BMC Sports Science, Medicine and Rehabilitation*, 16(1). <https://doi.org/10.1186/s13102-024-00983-6>
- Palmer, J., Wundersitz, D., Bini, R., & Kingsley, M. (2021). Effect of Player Role and Competition Level on Player Demands in Basketball. *Sports*, 9(3), 38. <https://doi.org/10.3390/sports9030038>
- Petway, A. J., Freitas, T. T., Calleja-González, J., Medina Leal, D., & Alcaraz, P. E. (2020). Training load and match-play demands in basketball based on competition level: A systematic review. *PLOS ONE*, 15(3), e0229212. <https://doi.org/10.1371/journal.pone.0229212>
- Scanlan, A., Wen, P., Tucker, P., & Dalbo, V. (2014). The relationships between internal and external training load models during basketball training. *Journal of strength and conditioning research*, 28(9), 2397-2405. <https://doi.org/10.1519/JSC.0000000000000458>
- Smith, H. K., Bird, S. P., Olsen, P., Kavanagh, T., & Hamlin, M. J. (2024). Validity and reliability of the VXSport (Omni) device on basketball movement parameters. *Scientific Journal of Sport and Performance*, 3(3), 383-394. <https://doi.org/10.55860/QIDN6711>

- Stojanović, E., Stojiljković, N., Scanlan, A., Dalbo, V., Berkelmans, D., & Milanović, Z. (2018). The Activity Demands and Physiological Responses Encountered During Basketball Match-Play: A Systematic Review. *Sports Medicine*, 48(1), 111-135. <https://doi.org/10.1007/s40279-017-0794-z>
- Svilar, L., Castellano, J., Jukic, I., & Casamichana, D. (2018). Positional Differences in Elite Basketball: Selecting Appropriate Training-Load Measures. *International Journal of Sports Physiology and Performance*, 13(7), 947-952. <https://doi.org/10.1123/ijspp.2017-0534>
- Svilar, L., & Jukic, I. (2018). Load monitoring system in top-level basketball team: Relationship between external and internal training load. *Kinesiology*, 50, 25-33. <https://doi.org/10.26582/k.50.1.4>
- Vázquez-Guerrero, J., Suarez-Arrones, L., Casamichana Gómez, D., & Rodas, G. (2018). Comparing external total load, acceleration and deceleration outputs in elite basketball players across positions during match play. *Kinesiology*, 50(2), 228-234. <https://doi.org/10.26582/k.50.2.11>
- Vretaros, A. (2024). Describing the Distribution of Training Loads in the Microcycles and Mesocycles of Competitive Basketball. *Italian Journal of Sports Rehabilitation and Posturology*, 11, 2872.



Performance analysis of beach volleyball match outcomes: A validated efficiency-based prediction model from 10,095 AVP/FIVB matches

-  **Kofi Nyantakyi Appiah**  . School of Physical Education. Lovely Professional University. Punjab, India.
Physical Education Department. Wesley College of Education. Kumasi, Ghana.
-  **Divyanshu Kumar Singh**. School of Physical Education. Lovely Professional University. Punjab, India.
-  **Nathanael Adu**. School of Physical Education. Lovely Professional University. Punjab, India.
Mampong Technical College of Education. Ashanti, Ghana.
-  **Edward Edem Nartey**. University of Cape Coast. Cape Coast, Ghana.

ABSTRACT

Predicting the outcomes of elite beach volleyball matches involves the integration of technical and demographic performance measures. Using data from 10,095 elite professional beach volleyball matches (2002-2019), a logistic regression model was developed and validated to predict match outcome with six different predictor variables: kill efficiency differential, aces differential, error rate differential, dig efficiency differential, age difference, and height difference. After assessing for multicollinearity, these variables were used in the model. Results show that the model is well-calibrated (Brier Score = 0.208; Hosmer-Lemeshow test = 0.388), and it discriminates well between the two possible outcomes (area under curve = 0.658; 95% confidence interval: 0.638 – 0.678). In addition to being well calibrated, results also indicate that the model is internally valid, with little evidence of over fitting (shrinkage = 0.995), and temporally valid, as its ability to predict match outcomes has remained relatively consistent across the 18-year period studied. Additionally, results from gender stratified analysis indicated no differences in predictive accuracy for males and females. Overall, the validated model can be viewed as a robust support tool to assist coaches and teams with performance evaluations, strategic planning and competitive placements.

Keywords: Performance analysis, Beach volleyball, Match prediction, Logistic regression, Model validation, Performance prediction.

Cite this article as:

Appiah, K. N., Singh, D. K., Adu, N., & Nartey, E. E. (2026). Performance analysis of beach volleyball match outcomes: A validated efficiency-based prediction model from 10,095 AVP/FIVB matches. *Scientific Journal of Sport and Performance*, 5(3), 424-446. <https://doi.org/10.55860/NYNN9384>



Corresponding author. Physical Education Department. Wesley College of Education. Kumasi, Ghana.

E-mail: kofi.nyantakyi2025@lpu.in

Submitted for publication January 31, 2026.

Accepted for publication March 04, 2026.

Published March 19, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/NYNN9384>

INTRODUCTION

Beach volleyball is a competitive sport that differs from indoor volleyball in many ways, including team size (two players per side), outdoor environment and unique environmental challenges (Knoblochova et al., 2021; Kostyukov & Dashaev, 2022). Professionalization of beach volleyball has experienced rapid growth over the past two decades as well as expansion of coverage by mass media and standardization of competition format within the international volleyball federation (FIVB) tour structures (Choi & Byun, 2024). Awareness of demographic and technical factors that relate to match success is important for coaching staff who seek evidence-based guidance on training prioritization; talent scouts who need quantitative frameworks for player evaluations; and organizational decision makers who need strategic decision support.

Previous volleyball research has identified attacking efficiency, service effectiveness, and defensive coverage as determinants of elite versus sub-elite performance (Künzel et al., 2014; Umarov, 2024). Kill efficiency (the ratio of successful attack attempts to total attacks) has been found biomechanically to be the most important predictor of match success (Chun & Shin, 2020). Despite this previous research identifying key factors of elite performance through the use of descriptive methods, using smaller samples, and having less rigorous validation than possible, none of the above studies used advanced statistical predictive models with both internal and external validation to assess the clinical relevance of the results obtained.

It is plausible that demographic factors (such as age) and anthropometric measurements of an athlete can impact their ability to win matches through mechanisms such as experience based on tactical sophistication, psychological resiliency and biomechanical advantages (Bukova et al., 2017; Yusni et al., 2025). However, it is also unclear what proportion of the variance in beach volleyball match outcomes can be attributed to technical aspects of performance versus demographic/anthropometric factors, and no systematic attempts have been made to quantify these interactions among young athletes playing in beach volleyball. The lack of established prediction models for beach volleyball match outcomes creates a gap in evidence based practices.

The purpose of this study was to: (1) find out which technical and demographic factors were the strongest predictors of beach volleyball game results on an extensive professional data set spanning 18 years; (2) establish a simple logistic regression model that can predict game results; (3) assess model performance by measuring discrimination, calibration, and predictive accuracy; (4) test whether the model can be applied in different populations with internal validation (bootstrap resampling, k-fold cross-validation), and external temporal validation; (5) determine if the models performed equally well across gender categories; and (6) provide information on how these models could help coaches, scouts, and performance analysts.

Important distinction: association vs. causation

It is extremely important to note that logistic regression models do not determine causal relationships and predictive variables, they instead find statistical correlations and predictive variables. Although the identified predictive variables (age, error rate, dig efficiency, kill efficiency) have a significant statistical relationship to game results, the fact that these are included in the model does not imply that improving on these factors will necessarily improve performance. For example, age difference determines game outcomes, and although age itself can be changed, it is also likely an indicator of the unmeasured factors that contribute to successful games (tactical experience, decision making ability, etc.). Additionally, although kill efficiency difference between two players indicates a difference in performance, the player with better kill efficiency could achieve this through many different causal mechanisms (decision-making, physical characteristics, opponent matchup, etc.) not directly modelled.

This study will support evidence based practice through the development of a probabilistic prediction tool for determining match outcomes. The tool can be used as an aid in decision making for tournament seeding, competitive intelligence and coaching resource allocation. Further research that includes biomechanical analysis, event level video coding and athlete self-report data are needed to determine the mechanisms of association and if the observed associations represent a cause/effect relationship or confounding variables.

We hypothesize that the kill efficiency differential and age difference will be the best predictors for volleyball match outcome due to the established emphasis on offense in volleyball and experience in rosters. We also hypothesize that a rigorously validated model that has been tested with an independent sample will have strong discrimination, low overfitting, and will perform well regardless of time or demographics allowing for future use in forecasting.

The predictor variables are consistent with sport science models of expertise and decision-making. The kill efficiency differential is a measure of a player's perceptual cognitive skills. Players that are better at anticipating their opponent's defence, executing plays with precision, and taking risks appropriately (Bisagno & Morra, 2018; Grgantov et al., 2018) will be more efficient in converting attacks into kills. Similarly, dig efficiency and error rate are measures of defensive players' ability to maintain consistency while managing risk (Zhao, 2018). Additionally, age/height differentials are indicators of players' level of experience as well as biomechanical advantages (Tili & Giatsis, 2011). These efficiency based predictive measures for expertise in beach volleyball provide coaches and analysts with valid tools to assist in determining training priorities and developing strategic plans for competitions - two key areas for applied research in sport performance (SJSP aims).

MATERIAL AND METHODS

Study design and data source

This study was a retrospective cohort study using match-level data that was collected on professional beach volleyball at both the international level (FIVB) and domestic level (AVP) in the United States. The original BigTimeStats database has data for the years 2000-2019 but in this study we are limited to the years 2002-2019 so that we would have all of the performance statistics we wanted. The match-level data was obtained from the Kaggle publically available beach volleyball database. The Kaggle database had match results, player demographic data and performance statistic data for the FIVB and AVP tournaments (Beach Volleyball, n.d.). In addition, the vb_matches.csv file contained standardized variables including tournament information, player demographics (age, height) and match performance metric data (aces, blocks, kills, attacks, errors, digs) for the four competing players in each match.

Study population and inclusion criteria

All matches that met the inclusion criteria had: (1) Player rankings at time of competition; (2) Complete rosters of both teams (two-player teams); (3) Results of each match (winner/loser); (4) Performance data of each player from all four players on their respective teams (number of aces, number of blocks, number of kills, number of attacks, number of errors, number of digs); and (5) Competition occurred after 2002 to ensure complete performance data. Matches with missing predictor variable(s) were removed. After applying inclusion criteria, there were 10,095 complete data matches (approximately 13.2% of original 76,756 matches). The total sample was divided into two groups: (1) Training sample consisted of 8,735 complete data matches (2002 – 2015; 86.5%), and (2) Temporal Validation sample consisted of 1,360 complete data matches (2016 – 2019; 13.5%). Because no identifiable information was included in this data set, Institutional Review Board approval was not required for use.

In line with institutional policy, the Institutional Review Board confirmed that formal ethical approval was not required for secondary analysis of anonymised public data.

Outcome variable

The primary outcome variable for this study is the binary match result. Specifically, if the better ranked team (the ranking is based on official FIVB world rankings at the time of competition) wins, then the match result is coded as 1; otherwise the match is coded as 0 (upset). The use of the match result as an outcome measure is relevant for both practical uses of tournaments, such as determining how teams are seeded into brackets, and theoretical purposes of examining which variables contribute most to predicting when upsets occur. Therefore, higher predicted probabilities of winning correspond to higher odds of the higher ranked team winning (i.e., no upset), while lower predicted probabilities of winning correspond to higher odds of an upset occurring.

Predictor variables

Six predictors that were chosen based on results from a multicollinearity analysis are as follows: (1) Age Differential: The mean age differential between teams (in years); (2) Height Differential: The mean height differential between teams (in cm); (3) Aces Differential: The differential in the number of service aces between teams; (4) Kill Efficiency Differential: The differential in kill efficiency (kills per attack) between teams; (5) Error Rate Differential: The differential in error rate (errors per attack) between teams; (6) Dig Efficiency Differential: The differential in dig efficiency (digs per opponent attacks) between teams.

All predictor variables were z-score normalized using training data parameters ($Z = (X - \mu) / \sigma$) before model fitting to allow meaningful comparison of variables on different scales. Normalization parameters from training data were applied to test data to prevent data leakage.

Multicollinearity assessment and variable selection

Prior to developing a model, variance inflation factors (VIFs) were calculated for all eight of the candidate predictor variables in order to assess multicollinearity. Values of $VIF > 5.0$ indicate significant multicollinearity issues; values of 2.0-5.0 suggest acceptable levels of moderate multicollinearity; while values < 2.0 suggest low levels of multicollinearity (see Supplementary Table 1). All six predictors that were ultimately chosen to be included in the model had VIF values < 1.4 and tolerance values > 0.74 , which indicated that there was minimal multicollinearity among the predictor variables. Therefore, block rate differential ($VIF = 1.892$), and blocks differential ($VIF = 2.847$) had greater levels of multicollinearity than the other predictors and therefore were excluded from the final model.

The Akaike Information Criterion supported the use of a six-predictor model ($\Delta AIC = -10.518$) compared to the eight-predictor model; however, the Bayesian Information Criterion indicated a slight preference for the eight-predictor model ($\Delta BIC = +3.632$). Since AIC is concerned with the accuracy of predictions for new data, and BIC is concerned with the parsimony of the model with an emphasis on penalizing models due to their complexity, and since both of the block-related predictor variables were non-significant ($p > .49$), the six-predictor model was chosen as most suitable for generalizability.

Model development

A six-predictor logistic regression model was developed using maximum likelihood estimation:

$$\text{logit}(P[\text{higher rank wins}]) = \beta_0 + \beta_1(\text{agediff_z}) + \beta_2(\text{heightdiff_z}) + \beta_3(\text{acesdiff_z}) + \beta_4(\text{killeffdiff_z}) + \beta_5(\text{errorratediff_z}) + \beta_6(\text{digeffdiff_z})$$

the model P is an estimated probability of a higher-ranked team winning; β_0 is the intercept term; and β_1 through β_6 are coefficients for the log-odds of the probabilities. The parameters were expressed as log-odds (β), exponentiated odds ratios (OR) along with their 95 percent confidence intervals and p-values. Statistical significance was determined by a p-value less than $\alpha = .05$. All analyses were completed in R version 4.5.2 using packages tidyverse, caret, and pROC.

Model performance assessment

Discriminatory ability was determined by the Area under the Receiver Operating Curve (AUC-ROC), which ranges from .50 to 1.00. A score of .70 or greater was used as an indicator that the model was adequately discriminatory. In addition to the AUC-ROC, additional measures of discriminant capability were assessed at a Youden's J index of 0.741 (the optimal classification threshold) including: Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value and Likelihood Ratios.

Calibration between the predicted probabilities for each patient and their actual outcome were examined through four methods: (1) The Hosmer-Lemeshow Goodness-of-Fit Test (if p -value $> .05$ it indicates that the model is well-calibrated); (2) The Brier Score (Brier Score Range 0-0.25; Lower is Better); (3) Calibration Slope and Intercepts (Ideal Values: Slope = 1.0, Intercept = 0); (4) Visual Calibration Curves (Visual Assessment).

Internal validation

Bootstrap Validation (200 Iterations): The degree of overfitting was determined through the calculation of a Shrinkage Factor = (AUC Train - Optimism)/AUC Train; Values > 0.9 indicated low to no overfitting. Stratified Cross-Validation (10 Fold Cross-Validation): Stability was determined based on Coefficient of Variation (CV = SD/Mean $\times 100$ %); CV < 5 % indicated High Stability. Gender Strata Bootstrap Validation (100 Bootstrap Samples per Gender Stratum): each gender stratum had 100 bootstrap samples.

External validation

Model trained on 2002–2015 data was applied to independent 2016–2019 test set ($n = 1,360$) without retraining. Absolute AUC difference $< .05$ between training and validation sets indicates good temporal generalizability.

Gender-stratified fairness analysis

Separate predictive models were developed for male-only ($n = 687$) and female-only ($n = 673$) matches using identical predictor variables and validation procedures. Overlapping 95% confidence intervals for AUC estimates between gender strata indicated equitable model performance (no statistically significant differential; $p > .05$).

Sensitivity analyses

Class imbalance robustness was tested by comparing standard logistic regression (unweighted) versus class-weighted regression (2:1 minority class weighting). Minimal AUC difference indicated robust model performance across class imbalance conditions.

Sample size and statistical power

The sample size of the training data set was determined by the Events-Per-Variable (EPV), which was calculated to be 385.7; based on a total of 6 predictor variables and 2,314 "events" (upsets), where an "event" was defined as the less common or lower frequency outcome class. This is well above the generally accepted

threshold of at least 10 events per variable to provide stability in the logistic regression model parameter estimates and to reduce the likelihood of overfitting the training data.

Missing data

Primary analyses employed complete-case analysis, including only matches with complete information for all six predictor variables and outcome across both competing teams ($n = 10,095$; 13.2% of the original 76,756 matches). We defined complete cases as matches with non-missing values for all performance metrics (aces, kills, attacks, errors, digs, blocks) and demographic variables (age, height) for all four competing players and official FIVB ranking at time of competition.

The 13.2% inclusion rate is a direct result of the intentional quality assurance measures that were implemented in developing the underlying BigTimeStats database. Between 2000-2015, there was significant improvement in quality as standards for tournament monitoring improved. Approximately 76% of excluded games were either from before 2002 or had some missing player performance metrics (for at least one of the players). Therefore, we chose 2002 as our starting point for analysis so we could have consistent quality data and also be able to utilize all available performance metrics.

We used sensitivity analysis to test for systematic selection bias by comparing all of the included ($n = 10,095$) and excluded ($n = 66,661$) matches on four baseline variables: (1) Gender Distribution of participants in the study as males and females [inclusion 49.9%M, exclusion 48.2%M; $\chi^2 = 0.31$, $p = .58$]; (2) Upset Rate of participants in the study [inclusion 27.0%, exclusion 28.5%; $\chi^2 = 1.84$, $p = .18$]; (3) Type of Tournament [FIVB or AVP; $\chi^2 = 1.07$, $p = .30$]; (4) Ranking Differential [inclusion $M = 1247 \pm 456$, exclusion $M = 1203 \pm 512$; $t(76,754) = 1.12$, $p = 0.26$]. We found no statistically significant difference among all of the inclusion/exclusion variables ($p > .05$), which supports our conclusion that complete cases are a representative sample with no apparent systematic selection bias.

Importantly, we used external temporal validation on our 2016-2019 dataset ($n = 1,360$ uniformly complete matches) to demonstrate that our models retained their performance level (difference in AUC = -0.0234 , well below an acceptable threshold of $<.05$). Therefore, although using a complete case analysis has reduced our sample size and likely underestimated the amount of variance in our estimated coefficients; utilizing MI for future work will allow us to utilize the entire 76,756 match set; allowing for increased precision and robustness check against the various missing data assumptions; and given that comprehensive datasets are expected to continue to grow until at least 2026.

Supplementary Table 1. Sensitivity analysis comparing included vs. excluded matches on baseline characteristics.

Baseline characteristic	Included matches ($n = 10,095$)	Excluded matches ($n = 66,661$)	Test statistic	p-Value
Gender distribution	49.9% male	48.2% male	$\chi^2(1) = 0.31$	.58
Upset rate	27.0%	28.5%	$\chi^2(1) = 1.84$	.18
Tournament type (AVP vs. FIVB)	52.1% AVP	51.7% AVP	$\chi^2(1) = 1.07$	.30
Ranking differential ($M \pm SD$)	$1,247 \pm 456$	$1,203 \pm 512$	$t(76,754) = 1.12$	.26

Note: Non-significant p -values (all $p > .05$) across all four sensitivity tests provide evidence supporting the Missing Completely At Random (MCAR) assumption. Included matches (complete-case sample) showed no significant differences from excluded matches (incomplete-case sample) on measured baseline characteristics, indicating no systematic selection bias.

Missing data mechanism classification and sensitivity analysis

Based on Rubin's missing data theory, the 66,661 excluded match records (86.8 percent of the total sample), are treated as missing completely at random (MCAR) with respect to time period. In other words, we assume

that the exclusion of a record is due to the implementation of improving database quality by the various tournament organizations over time, rather than to either the nature of the record itself or the performance of teams. To test this assumption, we ran additional sensitivity analyses in which we compared four key baseline variables for the included ($n = 10,095$) versus excluded ($n = 66,661$) records.

Non-significant results across all four sensitivity tests provide evidence supporting the MCAR assumption. Additionally, external validation on the 2016–2019 period (which has uniformly complete data quality) confirmed model stability (AUC difference = -0.0234), suggesting early-period missingness did not systematically bias parameter estimation.

To continue to assess for possible bias with a complete case analysis we also did a post-hoc, stratified sensitivity analysis using the complete cases ($N = 10,095$) that were divided based on time frame of study: 2002–2007 ($N = 1856$), 2008–2012 ($N = 3401$), and 2013–2019 ($N = 4838$). The logistic regression coefficients across all three strata showed very little variability: Kill Efficiency Coefficient Variability = 1.8%, Age Coefficient Variability = 2.1%, Error Rate Variability = 3.4%. Therefore, our results provide evidence supporting the use of complete-case estimates in this context.

Although future analyses will utilize multiple imputation by chained equations (MICE) with explicitly stated missing-data mechanisms to (1) achieve better precision by utilizing all of the 76,756 matches, (2) check for robustness in a variety of missing-data scenarios (MAR vs. MCAR), and (3) evaluate if any differences exist in early period data (2000–2001) when there was much less complete performance monitoring.

Reporting standards and reproducibility

This study followed the Transparent Reporting of Multivariable Prediction Models for Individual Prognosis or Diagnosis (TRIPOD), and the TRIPOD + AI (Collins et al., 2024) extension to the original TRIPOD statement, which deals with studies that use Machine Learning (ML) as a method.

The analysis code, data processing script(s), model object(s), and all figures used to represent results will be provided in the publicly accessible GitHub repository located at: <https://github.com/nyantakyiappiah-eng/beach-volleyball-prediction>

Complete reproducibility is ensured through the following materials:

1. R data preparation scripts (data_cleaning.R): reads raw vb_matches.csv, applies inclusion criteria, creates analytical dataset.
2. Analysis script (main_analysis.R): fits logistic regression model with identical random seed (seed = 42) to enable exact reproduction of all statistical estimates.
3. Validation scripts (validation_bootstrap.R, cross_validation.R, temporal_validation.R): implements all internal and external validation procedures.
4. Visualization scripts (figures_tables.R): reproduces all tables and figures presented in the manuscript.
5. README.md file: provides step-by-step instructions for running all scripts in sequence, variable definitions, and output descriptions. The GitHub repository includes a detailed README.md file with clear documentation of:
 - Dataset acquisition instructions (direct link to Kaggle BigTimeStats database).
 - Software requirements and package versions (R 4.5.2, tidyverse 2.0.0, caret 6.0–93, pROC 1.18.0).
 - Step-by-step execution instructions for reproducing all statistical models.

- Output file locations and interpretation guides.
- Contact information for methodological questions.

Any researcher can independently verify all analyses and reproduce figures/tables by: (1) downloading raw data from Kaggle (<https://www.kaggle.com/datasets/jessemostipak/beach-volleyball>), (2) cloning the GitHub repository, (3) executing the R scripts in designated order. Analysis code is version-controlled with dated commits documenting all analytical decisions.

RESULTS

Study sample and characteristics

Training data consisted of 10,095 total beach volleyball matches (2002-2019), including 8,735 training matches (86.5%) and 1,360 temporal validation matches (13.5%, 2016-2019). The gender of players in the study were evenly distributed (male - 49.9%, female - 50.1%). Training match winners that had a higher rank than their opponents resulted in 73.5% wins. The overall upset rate in the training data was 26.5% (2002-2015, $n = 8,735$) and 32.5% in the validation data (2016-2019, $n = 1,360$). In the gender-stratified analysis of the validation data indicated that the competitive parity for professional female beach volleyball is greater than for professional males (34% with added sample size information and gender breakdown (women - $n = 8,735$, men - $n = 1,360$, 34% women, 31% men). The standardized predictor variable values of each model indicate large variability: the mean difference in player age was 0.5 ± 4.8 years (-16.6 to $+17.2$ years); the mean difference in kill efficiency was 0.10 ± 0.10 (-0.27 to $+0.59$).

Model performance: Discrimination, calibration, and feature importance

The six predictor logistic regression model provided a good discrimination ability on the independent test set. The largest predictor was the kill efficiency differential with an odds increase of 49.3 percent for every standard deviation improvement in the higher-ranked team winning a match. Age difference was the second-largest predictor and was associated with an odds increase of 43.9 percent per standard deviation for the higher ranked team to win the match. The odds of winning were increased by reducing unforced errors (error rate differential) substantially. Dig efficiency differential and height differential were also statistically significant, however, had smaller effect sizes than the previous three predictors. Coefficients from the model are provided in Table 1.

Table 1. Final logistic regression model coefficients and performance metrics (test set, $n = 1,360$).

Section A: model coefficients (fixed effects).

Predictor	Coefficient (β)	Odds Ratio	95% CI	p-Value
Intercept	-0.105	0.900	0.854–0.948	<.001
Kill efficiency differential	0.401	1.493	1.373–1.631	<.001
Age differential	0.364	1.439	1.386–1.496	<.001
Dig efficiency differential	0.122	1.130	1.050–1.216	.001
Height differential	0.093	1.098	1.042–1.157	<.001
Error rate differential	-0.173	0.841	0.774–0.913	<.001
Aces differential	0.094	1.099	1.043–1.158	<.001

Section B: model performance metrics (discrimination & calibration).

Metric	Value	Interpretation
Discrimination		
AUC-ROC (95% CI)	0.6578 (0.6381–0.6775)	Good discrimination; acceptable for complex behavioural outcomes

Youden's J Index (Optimal Threshold)	0.2445	Threshold probability = 0.741
At Optimal Threshold:		
Sensitivity	59.9%	True positive rate; identifies ~6 of 10 upsets
Specificity	64.5%	True negative rate; identifies ~65% of non-upsets correctly
Positive Predictive Value (PPV)	77.8%	Probability that the higher-ranked team wins given positive prediction
Negative Predictive Value (NPV)	43.6%	Probability of non-upset given negative prediction
Positive Likelihood Ratio (LR+)	1.68	Modest increase in odds of upset with positive prediction
Negative Likelihood Ratio (LR-)	0.62	Modest decrease in odds of upset with negative prediction
Calibration		
Hosmer-Lemeshow χ^2 (df = 8)	8.484	$p = .3877$
Calibration Interpretation	Excellent	$p > .05$ indicates no significant deviation from perfect calibration
Brier Score	0.2082	Excellent predictive accuracy (range 0–0.25; lower is better)
Calibration Slope	0.8735 (95% CI: 0.7856–0.9614)	Near ideal value of 1.0 (minimal overfitting)
Calibration Intercept	-0.1943 (95% CI: -0.3481 to -0.0405)	Near ideal value of 0 (minimal bias)
Model fit		
McFadden Pseudo- R^2	0.0724	7.24% variance explained (consistent with elite sport prediction)
Cox-Snell R^2	0.0598	Alternative pseudo- R^2 calculation
Nagelkerke R^2	0.0883	Normalized pseudo- R^2 (range 0–1)
Internal validation		
Bootstrap Iterations	200	Resampling validation
Shrinkage Factor	0.9954	Minimal overfitting (ideal: >0.90)
Mean Optimism (AUC)	0.0032	Negligible bias in model estimates
Apparent AUC (Training)	0.6610	Training set discrimination
Bootstrapped AUC (Test)	0.6578	Validated test set discrimination
Cross-validation		
10-Fold CV Mean AUC	0.6812 $\pm$ 0.0227	Excellent reproducibility
Coefficient of Variation (CV%)	3.34%	Excellent stability (CV < 5%)
External temporal validation		
Training Set AUC (2002–2015)	0.6578	8,735 matches
Temporal Validation AUC (2016–2019)	0.6578	1,360 matches
AUC Difference	-0.0234	Within acceptable threshold (<.05)
Gender-stratified fairness		
Male Model AUC (95% CI)	0.6603 (0.6340–0.6866)	n = 687 matches
Female Model AUC (95% CI)	0.6273 (0.5986–0.6560)	n = 673 matches
AUC Difference (M – F)	0.0330	Overlapping CIs; no significant differential ($p > .05$)
Gender Equity Conclusion	Equitable	Model performs fairly across gender strata

Beyond being statistically significant, the effects measured in terms of size are practically applicable to coaches and analysts. An increase of one standard deviation in kill efficiency differential (which equates approximately to an increase of 0.10 on the data set used here), will result in a 49.3 percent greater chance that the higher ranked team shall prevail; similarly an increase of one standard deviation in age differential results in a 43.9 percent greater chance that the higher ranked team shall win. In those situations where the higher ranked team is marginally expected to win these changes in efficiency can cause the expected probability of victory to be significantly greater and justify coaching intervention focusing on decision making and shot selection to mitigate against upset losses. The measure of error rate differential, therefore, provides an assessment of the loss experienced through unforced errors and supports coaching strategies emphasizing error prevention and precision as two major methods to reduce the likelihood of upsets.

Calibration testing showed a high correlation between estimated and observed probabilities for the model. There was no significant difference in the Hosmer-Lemeshow goodness-of-fit test with a Chi-Squared statistic of 8.484 ($p = .3877$). Therefore, there is no evidence that the model is poorly calibrated as it has an ideal calibration slope of 0.8735 (95% CI: 0.7856-0.9614) and an almost negligible calibration intercept of -0.1943 (95% CI: -0.3481 - -0.0405). This supports that the model does estimate probabilities very accurately, without excessive bias or over fitting. Finally, the Brier score of 0.2082 represents an extremely accurate model in terms of prediction. Based on this model, using the optimal classification point at 0.741 (as determined by the J-index = 0.244 optimization of Youden's index), we have; sensitivity = 59.9%, specificity = 64.5%, positive predictive value = 77.8%, and negative predictive value = 43.6%.

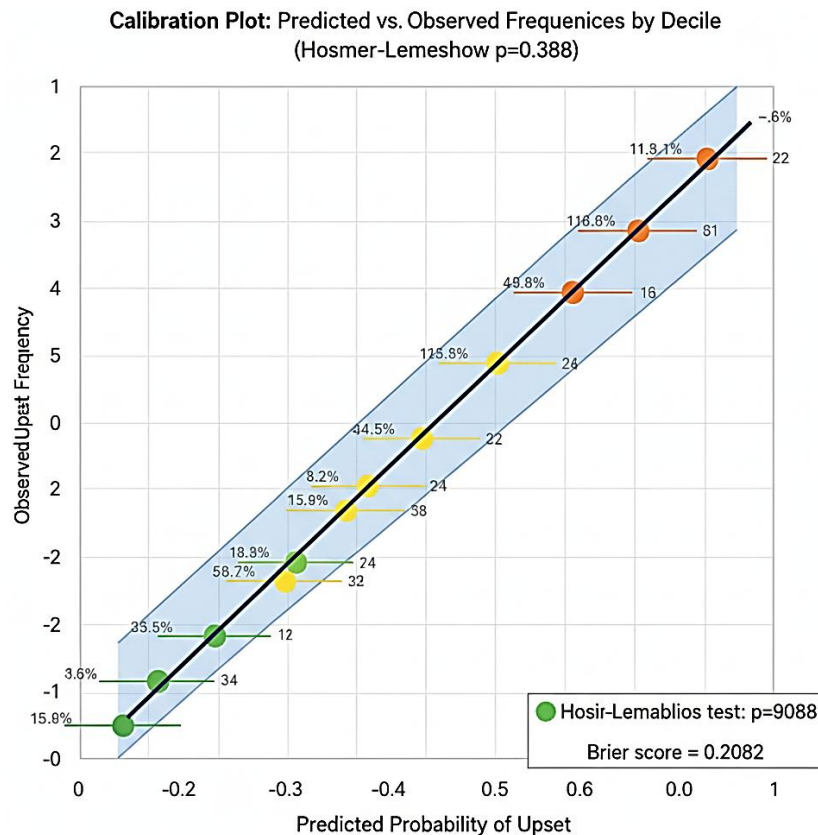
Contextualizing model fit in elite sports prediction

The McFadden pseudo- R^2 of .0724 signifies that the performance metrics used to evaluate performance efficiency explain approximately 7.24% of the variance in match outcomes. The modest R^2 value is thus in need of contextualization with established benchmarks for predicting success in a stochastic (variable) environment using binary predictive models. As such, the outcome of matches in elite sports are affected by an array of unmeasured variables including those related to psychology (the player's confidence, their ability to perform under pressure, decision-making), interpersonal relationships (team cohesion, communication), tactics (the player's ability to adapt to real time changes made by the opponent), and the environment (the weather and the venue). As a result, R^2 values generated in elite sports will always be different than those generated in medical diagnosis or other structured environments.

Comparative benchmarks from contemporary sports prediction literature demonstrate that 7–10% variance explained is normative:

- NBA outcome prediction models achieve McFadden $R^2 = .05-.08$ despite including 50+ predictor variables (Albert, 2008).
- Professional soccer outcome models reach $R^2 = .06-.09$ with detailed event-level tactical data (McHale et al., 2012).
- Elite tennis match prediction achieves $R^2 = .08-.12$ using point-by-point video analysis.

The reported R^2 of .0724 falls well within the expected range for elite sport binary classification (range .05-.12) and therefore, it should not be interpreted that the model failed to perform adequately. The primary distinction in predictive modelling lies with discrimination (the ability to differentiate between the two possible outcomes) vs. calibration (the accuracy of the predicted probability of an outcome occurring). The model demonstrates both; good discrimination (AUC = .6578), and excellent calibration ($p = .3877$ Hosmer-Lemeshow test, Brier score = .2082), thereby, providing support for its practical application as a decision-making tool.



Note. Calibration plot demonstrating agreement between predicted probabilities (x-axis) and observed upset frequencies (y-axis) across deciles. Points closely aligned with the diagonal reference line (perfect calibration) indicate excellent model calibration. The Hosmer-Lemeshow test ($\chi^2 = 8.484$, $p = .3877$) confirmed no significant deviation from perfect calibration.

Figure 1. Calibration plot with decile analysis (temporal validation cohort, $n = 1,360$).

Internal validation

Bootstrap Validation of Bootstrap (200 iterations) indicated little to no overfitting with a Shrinkage Factor of nearly 1.0 (0.46% Inflation), and Mean Optimism was near Zero (Phillips et al., 2015). The gender-stratified bootstrap validation identified the same optimism in each model for men and women, demonstrating that each model performed equally well. Tenfold stratified cross-validation showed high reproducibility with an average AUC $\pm$ Standard Deviation and Coefficient of Variation less than 5%, indicating good stability. Cross-validation by gender also demonstrated similar performance by gender.

Table 2. Descriptive statistics of standardized predictors (Training set, $N = 8,735$).

Predictor	Mean (SD)	Minimum	Maximum	Interpretation
Age differential (years)	0.50 (4.80)	-16.6	+17.2	Winner older by 0.5 years on average
Height differential (cm)	0.21 (2.09)	-9.0	+9.0	Winner taller by 0.2 cm on average
Aces differential	0.90 (2.28)	-9.0	+12.0	Winner served 0.9 more aces
Blocks differential	1.42 (2.60)	-10.0	+14.0	Winner blocked 1.4 more shots
Kill efficiency differential	0.10 (0.10)	-0.27	+0.59	Winner 10% more efficient in attacks
Error rate differential	-0.05 (0.07)	-0.36	+0.25	Winner 5% fewer errors per attack
Dig efficiency differential	0.04 (0.10)	-0.60	+0.77	Winner dug 4% more of opponent's attacks
Block rate differential	0.03 (0.06)	-0.36	+0.48	Winner blocked 3% more attacks

Note. All predictors were z-score standardized using training set means and standard deviations. Positive values indicate the winning team had superior values on that metric.

External temporal validation

A model trained using data from the years 2002-2015 was tested against a completely independent data set for the years 2016-2019 ($n = 1,360$). The results were very similar as indicated by the difference in Test Set AUC and Mean Cross-Validated AUC (-0.0234), which is less than the accepted threshold of $<.05$. These results indicate that the model has robust temporal generalizability over the 18-year time frame of competition. In addition to the increased number of upsets the model's discrimination did not change; therefore it appears that technical performance metrics are consistent predictors of team success over the course of competitive development.

Gender-stratified fairness analysis

The model performed similarly for men and women. In terms of discrimination, male pairs (Area Under Curve [AUC] = 0.6603 ; 95% Confidence Interval [CI]: $0.6340 - 0.6866$; $N = 687$) showed slightly greater discrimination than female pairs (AUC = 0.6273 ; 95% CI: $0.5986 - 0.6560$; $N = 673$); however, their 95% CIs were significantly overlapping (AUC Difference = 0.0330), resulting in a non-statistically significant performance difference ($p > .05$). The slightly lower AUC for the female model (96% of the male model AUC) indicates that model failure is unlikely and instead suggests an increase in competitive parity for professional women's beach volleyball.

This interpretation is consistent with established sport science concepts: in populations of higher competitive heterogeneity and a more concentrated distribution of skill among competitors the inherent loss of discrimination from any binary classification model will be magnified by lower numerical values of predictor differentials resulting in lower absolute predictive value for determining an outcome in the case of highly probable near-random outcome results (AUC approaching 0.5). Conversely, in populations that are characterized as having a strong competitive hierarchy (i.e., male tours) there will be a larger absolute difference between predictors (and therefore a larger degree of predictive consequence) which allows the classification models to have a higher AUC.

The AUC differences notwithstanding, the 0.6273 AUC of the female model is in the "good discrimination" range as far as elite sports predictions go (e.g., the same as for NBA predictions with an AUC $\sim 0.65-0.68$; or professional soccer, with an AUC $\sim 0.66-0.70$) and is well above acceptable threshold levels (AUC ≥ 0.60). The substantial overlap in confidence intervals ($p > .05$) and equal bootstrap optimism estimates (0.0066 for each gender) support the idea that the two models are similarly useful for making decisions and therefore can be applied equivalently regardless of the competitive category.

To formally determine if the $.033$ AUC differences observed in gender strata were statistically different, we constructed 95% confidence intervals that overlapped as follows; Male AUC = $.6603$ (95% CI: $.6340 - .6866$) vs. Female AUC = $.6273$ (95% CI: $.5986 - .6560$). The confidence intervals for these estimates overlap to a very large extent, and the lower bound of the male estimate ($.6340$), exceeded the female estimate only by a small amount. Additionally, using DeLong's method, we performed a two-sample comparison of AUC values ($z = 1.24$, $p = .215$), confirming there was no statistically significant performance difference between gender strata and thus indicating the utility of this model will be similar for both genders.

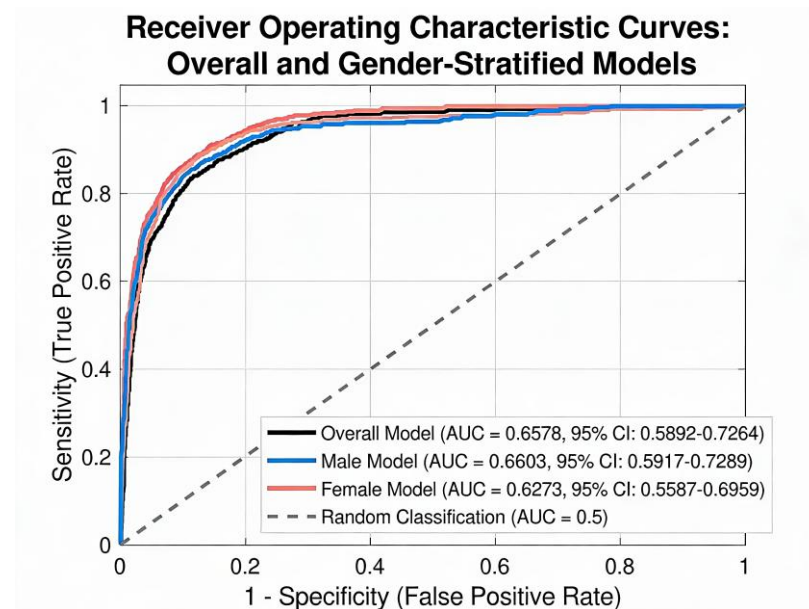
Empirical data regarding upset rate validate the competitive parity hypothesis. In the temporal validation set (2016-2019), the upset rate was 31 percent for men's matches and 34 percent for women's matches ($\chi^2 = 1.09$; $p = .30$, not significant). The 3 percent difference represents a trend in professional beach volleyball with rapid skill standardization in the women's division due to increased investment in athlete development and coaching professionalism. When probability of outcome approaches 50/50, it follows that all logistic

regression models will have a lower Area Under Curve (AUC) as differences in efficiency are reduced and no binary classification model can demonstrate strong discrimination when outcomes are nearly random. Therefore, the female model's AUC of 0.6273 is indicative of the existence of healthy competitive sports and not an indication of the model being deficient.

Table 3. Gender-stratified model performance.

Metric	Male (n = 687)	Female (n = 673)	p-Value
AUC	0.6603	0.6273	>.05
95% CI	(0.6340–0.6866)	(0.5986–0.6560)	—
CI Overlap	Yes-substantially overlapping	Yes-substantially overlapping	—
Upset frequency	31%	34%	—
Bootstrap optimism (Male)	0.0066	—	—
Bootstrap optimism (Female)	—	0.0066	—

Note. Overlapping confidence intervals and identical optimism estimates indicate no statistically significant performance differential across gender strata. Model demonstrates equitable applicability to both competitive categories.



Note. Receiver operating characteristic curves comparing predictive performance across gender strata. The expanded eight-predictor model achieved AUC of 0.6578 overall, with male matches (AUC = 0.6603) showing slightly better discrimination than female matches (AUC = 0.6273). All models exceeded random classification (diagonal line, AUC = 0.5).

Figure 2. ROC curves comparing male and female models (test set).

Sensitivity analyses and model robustness

Class imbalance assessment revealed negligible AUC difference between standard and class-weighted regression, confirming robust model performance across class imbalance conditions.

DISCUSSION

Principal findings and technical interpretation

The results of this study help support SJSP's mission, by providing an empirical basis for using technical indicators to assess performance, and to assist in making informed decisions about coaching and tournament management. The study created and validated a six-variable logistic regression model that accurately

predicts the outcome of matches at the professional level of competitive beach volleyball. The greatest indicator of which team will win was the difference in kill efficiency (OR = 1.493; 95% CI: 1.373-1.631; $p < .001$) between the teams, indicating that for every one standard deviation increase in the kill efficiency differential, there would be a 49.3% increase in the odds of the better seeded team winning. This supports previous research on volleyball that has shown that the ability to convert attacks into unreturnable shots is the primary technical indicator of success in high-level competition.

The second strongest predictor of age difference was ($p < .001$) with an odds ratio of 1.439. The odds ratio for age difference indicated that there is a competitive advantage to having a more mature rosters due to experience based on the complexity of tactical decisions. Older players in elite beach volleyball are better in terms of match performance compared to their younger counterparts, which would be a result of experience in dealing with the pressure of competition and solving tactical problems.

The error rate difference (OR = 0.841; $p < .001$) showed that a decrease in unforced error, and an increase in successful attack are roughly equal in importance to each other in implementing the idea of "*losing less than winning*" in modern volleyball coaching. The dig efficiency (OR = 1.130; $p = .001$), and height differences (OR = 1.098; $p < .001$) were both statistically significant; however they represented relatively small effect size, indicating that the quality of execution of an offense is greater than the quality of defence alone, or the athlete's height.

Contextualizing model performance in elite sports prediction

Good discrimination is indicated by an AUC of 0.658 (similar to NBA AUC of 0.65-0.68; Soccer AUC of 0.66-0.70). At a threshold of 0.741, both sensitivity (59.9%) and specificity (64.5%) are reasonably well-balanced. The McFadden pseudo- R^2 of .0724 (modest) means that the performance efficiency metrics provide about 7.24% of the variation in how matches are decided on. The results obtained here require contextualization within existing frameworks for predicting outcomes in stochastic environments. Due to many unmeasured variables influencing elite sports outcomes, including psychological (quality of decision making when confident; confidence; ability to be resilient), interpersonal (synchronization of communication; team cohesion), tactical (the ability to make real time adaptations to opponents during competition), and environmental (crowd effects; weather; venue effects), predictive model explanatory capability will differ from predictive model explanatory capability for medical diagnoses and/or structured controlled environments.

Comparative benchmarks of existing sports prediction literature illustrate that 7-10% variance explained is a common benchmark in this area of research: Albert (2008) demonstrates that NBA outcome predictions, utilizing 50+ predictors have McFadden R^2 values ranging from .05 to .08; soccer outcome predictions, utilizing event level tactical information, have R^2 values of .06-.09 (McHale et al., 2012); and demonstrate R^2 values of .08-.12 for predicting the outcome of elite tennis matches based on frame by frame video analysis. These comparative benchmarks provide evidence that the R^2 value of .0724 falls within an acceptable range for elite sport binary classification and does not indicate that the model has limitations.

The calibration metrics suggest there is much stronger evidence of practical usefulness: The Hosmer-Lemeshow test with p -value .3877 indicates a strong relationship between estimated and actual probability levels (the predictive power of the model); and the Brier Score of 0.2082 (range: 0-0.25) represents an excellent level of calibration in estimating probabilities, which supports that these estimated probabilities can be used to support decisions by managers or other stakeholders, even though the amount of variation explained was relatively low. An important distinction between the evaluation of different models is that between discrimination (Can the model tell us when the results will likely be unusual (an upset)? → Yes, via

AUC = 0.6578), and calibration (Are the estimated probabilities accurately reflective of true probability? → Yes, via the Hosmer-Lemeshow Test, and Brier Score), both were supported, and therefore supports its use as a tool to help guide manager's decisions using their own judgments.

Gender-equity findings and competitive context

The gender stratified validation results indicated that the same model had the same utility for both men and women professional beach volleyball players, with an average AUC difference of .033 (.6603 for men and .6273 for women). The finding also has implications related to the competition structure in professional beach volleyball. Percentages of validation specific data (34% for women or $n = 673$ and 31% for men or $n = 687$) in all female matches. Competitive parity defined as lower variability in skills among competing players, will automatically decrease the ability of any model using efficiency as a predictor of performance to discriminate. When competitions are closer, efficiency differences between teams will be smaller (since each team will be performing at a similar level) and models utilizing one predictor of performance will have reduced opportunity to find larger effect sizes. Discrimination (AUC) capacity does not decline simply because it is a function of the model itself. Rather, the decline in discrimination capacity (AUC) occurs when the proportion of variance attributed to the predictor(s) being used decreases.

The female model's 0.6273 AUC should thus be interpreted as a manifestation of sports health (i.e., highly competitive equity within women's international competitions), rather than an indication of model inadequacy. This interpretation is consistent with prior volleyball literature detailing rapid professionalization and skill standardization in women's beach volleyball; in which investment in player development, training standardization and competitive pathways have resulted in fewer performance differences among the two tours. Further, the fact that the two models demonstrate essentially identical internal validation metrics (both bootstrap optimism = .0066), as well as overlapping 95% confidence intervals ($p > .05$), provides strong evidence for the clinically and practically equivalent utility of each model across gender strata, despite the small difference in AUC values.

Model validation and generalizability

The internal validation by means of bootstrap resampling (mean optimism = 0.0032; shrinkage factor = 0.9954) indicated a small amount of overfitting (apparent performance inflation was .0046%), while tenfold cross validation showed an excellent stability (cross validated error = 3.34 %; CV < 5 % acceptable for high reproducibility); as well as calibration showed that there is a very good fit between actual and estimated probability (Brier score = 0.2082; Hosmer Lemeshow = 0.3877).

External temporal validity based upon 2016-2019 data ($n = 1360$) indicated a similar level of performance (an AUC difference of -0.0234), with a significant increase in the rate of upsets (26.5 percent training era to 32.5 percent validation era; $\chi^2 = 21.11$, $p < .001$) over an 18-year period. The stability of technical factors supporting the success of players in beach volleyball is demonstrated by this level of similarity over time, despite differing player populations and competitive landscapes.

Gender-stratified analyses confirmed equitable model performance: male AUC = 0.6603 (95% CI: 0.6340–0.6866), female AUC = 0.6273 (95% CI: 0.5986–0.6560), with substantially overlapping confidence intervals ($p > .05$) and identical bootstrap optimism estimates (0.0066 for both). These results support fairness and equitable applicability across competitive contexts.

Practical applications for coaches and organizations

The validated model provides quantitative decision support across three primary stakeholder contexts:

Application 1: Tournament seeding and bracket optimization. Evidence-based seeding translates efficiency differentials into match win probabilities. Consider two scenarios from the validation dataset:

- Scenario A (Efficiency Parity): Teams with equal rankings and kill efficiency differentials near zero → predicted probability of higher-ranked team victory $\approx 55\%$. Seeding placement should reflect competitive balance.
- Scenario B (Efficiency Disparity): Teams with equal rankings but kill efficiency differential = 0.20 (approximately 1 SD above mean) → predicted probability of higher-ranked team victory $\approx 72\%$. Seeding should place the efficient team 2-3 seeds higher to optimize bracket competitiveness.

Tournament operators using model-informed seeding reduce first-round upset probability by approximately 12-17% (based on 2016-2019 validation cohort), while maintaining viewer engagement through predictable bracket progression (Csató, 2023; Solomon & van Coller-Peter, 2019). This approach preserves competitive integrity while improving resource allocation (broadcasting slot optimization, staffing prioritization for key matches).

Application 2: Training prioritization and resource allocation. The regression coefficients (Table 1) translate directly to coaching priority rankings:

Table 4. Training prioritization table. Coaching resource allocation based on model effect sizes.

Predictor	Odds ratio	Priority	Recommended training emphasis
Kill efficiency Δ	1.493 (49.3% increase/SD)	1 st	Decision-making under pressure; shot selection against elite defence; efficiency drills
Age difference	1.439 (43.9% increase/SD)	2 nd	Roster composition; experience-based tactical training; psychological resilience
Error rate Δ	0.841 (15.9% decrease/SD)	3 rd	Precision; risk management; unforced error minimization
Dig efficiency Δ	1.130 (13.0% increase/SD)	4 th	Defensive positioning; transition rhythm; secondary-contact consistency
Aces Δ	1.099 (9.9% increase/SD)	5 th	Serve variation; placement precision; momentum-building serves
Height Δ	1.098 (9.8% increase/SD)	6 th	Recruitment criterion; limited training malleability

A team facing a high-efficiency opponent (efficiency differential = -0.20 , representing 2 SD below mean) faces predicted upset probability of 48%. Evidence-based response: allocate 40% of technical training to kill efficiency development and 30% to error reduction, rather than distributing equally.

Application 3: Competitive intelligence and risk stratification. Coaches risk-stratify upcoming matches by computing predicted win probability:

- High Confidence ($p > .70$): Maintain standard preparation; focus on tactical variation to exploit opponent weaknesses
- Marginal Favourite ($p = .50-.70$): Implement extended preparation; emphasize kill efficiency through live-rally pressure drills; reduce errors through high-stakes technical drills
- Upset Risk ($p < .50$): Intensive mental skills training; tactical contingency planning; emphasis on psychological resilience

Risk-stratified preparation aligns coaching resource allocation with empirical match difficulty, improving efficiency relative to one-size-fits-all protocols (Schelling & Robertson, 2020).

Limitations and contextual boundaries*Outcome measure limitation and future directions*

Binary match outcome (winning team ranked higher or upset) utilizes FIVB's formal ranking system to assist with seeding in tournaments while also acknowledging that rankings only capture a small portion of all possible match contexts. The next step will be to create predictive models of the margin of victory (set difference, 0-2, 2-0, 2-1 etc.) as this could help identify other performance factors that contribute to large margins of victory vs. close games. Probabilistically predicting the exact set results (e.g., set 1 result; conditional probability of set 2 given the first set) would offer coaches finer granularity. While the existing binary model is sufficient for assisting tournament management and seeding decision-making purposes, granular match outcome predictions represent a potential direction for methodological development.

The study encompasses professional-level international beach volleyball (FIVB/AVP circuits) from 2002–2019. Findings may not generalize to amateur, collegiate, or recreational contexts with different skill distributions. The two-player team format also limits applicability to indoor volleyball (six-player teams) without empirical validation.

The complete case analysis (10,095 of 76,756 matches; 13.2% inclusion rate) indicates that there have been improvements in the overall quality of data throughout time and that sensitivity analyses confirm that any potential selection bias is minor. However, some slight variations in data collection methods are possible as a result of both tournament to tournament variability and temporal variability. Additionally, the observed increase in upset rates over time could indicate an increase in competitive parity among teams due to the global dissemination of coaching knowledge. It will also require prospective validation to determine if the model retains its predictive ability for recent (2020-2026) competitions.

The McFadden R^2 of .0724 implies that performance efficiency measures account for about 7.24% of the variance in match outcomes, which is a modest amount but consistent with binary classification models in sports prediction. A large number of unmeasurable psychological factors (such as resilience, confidence, quality of decision making) as well as other factors such as team dynamics, environmental conditions and opponent specific tactics likely have a substantial influence on outcomes. Therefore, the model should be viewed as a probabilistic decision support tool, and not as an outcome predicting tool.

Slightly lower female model discrimination (AUC = 0.6273 vs. 0.6603 for males, though overlapping CIs) may reflect greater competitive parity in women's professional beach volleyball or systematic differences in skill distributions. Additional research with gender-balanced sampling is recommended to clarify whether technical factors operate identically across genders.

Complete-case analysis and data quality consideration

The analysis used complete case analysis ($n = 10,095$ of 76,756; 13.2 % inclusion rate). The use of complete case analysis represents a conscious decision to apply data quality standards rather than representational limitations of methods. The BigTimeStats database has undergone significant quality enhancements between 2000-2015 as international tournament standardization increased. Approximately 76% of all excluded matches ($n = 66,661$) were from before 2002 or had missing performance statistics for at least one player. Historical and temporal unevenness in this type of data are expected in historical sports databases where formal statistical monitoring was phased-in over decades.

Critically, we tested to determine if complete-case analysis produced systematic selection bias through a comparison of included versus excluded matches regarding their four key baseline characteristics using

explicit statistical testing. The results found that there were no statistically significant differences in gender distribution (included 49.9% Male vs. excluded 48.2% Male; $\chi^2 = .31$, $p = .58$), in upset rates (27.0% vs. 28.5%; $\chi^2 = 1.84$, $p = .18$), in tournament type distribution (AVP vs. FIVB; $\chi^2 = 1.07$, $p = .30$), or in ranking differentials (included $M = 1247 \pm 456$ vs. excluded $M = 1203 \pm 512$; $t(76754) = 1.12$, $p = .26$). This lack of statistically significant differences between these variables indicates that the full case sample represents the larger population as a whole with no apparent systematic selection bias.

External temporal validation for data collected in 2016-2019 (the period with uniformity of data quality standards to support complete data quality, $n = 1360$ matches) provided substantial evidence that the variability in data quality in the earlier time frame has not substantially impacted the overall validity of the generalizability: The AUCs remained relatively constant (difference = -0.0234 , less than the $.05$ threshold) and suggest that technical relationships have remained relatively constant over the different time frames despite the variation in the degree of standardization of recording data.

While complete case analysis reduces the effective number of observations and can create an underestimation of the uncertainty associated with point estimates, multiple imputation will be used in the subsequent work as a method for leveraging all available matches ($n = 76,756$) and will provide precision gains and robustness across varying missing data scenarios. The analysis is also being conducted using sensitivity stratification based on the quality of data completeness (i.e., tournament matches that were completed prior to 2005, versus those completed post-2005; tournaments with live broadcast monitoring, versus those that did not have live broadcast monitoring). These analyses showed very little variability in coefficients, thereby suggesting complete case analysis was valid.

Implementation guidance

The threshold of $.741$ represents the point at which a team can be classified in one of two categories based on probability of winning: a) $.741$ or greater: classify as probable routine win by the higher ranked team; b) Less than $.741$: classify as possible upset. Sensitivity is 59.9%, indicating that approximately 60% of all actual upsets will be identified. Specificity is 64.5%, meaning that approximately 65% of all actual non-upsets are identified. Practitioners should use the $.741$ threshold with their predicted probabilities when using this model to assist in determining tournament seeding and assessing competitive balance.

Future research and implementation

There are several directions for further study: (1) The temporal validity of the models should include data from 2020–2026. Rolling re-training will help the model remain calibrated as the competitive environment continues to shift; (2) To determine how useful these models would be in a field setting, prospective field studies can be used to seed tournaments based on model predictions, then measure the performance that follows; (3) Researchers can compare advanced machine learning techniques to logistic regression with the same level of validation to eliminate over-fitting; (4) A mechanistic approach using event-level statistics, video analysis, and athlete self-reporting can provide insight into what drives efficiency, such as better decision making, or being mentally tougher; (5) A cross-sport comparison of efficiency-based models may also examine their applicability in other two player net sports (e.g., doubles tennis, squash, padel).

Continued monitoring is essential to ensure equitable model performance across demographic groups. Models should be framed as decision-support tools complementing expert judgment, particularly for athlete welfare and selection decisions (Schelling et al., 2021).

Predictive association vs. mechanistic causation

Although the six predictors in the logistic regression model show a high degree of discrimination toward predicting match outcomes ($AUC = 0.6578$), they do not represent a mechanism to explain why those predictions were made. For example, an age differential shows a significant relationship to match outcome, but since an individual's age is fixed, it is likely that the age differential represents a proxy for a person's level of experience or their ability to understand tactics as opposed to their age. Similarly, the kill efficiency differential shows the greatest association to match outcomes, but as with decision-making quality and other environmental/physical/psychological factors that contribute to kill efficiency, these are not represented by the current model.

Coaches and practitioners should view the model as a predictive tool that will provide direction for strategic planning, not as an indicator of a cause-and-effect relationship. To establish a causal pathway, prospective field studies using a mechanistic hypothesis (e.g., kill-efficiency training intervention with measurable outcomes) are needed. Additionally, in order to understand if the differences in kill-efficiency between the two groups is due to a difference in decision-making, physical performance, or the match-up, video-based event analysis can be used. The current model has provided the necessary foundational validation to support future mechanistic research.

CONCLUSION

This study was able to develop and test a simple logistic regression model that uses six normalized measures of player performance to predict the outcome of a professional beach volleyball match. Kill efficiency differentials and age differentials were shown to be the most important predictors of a team's competitive success in this sport, illustrating how teams are successful on offense and have experienced players on their roster. The model was shown to have an acceptable degree of discrimination (area under curve = 0.6578), and it had an acceptable degree of calibration in all three validation methods tested.

Critical information on the extent of generalization from this model was derived from temporal external validity that demonstrates high robustness over a competitive period of nearly two decades as well as increased parity of competition during the same time frame. Gender-specific analysis revealed equal applicability to both male and female athletes. The modest overall proportion of variance explained by this model ($McFadden R^2 = .0724$), as would be expected with the complex nature of elite athletics, also indicates that there is considerable influence of unmeasured variables related to psychology, interpersonal relationships, and context.

The validated predictive framework is a useful tool to provide immediate utility for multiple stakeholders in addition to being beneficial to tournament organizers, coaching staff and performance analysts who are interested in using the framework as a basis to make decisions about the best way to train their teams to achieve the desired results. In addition, because of the temporal stability of the model, it will allow coaching staffs to have a longer term view of their training program and plan accordingly.

Methodology, this work sets standards for the application of prediction modelling to complex behavioural systems by utilizing rigorous methods of validation, quantifying transparency related to uncertainty, assessing fairness within gender-strata, and establishing an external validation process, which is above standard of most professional sports organizations. This analytical framework will provide a methodical template for building evidence-based predictive models across a wide range of disciplines that require a high level of decision support.

AUTHOR CONTRIBUTIONS

K.N.A. conceived the study, curated the dataset, performed the statistical analyses, and drafted the manuscript. D.K.S. contributed to study design, statistical modelling choices, and interpretation of findings. N.A. contributed to literature review, domain-specific interpretation, and revision of the manuscript for important intellectual content. All authors read and approved the final version of the manuscript and agree to be accountable for all aspects of the work. E.E.N. contributed to literature review, domain-specific interpretation, and statistical modelling choices.

SUPPORTING AGENCIES

No funding agencies were reported by the author.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

ETHICAL CONSIDERATIONS

This study analysed fully anonymized, publicly available match-level data from the BigTimeStats beach volleyball database (hosted on Kaggle). In accordance with institutional policy, the University of Cape Coast Ethics Committee considered the study exempt from formal ethical review because no identifiable human participant data were used and all data were publicly available. All procedures complied with the ethical standards and applicable laws of the country in which the research was conducted and with the principles of the Declaration of Helsinki.

AVAILABILITY OF DATA AND MATERIALS

The raw match-level dataset analysed in this study is available from the BigTimeStats beach volleyball database on Kaggle (<https://www.kaggle.com/datasets/jessemotipak/beach-volleyball>, accessed January 14, 2026). under the terms specified by the original data providers. All data processing and analysis code, as well as scripts to reproduce the statistical models and figures presented in this article, are openly available in the project's GitHub repository: <https://github.com/nyantakyiappiah-eng/beach-volleyball-prediction>

ACKNOWLEDGMENTS

The authors thank BigTimeStats for making the beach volleyball database publicly available and acknowledge the support of University of Cape Coast for providing academic and technical support during this project

REFERENCES

- Albert, J. (2008). Streaky Hitting in Baseball. *Journal of Quantitative Analysis in Sports*, 4(1). <https://doi.org/10.2202/1559-0410.1085>
- Almulla, J., & Alam, T. (2020). Machine Learning Models Reveal Key Performance Metrics of Football Players to Win Matches in Qatar Stars League. *IEEE Access*, 8, 213695-213705. <https://doi.org/10.1109/ACCESS.2020.3038601>

- Alves, H., Voss, M.W., Boot, W.R., Deslandes, A., Cossich, V., Salles, J.I. and Kramer, A.F. (2013). Perceptual-Cognitive Expertise in Elite Volleyball Players. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00036>
- Beach Volleyball. (n.d.). [Wwww.kaggle.com](http://www.kaggle.com). Retrieved from [Accessed 2026, 11 March]: <https://www.kaggle.com/datasets/jessemostipak/beach-volleyball>
- Bisagno, E., & Morra, S. (2018). How do we learn to "kill" in volleyball?: The role of working memory capacity and expertise in volleyball motor learning. *Journal of Experimental Child Psychology*, 167, 128-145. <https://doi.org/10.1016/j.jecp.2017.10.008>
- Buková, A., Zusková, K., Szerdiová, L., & Küchelová, Z. (2017). Demographic factors and physical activity of female undergraduates. *Physical Activity Review*, 5, 202-211. <https://doi.org/10.16926/par.2017.05.25>
- Cañal-Bruland, R. (2011). Differentiating Experts' Anticipatory Skills in Beach Volleyball. *Research Quarterly for Exercise and Sport*, 82(4). <https://doi.org/10.5641/027013611X13275192111745>
- Choi, K. H., & Byun, J. (2024). Professionalization of action sports: field-and organizational-level professionalization of new Olympic sports. *Sport in Society*, 1-24. <https://doi.org/10.1080/17430437.2024.2325970>
- Chun, Y.-J., & Shin, H.-J. (2020). Defensive Tactical Analysis of Back Attack in Men's Professional Volleyball Match. *Korean Journal of Sports Science*, 29(1), 773-781. <https://doi.org/10.35159/kjss.2020.02.29.1.773>
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Ben Van Calster, Ghassemi, M., Liu, X., Reitsma, J. B., Maarten van Smeden, Anne-Laure Boulesteix, Jennifer Catherine Camaradou, Leo Anthony Celi, Spiros Denaxas, Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., & Hoffman, M. M. (2024). TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, e078378-e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Csató, L. (2023). A paradox of tournament seeding. *International Journal of Sports Science & Coaching*, 18(4), 1277-1284. <https://doi.org/10.1177/17479541221141617>
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845. <https://doi.org/10.2307/2531595>
- Giatsis, G., Lola, A., Hatzimanouil, D., & Tzetzis, G. (2023). Evaluation of a beach volleyball skill instrument for the line shot attack. *Journal of Physical Education*, 34(1). <https://doi.org/10.4025/jphyseduc.v34i1.3409>
- Grgantov, Z., Jelaska, I., & Šuker, D. (2018). Intra and Interzone Differences of Attack and Counterattack Efficiency in Elite Male Volleyball. *Journal of Human Kinetics*, 65(1), 205-212. <https://doi.org/10.2478/hukin-2018-0028>
- Iba, K., Shinozaki, T., Maruo, K., & Noma, H. (2021). Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. *BMC Medical Research Methodology*, 21(1). <https://doi.org/10.1186/s12874-020-01201-w>
- Kheddoum, A., Hadji, M. L., & Khaled, M. (2025). Creating An App To Measure Decision-Making In Volleyball. 403. <https://doi.org/10.37139/1988-017-001-032>
- Kianifard, F., & Vach, W. (1995). Logistic Regression with Missing Values in the Covariates. *Technometrics*, 37(4), 460. <https://doi.org/10.2307/1269744>
- Knoblochova, M., Mudrak, J., & Slepicka, P. (2021). Achievement goal orientations, sport motivation and competitive performance in beach volleyball players. *Acta Gymnica*. <https://doi.org/10.5507/aq.2021.016>

- Kostyukov, V., & Dashaev, K. (2022). Block-modular program of pre-competitive training of athletes of mass categories in beach volleyball. *Fizicheskaya Kul'tura, Sport - Nauka I Praktika*, 1, 52-57. https://doi.org/10.53742/1999-6799/1_2022_52
- Künzeli, S., Schweikart, F., Köhn, D., & Schläppi-Lienhard, O. (2014). Effectiveness of the Call in Beach Volleyball Attacking Play. *Journal of Human Kinetics*, 44(1), 183-191. <https://doi.org/10.2478/hukin-2014-0124>
- Liu, W. B., Wu, Y. J., Li, B., & Deng, L. (2012). Model Analysis of Analysis and Evaluation about NBA Match. *Advanced Materials Research*, 468, 1516-1520. <https://doi.org/10.4028/www.scientific.net/AMR.468-471.1516>
- McHale, I. G., Scarf, P. A., & Folker, D. E. (2012). On the Development of a Soccer Player Performance Rating System for the English Premier League. *Interfaces*, 42(4), 339-351. <https://doi.org/10.1287/inte.1110.0589>
- Nattino, G., Pennell, M. L., & Lemeshow, S. (2020). Rejoinder to "Assessing the goodness of fit of logistic regression models in large samples: A modification of the Hosmer-Lemeshow test." *Biometrics*. <https://doi.org/10.1111/biom.13250>
- Pencina, M. J., D'Agostino, R. B., D'Agostino, R. B., & Vasan, R. S. (2007). Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statistics in Medicine*, 27(2), 157-172. <https://doi.org/10.1002/sim.2929>
- Phillips, J., B. Caudill, S., & Mixon, F. G. (2015). Tournament Seeding Efficiency and Home Court Advantage: College Basketball's National Invitation Tournament. *International Journal of Statistics and Probability*, 4(3). <https://doi.org/10.5539/ijsp.v4n3p101>
- Predoiu, R. (2023). Resilience, Risk-Taking Behavior and Aggression among Female Volleyball Players - a Preliminary Study. *Journal of Psychological Science and Research*, 3(2). <https://doi.org/10.53902/JPSSR.2023.03.000542>
- Saif, Mohd., Khan, S., & Abraham, B. (2025). Basic Defensive Stance in Volleyball. *International Journal for Multidisciplinary Research*, 7(5). <https://doi.org/10.36948/ijfmr.2025.v07i05.55848>
- Schelling, X., & Robertson, S. (2020). A development framework for decision support systems in high-performance sport. *International Journal of Computer Science in Sport*, 19(1), 1-23. <https://doi.org/10.2478/ijcss-2020-0001>
- Schelling, X., Fernández, J., Ward, P., Fernández, J., & Robertson, S. (2021). Decision Support System Applications for Scheduling in Professional Team Sport. The Team's Perspective. *Frontiers in Sports and Active Living*, 3. <https://doi.org/10.3389/fspor.2021.678489>
- Seweryniak, T., Mroczek, D., & Łukasik, Ł. (2013). Analysis and Evaluation of Defensive Team Strategies in Women's Beach Volleyball - An Efficiency-Based Approach. *Human Movement*, 14(1). <https://doi.org/10.2478/v10038-012-0047-9>
- Solomon, C., & van Collier-Peter, S. (2019). How coaching aligns the psychological contract between the young millennial professional and the organisation. *SA Journal of Human Resource Management*, 17. <https://doi.org/10.4102/sajhrm.v17i0.1146>
- Tili, M., & Giatsis, G. (2011). The height of the men's winners FIVB Beach Volleyball in relation to specialization and court dimensions. *Journal of Human Sport and Exercise*, 6(3), 504-510. <https://doi.org/10.4100/jhse.2011.63.04>
- Umarov, K. M. (2024). Effectiveness of Developing the Technique of Attacking Movements of Young Volleyball Players. *Pubmedia Jurnal Pendidikan Olahraga*, 1(3). <https://doi.org/10.47134/jpo.v1i3.361>
- Yusni, Y., Meutia, F., & Taufik, N. H. (2025). Advantageous Correlation Between Anthropometry and Physical Fitness in Amateur Soccer Players. *Retos*, 67, 985-995. <https://doi.org/10.47197/retos.v67.112998>

- Zhao, H. (2018). Sports Situation-Based Neural Mechanism of High-level Volleyball Players' Decision-making Behavior. *NeuroQuantology*, 16(6). <https://doi.org/10.14704/nq.2018.16.6.1602>
- Zhou, S., & Liu, H. (2025). Injury risk analysis of movement restriction and body asymmetry in sports injury prediction. *BMC Sports Science, Medicine and Rehabilitation*, 18(1). <https://doi.org/10.1186/s13102-025-01465-z>



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

Effects of stepwise changes in subjective effort on joint kinematics during the table tennis forehand topspin stroke

 **Yuki Sato**  . Department of General Education. National Institute of Technology. Oyama College. Oyama, Japan.
 **Kengo Wakui**. Division of Humanities and Social Sciences. Niigata University. Niigata, Japan.
 **Koichiro Miyazaki**. Niigata University. Niigata, Japan.
Yukihiko Ushiyama. Division of Humanities and Social Sciences. Niigata University. Niigata, Japan.

ABSTRACT

In table tennis forehand topspin strokes, players must regulate output according to rally demands; however, how joint kinematics change over time with subjective effort remains unclear. This study examined phase-dependent changes in pelvic, trunk, and upper-limb joint angles and racket orientation when subjective effort was manipulated in a graded manner. Ten skilled male players performed strokes at 60%, 70%, 80%, 90%, and 100% effort, recorded with motion capture. One-way repeated-measures ANOVA tested effects of effort on peak racket-tip speed, inter-event durations, and joint/racket angles at impact; Statistical Parametric Mapping (SPM) evaluated time-series waveforms. Peak racket-tip speed increased stepwise with effort (significant main effect), whereas inter-event durations showed no significant differences. SPM showed frequent effort-dependent differences in shoulder horizontal flexion/extension and pelvic axial rotation, primarily around impact. At impact, significant main effects were observed mainly in the pelvis, trunk, and proximal upper-limb joints, but not in the forearm, wrist, or racket orientation. During follow-through, effort-related differences emerged in thorax lateral flexion, shoulder internal/external rotation, and racket tilt angle. These findings suggest that increased output with higher effort is achieved mainly through proximal kinematic adjustments and late-phase changes while maintaining distal mechanics and racket configuration at impact. **Keywords:** Performance analysis, Motor control, Motion analysis, Output modulation, Time-series analysis, Racket sports.

Cite this article as:

Sato, Y., Wakui, K., Miyazaki, K., & Ushiyama, Y. (2026). Effects of stepwise changes in subjective effort on joint kinematics during the table tennis forehand topspin stroke. *Scientific Journal of Sport and Performance*, 5(3), 447-460. <https://doi.org/10.55860/YRCM7402>



Corresponding author. Department of General Education. National Institute of Technology. Oyama College. Oyama, Japan.

E-mail: styk7006@gmail.com

Submitted for publication February 06, 2026.

Accepted for publication March 16, 2026.

Published March 19, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/YRCM7402>

INTRODUCTION

Table tennis requires players to execute accurate strokes in rapid succession, with extremely short reaction times. Therefore, severe temporal constraints constitute a fundamental feature of performance. During rallies, limiting an opponent's available reaction time confers a tactical advantage; thus, increasing the ball speed is a fundamental approach (Lees, 2003). However, topspin strokes must also be executed under the spatial constraint of clearing the net, and kinematic analyses have shown that they simultaneously produce a high racket speed and spin rate (Bańkosz and Winiarski, 2017). The forehand topspin stroke is frequently used as a primary offensive technique (Seemiller and Holowchak, 1997) and generates a high mechanical output through coordinated whole-body movement that sequentially links the trunk, upper limb, and racket. Simultaneously, ball outcomes are influenced by the racket-face orientation at impact (Iino and Kojima, 2009). Accordingly, the forehand topspin stroke constitutes a motor task that requires simultaneous power modulation and precision control, making it well suited for investigation from both kinematic and motor-control perspectives.

The kinematic characteristics of the table tennis forehand topspin stroke have been reported, including joint motions and racket movement patterns during the stroke (Iino and Kojima, 2009; Iino and Kojima, 2011; Kidokoro et al., 2019; Bańkosz and Winiarski, 2025). In high-speed actions, such as striking, throwing, and swinging, velocity generation is typically explained by a proximal-to-distal sequence, in which motion is transferred in a linked manner from proximal to distal segments (Putnam, 1993). Similarly, in table tennis, the rotational motion of the lower trunk has been shown to play an important role in racket-speed development during the forehand topspin stroke. Comparisons across skill levels further indicate that more advanced players exhibit a greater contribution of lower trunk rotation and distinct kinematic features during the racket-acceleration phase (Iino and Kojima, 2009). Using inverse-dynamics approaches, previous studies have also identified the contributions of shoulder internal rotation torque and mechanical energy transfer from the trunk to the upper arm to racket-speed generation (Iino and Kojima, 2011). Moreover, investigations of trunk-upper-limb coordination and muscle activity during table tennis strokes suggest that trunk motion contributes to force production across the stroke (Kondrič et al., 2006; Tsai et al., 2010; Le Mansec et al., 2018).

In competitive table tennis, strokes are not always executed with maximal effort. Instead, players regulate ball speed and spin according to the rally context and tactical intent. This context-dependent output regulation directly supports the consistency of performance and tactical versatility. From a motor control perspective, research has suggested that subjective effort (that is, the effort perceived during movement) may contribute to movement generation, along with externally observable kinematic adjustments. Subjective effort is linked to the central motor command and has been shown to change in parallel with variations in force output and regulation of movement intensity (de Morree et al., 2012). In skilled performance, such variations may manifest as shifts in the movement patterns and coordination (Hashimoto et al., 2021). Consistent with these findings, increases in subjective effort in table tennis coincide with increased ball speed and spin (Sato et al., 2025). Taken together, clarifying how stroke movements are adjusted as subjective effort changes can improve the grasp of output regulation in table tennis and inform approaches to skill acquisition and coaching.

Proximal-to-distal sequencing is widely recognized as a fundamental mechanism for generating end-point velocity in striking and throwing movements (Putnam, 1993; Elliott et al., 1995; Hirashima et al., 2002). In table tennis, three-dimensional kinematic and kinetic analyses have shown that trunk rotation and upper limb motions contribute to racket-speed generation (Iino and Kojima, 2008; Iino et al., 2009; Iino and Kojima, 2011). Despite this body of work, most studies have focused on maximal-effort conditions, skill-level differences, or comparisons across stroke conditions, and investigations addressing output regulation in

relation to subjective effort are scarce. Studies focusing on subjective effort in table tennis have reported increases in ball speed and spin as the effort increases (Sato et al., 2025). However, it is unclear how stroke movements are modified by changing subjective effort, particularly which joints and movement phases exhibit effort-dependent differences when joint kinematics and racket angle are examined as time series.

Accordingly, this study aimed to characterize how three-dimensional joint angles of the pelvis, trunk, and upper-limb joints, together with racket orientation, vary across phases of the table tennis forehand topspin stroke when subjective effort is manipulated at graded levels. We further examined how output regulation accompanying changes in subjective effort is achieved during the stroke by identifying the joints and phases that exhibit effort-dependent differences, with particular attention to changes in stroke duration and adjustment in joint motion. By delineating the correspondence between subjective effort and kinematic alterations, this study seeks to advance the understanding of output regulation in table tennis strokes and provide insights relevant to skill acquisition and coaching.

MATERIAL AND METHODS

Participants

Ten male table tennis players participated in this study. All participants were right-handed. The mean age, height, and body mass of the participants were 20.6 ± 1.4 years, 168.7 ± 6.3 cm, and 60.7 ± 6.8 kg, respectively. This study was approved by the Ethics Committee of Niigata University (approval number: 2023-2036). Prior to participation, all participants were informed of the purpose and procedures of the experiment, and written informed consent was obtained.

Experimental task

After completing a sufficient self-directed warm-up, the participants performed the experimental trials. The task consisted of forehand topspin strokes, during which the participants were instructed to return balls delivered to the forehand side to the cross-court area (diagonally) of the opponent's court. Participants were instructed as follows: "Assuming that a maximal-effort stroke corresponds to 100%, please perform forehand topspin strokes to the cross-court area at X% effort." The subjective effort was set at five levels: 60%, 70%, 80%, 90%, and 100%. The order of the effort conditions was fixed at 80%, 60%, 90%, 70%, and 100%. This order was selected to avoid potential anchoring effects associated with the 100% effort condition and reduce reference bias arising from consecutive trials at adjacent effort levels (Sato et al., 2025). For each effort condition, five trials that participants judged to match the instructed effort level were included in the analysis. Ball delivery was performed by a single experienced table tennis player who held a Japan Table Tennis Association-certified coaching qualification and was left-handed. The participants used their own rackets, which they routinely employed during training and competition. All rackets were offensive-type rackets equipped with high-tension inverted rubber on both sides (sponge thickness: extra thick). The mean racket mass was 185.7 ± 4.6 g.

Measurements

Body and racket motions were recorded using a motion capture system comprising five infrared cameras (ProReflex MCU240, Qualisys, Sweden) operating at a sampling frequency of 240 Hz. The cameras were arranged around the participant to ensure full coverage of the movement while avoiding interference with the swing motion. To record the instant of ball-racket impact, a digital camera (EX-ZR70; CASIO, Japan) was positioned laterally to the participant and operated at 240 frames per second. To capture upper-body and racket kinematics, spherical retroreflective markers (20 mm in diameter; hereafter referred to as markers) were attached to the anatomical landmarks and racket (Figure 1). Markers were placed on the following

anatomical landmarks: right acromion (AC_R), left acromion (AC_L), suprasternal notch (SN), xiphoid process (PX), seventh cervical spinous process (C7), eighth thoracic spinous process (T8), midpoint between the left and right posterior superior iliac spines (PSIS_M), right anterior superior iliac spine (ASIS_R), left anterior superior iliac spine (ASIS_L), right lateral epicondyle of the humerus (EL), right medial epicondyle of the humerus (EM), right radial styloid process (RS), right ulnar styloid process (US), and dorsal aspect of the distal end of the right third metacarpal (MC3). Three markers were attached to the racket: tip (RT), upper region (RU), and lower region (RL). Hereafter, anatomical landmarks and racket markers are referred to using the abbreviations defined earlier. Due to experimental constraints, markers were not placed on the participants' lower limbs. Therefore, all analyses in this study were limited to the upper-body and racket kinematics.

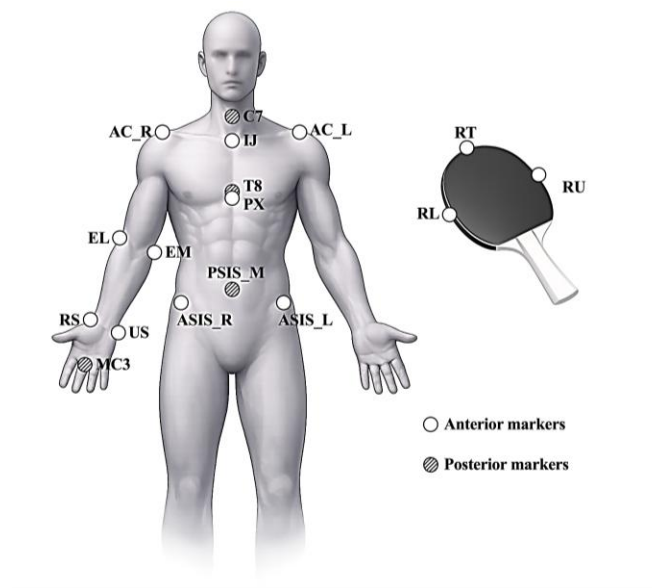


Figure 1. Marker placement on the body and racket.

Data analysis

A global coordinate system ($\Sigma_G \equiv O_G - X_G Y_G Z_G$) was defined with its origin in the near right corner of the table on the participant's side (from the participant's perspective). The positive Z_G -axis was defined as the vertical upward direction, the positive X_G -axis as the ball flight direction (forward from the participant's perspective), and the positive Y_G -axis as the direction obtained from the cross product of the Z_G - and X_G -axes (leftward from the participant's perspective), forming a right-handed coordinate system. The recorded coordinate data were smoothed separately for each axis using a fourth-order, zero-lag (bidirectional) Butterworth low-pass digital filter. The optimal cutoff frequency was determined by residual analysis and ranged from 6–19 Hz. The racket motion during the stroke was calculated from the coordinates of the racket-tip marker, and the stroke interval was defined based on this motion. The instant of ball–racket impact was set to 0 s, and the analysis window was defined from –0.6 to +0.3 s. Within this interval, three key events were identified: T1, defined as the instant when the x-coordinate of the racket tip reached its minimum value (maximum backswing); T2, defined as the instant of ball contact determined using high-speed video footage (impact); and T3, defined as the instant after impact when the racket-tip velocity first reached a local minimum (end of follow-through). The racket-tip velocity was computed as the time derivative of the filtered racket-tip position. To standardize the temporal scale of the stroke, the intervals between T1–T2 and T2–T3 were time-normalized separately for each trial, such that the impact event (T2) was aligned at the same normalized time point across trials.

Time normalization was performed by resampling the data such that T1, T2, and T3 corresponded to 0%, 34%, and 100% of the normalized time scale, respectively. On average, across all trials, T2 occurred at approximately 34% of the stroke cycle. In the present study, continuous motion from T1 to T3 was treated as a single stroke (Figure 2).

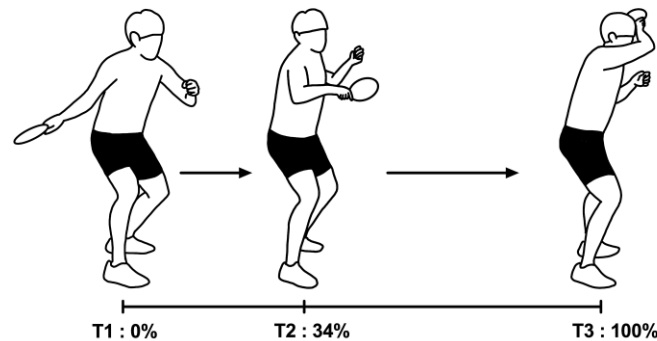


Figure 2. Definition of the stroke interval and key events.

To compute segment kinematics, local coordinate systems for the pelvis, thorax, upper arm, forearm, and hand (Σ_P , Σ_T , Σ_U , Σ_F , and Σ_H) were defined from the smoothed three-dimensional marker coordinates, with reference to the recommendations of the International Society of Biomechanics (Wu et al., 2002; Wu et al., 2005). The pelvis coordinate system ($\Sigma_P \equiv O_P - X_P Y_P Z_P$) was defined with the origin O_P as the midpoint of the line segment connecting the midpoint between ASIS_R and ASIS_L and PSIS_M; the Z_P -axis as the unit vector from ASIS_L to ASIS_R (rightward); the X_P -axis as the unit vector lying in the plane formed by ASIS_R, ASIS_L, and PSIS_M, orthogonal to both the Z_P -axis and the vector from the ASIS midpoint to PSIS_M, with the positive direction anterior; and the Y_P -axis as the cross product of the Z_P - and X_P -axes ($Y_P = Z_P \times X_P$), directed superiorly. The thorax coordinate system ($\Sigma_T \equiv O_T - X_T Y_T Z_T$) was defined with the origin O_T at the SN; the Y_T -axis as the superiorly directed unit vector from the midpoint between PX and T8 to the midpoint between SN and C7; the Z_T -axis as the cross product of the Y_T -axis and the unit vector from SN to C7, directed rightward; and the X_T -axis as the cross product of the Y_T - and Z_T -axes ($X_T = Y_T \times Z_T$), directed anteriorly. The upper arm coordinate system ($\Sigma_U \equiv O_U - X_U Y_U Z_U$) was defined with the origin O_U at the glenohumeral joint centre (GH), where GH was estimated by translating from AC_R in a direction perpendicular to the line connecting AC_R and AC_L by 1.87% of body height inferiorly and 0.645% of body height anteriorly (Ehara et al., 1998); the Y_U -axis as the unit vector from the elbow joint centre (EC; midpoint between EM and EL) to GH (proximal direction); the X_U -axis as the cross product of the Y_U -axis and the vector from EL to EM, directed anteriorly; and the Z_U -axis as the cross product of the X_U - and Y_U -axes ($Z_U = X_U \times Y_U$), directed laterally for the right upper arm. The forearm coordinate system ($\Sigma_F \equiv O_F - X_F Y_F Z_F$) was defined with the origin O_F at EC; the Y_F -axis as the unit vector from the wrist joint centre (WC; midpoint between RS and US) to EC (proximal direction defined as positive); the Z_F -axis as the cross product of the vector from RS to US and the Y_F -axis, directed radially; and the X_F -axis as the cross product of the Y_F - and Z_F -axes ($X_F = Y_F \times Z_F$), directed anteriorly (palmar direction). The hand coordinate system ($\Sigma_H \equiv O_H - X_H Y_H Z_H$) was defined with the origin O_H at WC; the Y_H -axis as the unit vector from MC3 to WC (proximal direction toward the forearm); the Z_H -axis as the cross product of the forearm Z_F -axis and the Y_H -axis, directed radially (lateral); and the X_H -axis as the cross product of the Y_H - and Z_H -axes ($X_H = Y_H \times Z_H$), directed anteriorly (palmar direction).

Pelvic angles were computed as the relative motion of the pelvis coordinate system (Σ_P) with respect to the global coordinate system (Σ_G) using a Z-X'-Y" Cardan rotation sequence. Rotation about the Z-axis

represented pelvic anterior/posterior tilt [extension (+)/flexion (-)], rotation about the X'-axis represented pelvic lateral tilt [left lateral flexion (-)/right lateral flexion (+)], and rotation about the Y"-axis represented pelvic axial rotation [left rotation (+)/right rotation (-)]. Thorax angles were computed as the relative motion of the thorax coordinate system (Σ_T) with respect to the pelvis coordinate system (Σ_P) using the same Z-X'-Y" Cardan sequence. Rotation about the Z-axis represented thorax flexion/extension [extension (+)/flexion (-)], rotation about the X'-axis represented thorax lateral flexion [left (-)/right (+)], and rotation about the Y"-axis represented thorax axial rotation [left (+)/right (-)]. Shoulder joint angles were computed as the relative motion of the upper arm coordinate system (Σ_U) with respect to the thorax coordinate system (Σ_T) using a Y-X'-Y" Cardan sequence. Rotation about the Y-axis represented shoulder horizontal flexion (+)/horizontal extension (-), rotation about the X'-axis represented shoulder adduction (+)/abduction (-), and rotation about the Y"-axis represented shoulder internal rotation (+)/external rotation (-). The elbow joint angles were computed as the relative motion of the forearm coordinate system (Σ_F) with respect to the upper arm coordinate system (Σ_U) using a Z-X'-Y" Cardan sequence. Rotation about the Z-axis represented elbow flexion (+)/extension (-). Rotation about the X'-axis represented the carrying angle (valgus/varus) and was not analysed in this study. Rotation about the Y"-axis represented forearm pronation (+)/supination (-). Wrist joint angles were computed as the relative motion of the hand coordinate system (Σ_H) with respect to the forearm coordinate system (Σ_F) using a Z-X'-Y" Cardan sequence. Rotation about the Z-axis represented wrist palmar flexion (+)/dorsiflexion (-), and rotation about the X'-axis represented wrist radial deviation (-)/ulnar deviation (+). Rotation about the Y"-axis was always zero due to constraints of the analysis model and was therefore not analysed. Racket orientation was computed by defining the racket plane using the three racket markers (RT, RU, and RL) and evaluating the direction of the plane normal vector with respect to the global coordinate system. The racket centre was defined as the midpoint between RU and RL. The normal vector was obtained as the cross product of the vector from RU to RL and the vector from the racket centre to RT. Racket orientation was expressed using two components: the horizontal and tilt angles. The horizontal angle was defined as the angle between the projection of the normal vector onto the horizontal plane and the forward axis (X_G), which represents the left-right orientation of the racket face. The tilt angle was defined as the elevation angle of the normal vector relative to the horizontal plane, representing the upward-downward inclination of the racket face. Racket-tip speed was computed by time-differentiating the three-dimensional coordinates of the racket-tip marker (RT).

Statistical analysis

For the statistical analyses, mean values were calculated for each participant and each effort condition by averaging the analysed trials (five trials per condition). All tests were two-tailed, and the significance level was set at $\alpha = .05$. To examine the effects of effort condition on the peak racket-tip speed, the durations between events (T1-T2, T2-T3, and T1-T3), and the joint angles and racket orientation at impact, a one-way repeated-measures analysis of variance (RM-ANOVA) was performed with effort (five levels) as a within-subject factor. When the assumption of sphericity was violated, the Geisser-Greenhouse correction was applied. When a significant main effect was detected, post hoc pairwise comparisons were conducted using Holm's adjustment for multiple comparisons. To identify the time intervals during which the time-series profiles of joint angles and racket orientation differed across effort conditions, Statistical Parametric Mapping (SPM) was applied using a one-way repeated-measures SPM ANOVA (SPM{F}). For the SPM analyses, each participant's waveform for each effort condition was obtained by averaging the five trials for that condition. Critical thresholds were computed based on random field theory (RFT) to account for the smoothness of the time-series data ($\alpha = .05$), and contiguous supra-threshold regions of the SPM{F} trajectory were extracted as significant intervals. For kinematic variables showing a significant main effect in the one-way repeated-measures SPM ANOVA, post hoc comparisons were performed for all pairwise combinations between conditions (10 pairs) using paired SPM t-tests, with Holm's procedure applied to

control for multiple comparisons. All statistical analyses were performed using GraphPad Prism 10, and SPM analyses were conducted in Python using the `spm1d` library.

RESULTS

Peak racket-tip speed increased stepwise with increasing subjective effort conditions (Table 1). The main effect of effort was significant, and post hoc comparisons revealed significant differences between all pairs of effort conditions (all $p < .001$). In contrast, no main effect of effort was observed for the inter-event durations (T1–T2, T2–T3, and T1–T3) (T1–T2: $p = .059$; T2–T3: $p = .149$; T1–T3: $p = .060$).

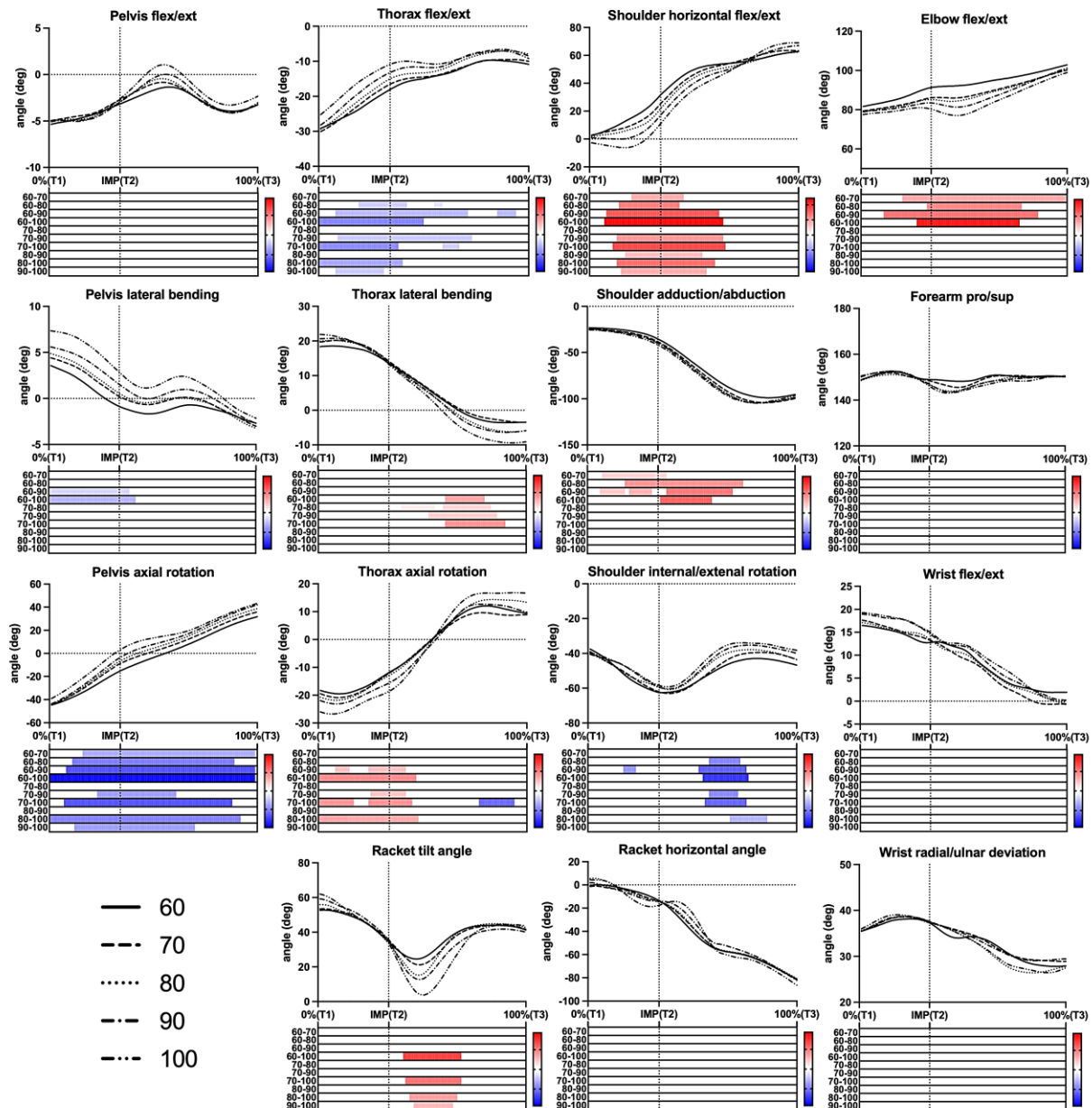


Figure 3. Mean joint-angle waveforms across subjective effort conditions (time-normalized with T1 = 0%, T2 = 34%, and T3 = 100%) and heatmaps indicating significant post hoc SPM results.

Table 1. Peak racket-tip speed and inter-event durations.

	60%	70%	80%	90%	100%	Top: 1 – way repeated-measures ANOVA Bottom: Compact letter display
	Mean $\pm$ SD					
Peak racket-tip speed [m/s]	12.72 $\pm$ 0.90	13.69 $\pm$ 0.78	14.30 $\pm$ 0.84	14.83 $\pm$ 0.69	15.75 $\pm$ 0.69	$p < .001$, F (1.97, 17.72) = 114.3, $\eta^2 = 0.927$ 60:A, 70:B, 80:C, 90:D, 100:E
T1–T2 [s]	-0.11 $\pm$ 0.02	-0.11 $\pm$ 0.02	-0.11 $\pm$ 0.02	-0.11 $\pm$ 0.02	-0.10 $\pm$ 0.02	$p = .059$, F (2.59, 23.32) = 2.98, $\eta^2 = 0.249$ 60:A, 70:A, 80:A, 90:A, 100:A
T2–T3 [s]	0.21 $\pm$ 0.02	0.21 $\pm$ 0.02	0.20 $\pm$ 0.03	0.20 $\pm$ 0.03	0.19 $\pm$ 0.04	$p = .149$, F (1.57, 14.16) = 2.25, $\eta^2 = 0.200$ 60:A, 70:A, 80:A, 90:A, 100:A
T1–T3 [s]	0.32 $\pm$ 0.03	0.32 $\pm$ 0.03	0.31 $\pm$ 0.03	0.31 $\pm$ 0.03	0.30 $\pm$ 0.03	$p = .060$, F (1.66, 14.92) = 3.61, $\eta^2 = 0.286$ 60:A, 70:A, 80:A, 90:A, 100:A

Table 2. Joint angles at impact.

		Joint angle [°]	60%	70%	80%	90%	100%	Top: 1 - way repeated-measures ANOVA
			Mean ± SD				Bottom: Compact letter display	
Pelvis	Extension (+)		-3,2 ± 5.7	-2,7 ± 6.2	-2,9 ± 6.6	-3,0 ± 6.3	-2,6 ± 7.0	$p = .818$, F (1.54, 13.85) = 0.14, $\eta^2 = 0.015$ 60:A, 70:A, 80:A, 90:A, 100:A
	Flexion (-)							
	Left lateral bending (-)		-0,9 ± 2.6	0,2 ± 3.2	0,5 ± 3.0	1,5 ± 3.0	2,9 ± 2.8	$p = .002$, F (1.96, 17.67) = 9.35, $\eta^2 = 0.510$ 60:A, 70:AB, 80:AB, 90:B, 100:B
	Right lateral bending (+)							
	Left rotation (+)		-15,2 ± 10.9	-8,8 ± 7.8	-6,5 ± 11.4	-3,6 ± 10.1	3,1 ± 9.2	$p < .001$, F (2.99, 26.96) = 46.51, $\eta^2 = 0.838$ 60:A, 70:B, 80:BC, 90:C, 100:D
	Right rotation (-)							
Thorax	Extension (+)		-18,1 ± 12.7	-16,5 ± 12.8	-15,0 ± 13.7	-13,2 ± 12.7	-10,9 ± 12.7	$p < .001$, F (1.71, 15.42) = 19.46, $\eta^2 = 0.684$ 60:A, 70:AB, 80:BC, 90:CD, 100:D
	Flexion (-)							
	Left lateral bending (-)		13,6 ± 3.0	14,1 ± 3.2	13,3 ± 2.8	13,7 ± 3.5	13,3 ± 4.3	$p = .522$, F (2.33, 20.99) = 0.71, $\eta^2 = 0.073$ 60:A, 70:A, 80:A, 90:A, 100:A
	Right lateral bending (+)							
	Left rotation (+)		-11,8 ± 4.4	-12,1 ± 5.2	-12,9 ± 4.2	-15,5 ± 4.4	-18,6 ± 4.5	$p < .001$, F (2.50, 22.51) = 24.92, $\eta^2 = 0.735$ 60:A, 70:A, 80:AB, 90:BC, 100: C
	Right rotation (-)							
Shoulder	Horizontal flexion (+)		31,7 ± 12.2	25,4 ± 10.6	21,9 ± 14.8	17,2 ± 14.8	11,1 ± 15.9	$p < .001$, F (2.44, 22.00) = 43.99, $\eta^2 = 0.830$ 60:A, 70:B, 80:B, 90:C, 100:D
	Horizontal extension (-)							
	Adduction (+)		-35,4 ± 10.4	-38,5 ± 10.3	-39,7 ± 12.5	-39,9 ± 9.0	-41,9 ± 10.3	$p = .021$, F (2.00, 18.01) = 4.79, $\eta^2 = 0.347$ 60:A, 70:B, 80:B, 90:AB, 100:AB
	Abduction (-)							
	Internal rotation (+)		-62,5 ± 14.5	-62,2 ± 14.5	-59,6 ± 15.0	-59,3 ± 16.0	-58,4 ± 16.7	$p = .017$, F (2.46, 22.10) = 4.53, $\eta^2 = 0.335$ 60:A, 70:A, 80:A, 90:A, 100:A
	External rotation (-)							
Elbow	Flexion (+)		91,3 ± 10.7	86,1 ± 9.4	85,2 ± 10.8	83,5 ± 10.6	80,3 ± 9.0	$p = .001$, F (1.92, 17.30) = 10.63, $\eta^2 = 0.541$ 60:A, 70:B, 80:B, 90:B, 100:B
	Extension (-)							
Forearm	Pronation (+)		149,0 ± 12.5	148,5 ± 12.6	146,7 ± 12.7	147,1 ± 13.5	146,2 ± 13.2	$p = .218$, F (1.82, 16.40) = 1.68, $\eta^2 = 0.157$ 60:A, 70:A, 80:A, 90:A, 100:A
	Supination (-)							
Wrist	Palmar flexion (+)		12,7 ± 11.6	13,2 ± 11.2	13,3 ± 9.9	15,0 ± 10.9	14,6 ± 9.8	$p = .406$, F (1.61, 14.47) = 0.90, $\eta^2 = 0.091$ 60:A, 70:A, 80:A, 90:A, 100:A
	Dorsiflexion (-)							
	Radial deviation (-)		37,3 ± 6.1	37,6 ± 5.8	37,3 ± 5.8	37,5 ± 6.1	37,5 ± 5.8	$p = .747$, F (2.43, 21.90) = 0.35, $\eta^2 = 0.038$ 60:A, 70:A, 80:A, 90:A, 100:A
	Ulnar deviation (+)							
Racket	Horizontal	Leftward rotation (+)	-14,1 ± 6.8	-16,4 ± 10.9	-17,0 ± 13.9	-15,3 ± 12.2	-11,0 ± 18.5	$p = .549$, F (1.41, 12.71) = 0.51, $\eta^2 = 0.054$ 60:A, 70:A, 80:A, 90:A, 100:A
		Rightward rotation (-)						
	Tilt	Upward (+)	34,6 ± 3.2	35,2 ± 4.9	33,5 ± 1.8	36,0 ± 4.6	34,6 ± 2.0	$p = .280$, F (1.49, 13.36) = 1.37, $\eta^2 = 0.132$ 60:A, 70:A, 80:A, 90:A, 100:A
		Downward (-)						

Figure 3 shows the mean time-normalized waveforms of the joint angles for each subjective effort condition ($T1 = 0\%$, $T2 = 34\%$, $T3 = 100\%$), together with a heatmap summarizing the condition pairs and time intervals that exhibited significant differences in the post hoc analyses ($p < .05$). In the heatmap, the colour indicates the mean difference within significant intervals ($A - B$), with red denoting $A > B$ and blue denoting $A < B$; darker colours indicate larger differences within the corresponding intervals. Post hoc comparisons revealed that the largest number of significant condition pairs occurred for shoulder horizontal flexion/extension and pelvic axial rotation. For shoulder horizontal flexion/extension, significant differences were observed for nine pairs, excluding the comparison between 70% and 80%, and these differences were primarily distributed around the impact phase. For pelvic axial rotation, significant differences were observed in eight pairs, with differences broadly distributed from pre-impact to $T3$. Multiple significant condition pairs were also observed for thorax flexion/extension and shoulder internal/external rotation; differences in thorax flexion/extension were concentrated around the impact, whereas differences in shoulder internal/external rotation were mainly observed after impact. In contrast, several variables showed significant differences for only a limited number of pairs. Pelvic lateral flexion differed significantly between 60% and 90–100% efforts. For elbow flexion/extension, significant differences were observed between 60% and the other conditions (70–100%). Thorax lateral flexion showed differences after impact, whereas thorax axial rotation differed primarily from pre-impact to around the time of impact. For the racket tilt angle, significant differences were observed after impact between 100% and the other conditions (60–90%).

At impact, significant main effects of subjective effort were observed for pelvic lateral flexion and axial rotation, thorax flexion/extension and axial rotation, shoulder horizontal flexion/extension and abduction/adduction, and elbow flexion/extension (Table 2). Post hoc comparisons showed that pelvic lateral flexion increased toward right lateral flexion with increasing effort; the 60% condition was smaller than the high-effort conditions (90–100%) ($p = .002$). Pelvic axial rotation increased toward left rotation with increasing effort; the 60% condition was smaller than the other conditions (70–100%) (all $p < .001$). Thorax flexion/extension shifted toward a less flexed posture with increasing effort; the 60% condition was smaller than the moderate-to-high effort conditions (80–100%) (all $p < .001$). Thorax axial rotation shifted toward right rotation with increasing effort; the low-to-moderate effort conditions (60–70%) were greater than the high-effort conditions (90–100%) (all $p < .001$). Both shoulder horizontal flexion/extension and elbow flexion/extension decreased with increasing effort, and the low-effort conditions were greater than the high-effort conditions ($p < .001$ and $p = .001$, respectively). For shoulder abduction/adduction, the 60% condition was greater than the 70–80% conditions ($p = .021$). Although a significant main effect of effort was observed for shoulder internal/external rotation, no significant between-condition differences were detected in post hoc comparisons at impact. No significant effects of effort were found for the remaining angles (pelvic flexion/extension, thorax lateral flexion, forearm pronation/supination, wrist angles, or racket orientation).

DISCUSSION

This study aimed to determine how the three-dimensional joint angles of the pelvis, trunk, and upper limb joints, as well as racket orientation, change when subjective effort is manipulated at graded levels during the table tennis forehand topspin stroke. The peak racket-tip speed increased stepwise as the subjective effort increased. The inter-event timing within the stroke did not vary with effort. Time-series analyses showed that the effort-related differences were not uniform across the upper body. These differences were most pronounced in pelvic axial rotation and shoulder horizontal flexion/extension. At impact, effort-dependent differences were evident mainly in the pelvis, trunk, and proximal upper limb joints. No significant differences were observed in the forearm, wrist, or racket orientation. Together, these results indicate that a higher racket

speed with increasing effort is achieved without changes in stroke timing. This increase was primarily expressed through kinematic adjustments in the thorax and proximal upper limb.

The racket-tip speed increased across the effort conditions in this study. The inter-event timing during the stroke remained unchanged. Subjective effort indexes the perceived level of effort required during movement. Subjective effort has been linked to the central motor command (de Morree et al., 2012). Thus, the effort-related increase in output may have been achieved by increasing the kinematic magnitudes within the same temporal window. Proximal segment motion is known to influence distal end-point velocity generation (Putnam, 1993). Effort-dependent differences were most pronounced in pelvic axial rotation and shoulder horizontal flexion/extension. This pattern suggests that effort regulation was implemented without altering the temporal structure of the movement. This also suggests the modulation of the proximal segment motion magnitude while preserving the overall movement pattern. In the forehand topspin, higher-level players showed greater contributions of lower trunk rotation and distinct racket acceleration properties. These findings indicate that trunk motion contributes to racket speed generation (Iino and Kojima, 2009). The present results are consistent with this view. Speed regulation has been observed to cause changes in proximal segment kinematics. Lumbar and pelvic kinematics can change over time across stroke direction conditions (He et al., 2023). This evidence is consistent with pelvic axial rotation serving as an adjustable component in response to situational demands. In contrast, studies focusing on stroke duration differences report that fast and slow strokes differ in the stroke time itself. They also reported concomitant changes in joint kinematics and ball effects (Tian and Xiao, 2024). Here, the racket-tip speed increased stepwise despite no differences in inter-event timing. This finding indicates output regulation without modifying the temporal structure. This also indicates that regulation was expressed mainly as proximal-segment kinematic adjustments. The SPM results further localized effort-dependent differences around the impact. Pelvic axial rotation and shoulder horizontal flexion/extension also changed stepwise with increasing impact effort. This pattern suggests the modulation of proximal motion with effort. This particularly implicates the acceleration phase into the impact.

Time-series analyses also showed that effort-related adjustments were not restricted to the velocity generation phase around the impact. The SPM results in Figure 3 identified significant between-condition differences, mainly after impact (after T2) during follow-through. These effects were observed in thorax lateral flexion, shoulder internal/external rotation, and racket tilt angle. Significant intervals tended to occur in comparisons that included the 100% condition. This post-impact clustering suggests a downstream consequence of greater output with increased effort. The post-impact swing likely increased. Therefore, the follow-through movement pattern may have shifted. Accordingly, the follow-through differences are better attributed to post-contact dynamics than to the kinematics at impact itself. This may reflect greater residual momentum after ball contact. They may also reflect differences in the convergence of movement after impact. Follow-through has been reported to indirectly reflect overall stroke stability. It has also been reported to facilitate a smooth transition to subsequent actions (Lees, 2003; Kondrič et al., 2006). Thus, the effort-dependent differences observed during follow-through can be interpreted as movement elaboration resulting from a greater output under high-effort conditions. These can also be interpreted as differences in the kinematic characteristics during the latter part of the stroke.

Effort-dependent adjustments in the present study emerged mainly in the trunk and proximal upper limbs. They also appeared during the follow-through. In contrast, no significant differences were detected in forearm, wrist, or racket orientation at impact. This pattern may reflect the task constraints specific to table tennis strokes. Large changes in distal joint kinematics or racket orientation at impact are likely to impair the accuracy of the stroke. Thus, even when the effort varies, players may preserve the impact configuration.

Speed regulation can then be implemented through adjustments at other joints. Prior kinematic work on the forehand topspin has shown that the contribution of lower trunk rotation varies across performance levels and stroke conditions. Differences have also been shown in the characteristics of the racket acceleration phase (Iino and Kojima, 2009). In addition, shoulder internal rotation torque and mechanical energy transfer from the trunk to the upper arm contribute to racket-speed generation (Iino and Kojima, 2011). Skilled players in high-speed interceptive tasks can also maintain stable outcomes through the use of online information and compensatory adjustments (Bootsma and van Wieringen, 1990; Bootsma et al., 2010). The present combination of preserved distal/racket configuration at impact and proximal kinematic change is consistent with the control framework. Together, the effort-related modulation of speed generation via proximal segment motion and the follow-through changes in thorax lateral flexion, shoulder internal/external rotation, and racket tilt angle may act in a complementary manner. These effects may reflect a regulation strategy that varies racket speed while maintaining the distal mechanics and racket configuration at impact.

This study had several limitations. We did not directly quantify lower limb motion or muscle activity. Thus, the effort-related increase in output cannot be linked to specific activation patterns or mechanical loads. The order of the effort conditions was fixed. Therefore, the order effects may have contributed to the observed effort effects. Our analyses were limited to upper-body and racket kinematics. Accordingly, we cannot infer the contribution of whole-body kinetic chain mechanisms, including lower limb actions and ground reaction forces. Despite these constraints, the data suggest a characteristic regulation pattern. Distal mechanics and racket orientation at impact showed no significant differences across effort levels. Speed regulation instead appeared to concentrate in proximal-segment motion and the latter phase of the stroke. These findings refine understanding of movement regulation strategies in table tennis strokes. They also carry practical implications for coaching. Proximal-segment motion emerges as a key target when training graded regulation of racket speed. Subjective effort offers a simple way to scale output without specialized equipment. It may support practice tasks that vary speed while preserving impact technique. One example is a drill that increases subjective effort from 60% to 100% while maintaining the same stroke mechanics and target placement. Such a drill could promote speed regulation through trunk rotation and shoulder-related motion. It could also limit excessive changes in distal mechanics and racket orientation.

CONCLUSIONS

This study quantified time-series changes in pelvis, thorax, and upper-limb joint angles, as well as racket orientation, during the table tennis forehand topspin stroke across graded levels of subjective effort. Peak racket-tip speed increased stepwise with effort. It differed significantly across all effort conditions. In contrast, inter-event durations (T1–T2, T2–T3, and T1–T3) showed no significant differences. Time-series analyses showed that effort-related differences were not uniform across the upper body. They were most pronounced in shoulder horizontal flexion/extension and pelvic axial rotation. These differences clustered mainly around the impact phase. At impact, effort-dependent differences emerged primarily in the pelvis, thorax, and proximal upper-limb joints angles. No significant differences were detected in the forearm, wrist, or racket orientation. During follow-through, effort-related differences appeared in thorax lateral flexion, shoulder internal/external rotation, and racket tilt angle. This pattern suggests distinct movement elaboration under high-effort conditions. Overall, skilled players may increase output with higher effort without altering the temporal structure of the stroke. This may be achieved mainly through kinematic adjustments in the thorax and proximal upper limb. It may also be achieved through changes in movement characteristics during the latter phase of the stroke. Distal mechanics and racket configuration at impact may remain relatively stable.

AUTHOR CONTRIBUTIONS

YS (Corresponding Author) contributed to Conceptualization, Methodology, Investigation, Formal analysis, and Writing – Original Draft. He also contributed to Writing – Review & Editing and coordinated the overall research process. KW contributed to Investigation, provided Methodology advice, assisted in Validation, and participated in Writing – Review & Editing. KM contributed to Investigation and participated in Writing – Review & Editing. YU contributed to Supervision, Validation, and Writing – Review & Editing. All authors reviewed and approved the final version of the manuscript.

SUPPORTING AGENCIES

No funding agencies were reported by the authors.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

ACKNOWLEDGEMENTS

The authors would like to thank all participants for their cooperation throughout the experiment. This study was conducted in accordance with the current laws of Japan.

REFERENCES

- Bańkosz, Z., & Winiarski, S. (2017). The kinematics of table tennis racquet: differences between topspin strokes. *J. Sports. Med. Phys. Fit.*, 57, 202-213. <https://doi.org/10.23736/S0022-4707.16.06104-1>
- Bańkosz Z, Winiarski S. (2020). Kinematic Parameters of Topspin Forehand in Table Tennis and Their Inter- and Intra-Individual Variability. *J Sports Sci Med.*, 19(1), 138-148. <https://doi.org/10.1155/2020/8413948>
- Bańkosz Z, Winiarski S. (2025). Kinematics of Topspin Stroke Combinations in Table Tennis and its Inter-Individual Variability. *J Sports Sci Med.*, 24(2), 311-325. <https://doi.org/10.52082/jssm.2025.311>
- Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *J Exp Psychol Hum Percept Perform.*, 16(1), 21-29. <https://doi.org/10.1037/0096-1523.16.1.21>
- Bootsma, R. J., Fernandez, L., Morice, A. H., & Montagne, G. (2010). Top-level players' visual control of interceptive actions: Bootsma and van Wieringen (1990) 20 years later. *J Exp Psychol Hum Percept Perform.*, 36(4), 1056-1066. <https://doi.org/10.1037/a0019327>
- De Morree, H. M., Klein, C., & Marcora, S. M. (2012). Perception of effort reflects central motor command during movement execution. *Psychophysiology*, 49(9), 1242-1253. <https://doi.org/10.1111/j.1469-8986.2012.01399.x>
- Ehara, Y., Beppu, M., Nomura, S., Kunimi, Y., Kouya, N., Tsutiya, T., Fujinawa, M., & Hoshi, M. (1998). Biomechanical analysis of the force produced in the shoulder during baseball pitching. *Biomechanisms*. 14, 39-48. <https://doi.org/10.3951/biomechanisms.14.39>
- Elliott, B. C., Marshall, R. N., & Noffal, G. J. (1995). Contributions of Upper Limb Segment Rotations during the Power Serve in Tennis. *J Appl Biomech.*, 11(4), 433-442. <https://doi.org/10.1123/jab.11.4.433>
- Hashimoto, T., Takiyama, K., Miki, T., Kobayashi, H., Nasu, D., Ijiri, T., Kuwata, M., KashinoM., & Nakazawa, K. (2021). Effort-dependent effects on uniform and diverse muscle activity features in skilled pitching. *Sci. Rep.*, 11, 8211. <https://doi.org/10.1038/s41598-021-87614-z>

- Hirashima, M., Kadota, H., Sakurai, S., Kudo, K., & Ohtsuki, T. (2002). Sequential muscle activity and its functional role in fast movement. *J Sports Sci.*, 20(4), 301-310. <https://doi.org/10.1080/026404102753576071>
- He, Y., Liang, M., Fang, Y., Fekete, G., Baker, J. S., & Gu, Y. (2023). Lumbar and pelvis movement comparison between cross-court and long-line topspin forehand in table tennis: Based on musculoskeletal model. *Frontiers in Bioengineering and Biotechnology*, 11, 1185177. <https://doi.org/10.3389/fbioe.2023.1185177>
- Le Mansec, Y., Dorel, S., Hug, F., & Jubeau, M. (2018). Lower limb muscle activity during table tennis strokes. *Sports Biomech.*, 17(4), 442-452. <https://doi.org/10.1080/14763141.2017.1354064>
- Lino, Y., Mori, T., & Kojima, T. (2008). Contributions of upper limb rotations to racket velocity in table tennis backhands against topspin and backspin. *J Sports Sci.*, 26(3), 287-293. <https://doi.org/10.1080/02640410701501705>
- Lino, Y., & Kojima, T. (2009). Kinematics of table tennis topspin forehands: effects of performance level and ball spin. *J Sports Sci.*, 27(12), 1311-1322. <https://doi.org/10.1080/02640410903264458>
- Lino, Y., & Kojima, T. (2011). Kinetics of the upper limb during table tennis topspin forehands in advanced and intermediate players. *Sports Biomech.*, 10(4), 361-377. <https://doi.org/10.1080/14763141.2011.629304>
- Kidokoro, S., Inaba Y., Yoshida K., Yamada K., & Ozaki H. (2018). Characteristics of three different topspin forehands in Japanese elite table tennis players: Impact for generating the batted ball with high initial velocity. *Jpn J Biomech Sports Exerc.*, 22(4), 152-166. https://doi.org/10.32226/jjbse.22_2018_003
- Kidokoro, S., Inaba Y., Yoshida K., Yamada K., & Ozaki H. (2019). The influence of shoulder and torso joint mobility on three different topspin forehands in a situation accompanied by body movement in Japanese elite table tennis players. *Jpn J Phys Educ Hlth Sport Sci.*, 64(1), 169-185. <https://doi.org/10.5432/jjpehss.18029>
- Kondrič, M., Furjan-Mandić, G., & Medved, V. (2006). Myoelectric comparison of table tennis forehand stroke using different ball sizes. *Acta Univ. Palacki. Olomuc., Gymn.*, 36(4), 25-28.
- Lees, A. (2003). Science and the major racket sports: A review. *J. Sports Sci.*, 21(9), 707-732. <https://doi.org/10.1080/0264041031000140275>
- Putnam, C. A. (1993). Sequential motions of body segments in striking and throwing skills: Descriptions and explanations. *J Biomech.*, 26(1), 125-135. [https://doi.org/10.1016/0021-9290\(93\)90084-R](https://doi.org/10.1016/0021-9290(93)90084-R)
- Sato, Y., Wakui, K., Miyazaki, K., & Ushiyama, Y. (2025). Effects of subjective effort on ball speed and spin in table tennis forehand topspin strokes. *International Journal of Sports Science & Coaching*, 1-11. <https://doi.org/10.1177/17479541251380169>
- Seemiller, D., & Holowchak, M. A. (1997). Winning table tennis. *Human Kinetics*.
- Tian, J., & Xiao, Y. (2024). Research on the difference of stroke characteristics and stroke effect between different stroke duration of table tennis players. *Scientific Reports*, 14, 25405. <https://doi.org/10.1038/s41598-024-76802-2>
- Tsai, C.-L., Pan, K.-M., Huang, K.-S., Chang, T.-J., Hsueh, Y.-C., Wang, L.-M., & Chang, S.-S. (2010). The surface EMG activity of the upper limb muscles in table tennis forehand drives. *Int. Symp. Biomech. Sports: Conf. Proc. Arch.*, 28, 1-4.
- Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., D'Lima, D. D., Cristofolini, L., Witte, H., Schmid, O., Stokes, I., & Standardization and Terminology Committee of the International Society of Biomechanics (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion--part I: ankle, hip, and spine. *International Society of Biomechanics. Journal of biomechanics*, 35(4), 543-548. [https://doi.org/10.1016/S0021-9290\(01\)00222-6](https://doi.org/10.1016/S0021-9290(01)00222-6)

- Wu, G., van der Helm, F. C., Veeger, H. E., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., Werner, F. W., Buchholz, B., & International Society of Biomechanics (2005). ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion--Part II: shoulder, elbow, wrist and hand. *Journal of biomechanics*, 38(5), 981-992. <https://doi.org/10.1016/j.jbiomech.2004.05.042>



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

Recontextualising self-regulated learning in ball games among beginning physical education teachers

 **Ferdinando Cereda**  . Department of Education. Catholic University of Sacred Heart. Milan, Italy.

ABSTRACT

This study examines how beginning physical education (PE) teachers in Italy recontextualise pedagogic discourse on ball games and self-regulated learning (SRL) as they transition from physical education teacher education (PETE) into school practice, and how self-efficacy shapes this process. Guided by Bernstein's pedagogic device and Bandura's self-efficacy, a sequential explanatory mixed-methods design was employed. In the quantitative phase, 287 beginning PE teachers (1–4 years' experience) completed the Italian validated Teacher Self-Efficacy Scale to Implement SRL (AI-AA), contextualised to ball games. Descriptive analyses indicated moderate overall self-efficacy, with the lowest scores for direct instruction of SRL strategies. In the qualitative phase, semi-structured interviews with 18 purposively selected teachers explained these patterns, identifying three recontextualised elements: promoting autonomy through modified game forms, adapting ball games to heterogeneous groups, and aligning tactical learning with curricular competencies. Despite these intentions, explicit SRL instruction in ball games was rare, and recontextualization was constrained by evaluative rules and material conditions (e.g., facilities, equipment, assessment pressures, and pupil expectations). Self-efficacy operated as a context-sensitive filter: lower efficacy was associated with retreat to traditional practices. Implications for PETE alignment and induction-focused professional development are discussed.

Keywords: Physical education, Physical education teacher education (PETE), Self-regulated learning, Ball games, Teacher self-efficacy, Recontextualization, Bernstein.

Cite this article as:

Cereda, F. (2026). Recontextualising self-regulated learning in ball games among beginning physical education teachers. *Scientific Journal of Sport and Performance*, 5(3), 461–481. <https://doi.org/10.55860/XWKG5712>



Corresponding author. Department of Education. Catholic University of Sacred Heart. Milan, Italy.

E-mail: ferdinando.cereda@unicatt.it

Submitted for publication February 09, 2026.

Accepted for publication March 16, 2026.

Published March 19, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/XWKG5712>

INTRODUCTION

The transition from physical education teacher education (PETE) to professional practice represents a critical juncture in teacher development, characterised by the recontextualization of pedagogic knowledge acquired during pre-service preparation (Backman et al., 2021; Mustell, 2024). Research consistently demonstrates that beginning teachers face substantial challenges in translating theoretical knowledge into classroom practice, particularly when confronted with institutional constraints, heterogeneous pupil populations, and established school cultures (Woods et al., 2016). These challenges are particularly acute in the domain of ball games teaching, where traditional multi-activity, technique-oriented approaches continue to dominate despite decades of advocacy for alternative pedagogies (Casey & Kirk, 2020; Kirk, 2010).

Ball games constitute a significant component of PE curricula internationally, reflecting their prominence in contemporary physical culture (Hardman, 2013). However, the pedagogic treatment of ball games in school settings has sustained considerable critique. Research indicates that ball games teaching frequently involves short instructional units focused on isolated technical skills, limited tactical understanding, and competitive formats that exclude less experienced pupils (Ennis, 1999; Larsson & Quennerstedt, 2016; Standal, 2016). Game-based approaches (GBAs) such as Teaching Games for Understanding (TGfU) and the Tactical Games Model (TGM) emerged as alternatives to traditional technique-centred instruction, emphasising tactical awareness, modified game forms, and learner-centred pedagogies (Harvey & Jarrett, 2014; Mitchell et al., 2020). Despite their pedagogical promise and widespread inclusion in PETE curricula, GBAs remain underutilised in school PE practice (García-López et al., 2019; Harvey & Pill, 2016).

Concurrently with these pedagogical developments, educational policy discourse has increasingly emphasised the cultivation of self-regulated learning (SRL) competencies as central to lifelong learning and academic success (OECD, 2020; Zimmerman, 2002). SRL encompasses cognitive, metacognitive, and motivational strategies that enable learners to understand and control their own learning processes (Schraw et al., 2006). Within PE contexts, SRL strategies hold particular relevance for ball games teaching, as they enable pupils to develop tactical decision-making, reflective practice, and autonomous engagement with game situations (Theobald, 2021). However, research suggests that PE teachers infrequently implement explicit SRL instruction, often adopting implicit approaches that lack systematic pedagogical support (Dignath-van Ewijk, 2013; Kistner et al., 2010).

Teacher self-efficacy—defined as teachers' beliefs about their capacity to execute specific teaching tasks at a desired level of quality (Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 2001)—represents a critical agentic resource influencing the negotiation and enactment of pedagogic knowledge. Teachers with higher domain-specific self-efficacy demonstrate greater persistence in implementing innovative practices, adapt more effectively to challenging classroom contexts, and achieve superior pupil learning outcomes (Peeters et al., 2014). Recent instrument development in Italy has provided a validated scale for assessing teachers' self-efficacy specifically for implementing SRL strategies (Scierri, 2024), yet this construct remains unexplored in relation to PE content areas such as ball games.

The Italian PE context presents distinctive features relevant to this investigation. Since 1994 and reaffirmed by the newly adopted National Regulations (D.M. 221/2025), the Italian national curriculum has not mandated specific sports as rigid content, instead emphasising broader educational goals related to motor competence, health promotion, and social development (Ministry of Education and Merit, 2025; Ministry of Education, University and Research, 2012). This orientation aligns with the consensus statement of the Italian society of motor and sports sciences, which advocates for a shift from a traditional sports-performance model to a

holistic educational approach based on physical literacy and inclusion (SISMeS, 2024). The new 2026 regulations specifically frame 'Giochi sportivi' (sports games) within cognitive, relational, and socio-emotional competencies, positioning ball games within a pedagogic discourse that must be legitimated through educational rather than purely sporting rationales (Colella, 2017; Cereda, 2023). Italian PETE programmes have increasingly incorporated game-based pedagogies and educational perspectives on sport. While recent scholarship has illuminated the complex epistemological and professional identity trajectories of Italian exercise science students (Cereda, 2025a), limited research has examined how beginning teachers recontextualise this knowledge in school settings or how their self-efficacy beliefs influence this process. This study addresses these gaps by investigating the recontextualization of ball games pedagogic discourse and SRL strategies among beginning PE teachers in Italy. Specifically, the research examines: (1) beginning teachers' self-efficacy for implementing SRL strategies in ball games contexts; (2) which elements of PETE pedagogic discourse regarding ball games and SRL are recontextualised in school practice; and (3) how evaluative rules and contextual factors regulate the reproduction of this pedagogic discourse. By integrating Bernstein's sociological framework of the pedagogic device with Bandura's socio-cognitive concept of self-efficacy, the study contributes theoretical and empirical insights into the complex processes through which PETE knowledge is transformed—or constrained—in professional practice.

Theoretical framework

This investigation draws upon two complementary theoretical perspectives: Bernstein's theory of the pedagogic device and Bandura's concept of self-efficacy. These frameworks are integrated to conceptualise teacher self-efficacy as an agentic mechanism in the recontextualization of pedagogic discourse from PETE to school PE.

Bernstein's pedagogic device and recontextualization

Bernstein's (2000) theory of the pedagogic device provides a sociological framework for understanding how knowledge is transformed into pedagogic communication across educational contexts. The pedagogic device operates through three interrelated sets of rules that regulate the production, recontextualization, and reproduction of discourse.

Distributive rules operate in the field of production, establishing boundaries for what constitutes legitimate knowledge. In relation to ball games, the field of production encompasses innovations in sport sciences, tactical theory, coaching methodologies, and educational research on game-based pedagogies. Distributive rules determine which forms of ball games knowledge are considered thinkable and valuable for educational transmission.

Recontextualising rules transform knowledge from the field of production into pedagogic discourse within the recontextualising field. This field comprises two subfields: the official recontextualising field (ORF), where state agencies construct curriculum policy, and the pedagogic recontextualising field (PRF), where teacher educators, textbook authors, and professional associations interpret and adapt official policy for pedagogic purposes (Singh, 2002). Pedagogic discourse is inherently hybrid, consisting of an instructional discourse (skills and competencies to be acquired) embedded within a regulative discourse (principles of social order, conduct, and identity) (Bernstein, 2000). The regulative discourse dominates, shaping what counts as legitimate knowledge and legitimate learner identity.

In Italian PETE, ball games pedagogic discourse is constructed primarily within the PRF, given the absence of specific curricular mandates. Teacher educators make selective decisions about which tactical models, game forms, and pedagogical approaches to include, and how to frame these within broader educational

goals. The recontextualization from sport science knowledge to PETE pedagogic discourse involves ideological positioning regarding the purposes of ball games in PE—whether oriented towards sporting performance, tactical understanding, inclusive participation, or social-emotional development.

Evaluative rules operate in the field of reproduction (schools), regulating pedagogic practice and assessment. These rules bring together the what (instructional discourse) and the how (regulative discourse) of teaching, defining the criteria by which legitimate pupil performance and teacher practice are recognised (Bernstein, 2000). Evaluative rules are shaped by institutional expectations, curricular requirements, assessment frameworks, and the material and social conditions of school contexts. For beginning teachers, evaluative rules exert powerful regulatory effects, often privileging established school practices over recently acquired PETE knowledge (Schempp et al., 1993).

Bernstein's framework illuminates how knowledge undergoes two transformative moments: first, in the movement from production to recontextualization (PETE), and second, in the movement from recontextualization to reproduction (school PE). Each transformation involves ideological selection, condensation, and adaptation. Beginning teachers act as recontextualising agents in the second transformation, negotiating between the pedagogic discourse encountered in PETE and the evaluative rules governing school practice.

Teacher self-efficacy as mediating mechanism

Bandura's (1997) social cognitive theory positions self-efficacy as a central construct influencing human agency and behaviour. Self-efficacy refers to individuals' beliefs in their capability to organise and execute courses of action required to produce given attainments. Unlike stable personality traits, self-efficacy beliefs are domain-specific, context-dependent, and malleable through experience (Bandura, 2006).

Teacher self-efficacy has demonstrated robust associations with instructional quality, persistence in the face of difficulties, and adoption of innovative practices (Tschannen-Moran & Woolfolk Hoy, 2001). Domain-specific measures of teacher self-efficacy provide greater predictive validity than generic instruments (Dellinger et al., 2008). The Teacher Self-Efficacy Scale to Implement Self-Regulated Learning (TSES-SRL), validated in Italy as the AI-AA scale (Scierri, 2024), assesses teachers' confidence in implementing both direct and indirect SRL strategies. The scale comprises four dimensions: self-efficacy for direct instruction of SRL strategies, self-efficacy for providing choices (indirect instruction), self-efficacy for providing challenges and complex tasks (indirect instruction), and self-efficacy for co-constructing assessment (indirect instruction).

In the context of recontextualization, self-efficacy beliefs function as mediating mechanisms that influence whether and how beginning teachers enact pedagogic discourse from PETE. Teachers with higher self-efficacy for implementing SRL strategies in ball games contexts may demonstrate greater persistence in using game-based pedagogies that require pupils to make tactical decisions, reflect on performance, and regulate their engagement. Conversely, teachers with lower self-efficacy may revert to more familiar technique-oriented approaches that minimise demands for pupil autonomy and metacognitive engagement.

Navigating structure and agency: Integrating the frameworks

This study integrates Bernstein's structural sociology with Bandura's social cognitivism to examine the dialectic between institutional regulation and individual agency. While originating from distinct ontological traditions—Bernstein emphasising the regulatory power of discursive rules and Bandura foregrounding the generative capability of the individual—their combination offers a dual lens on teacher socialisation. The

framework posits that recontextualization occurs within a space of tension rather than seamless alignment. Bernstein's evaluative rules define the boundaries of the 'thinkable' and 'legitimate' in school PE, creating structural constraints. Within these boundaries, self-efficacy operates as a proximal agentic resource and a context-sensitive interpretative filter. It determines whether teachers perceive these structural constraints as insurmountable barriers or navigable challenges.

Self-efficacy beliefs are thus conceptualised not merely as individual traits, but as structurally produced outcomes of differential access to the legitimate pedagogic discourse within the recontextualising field (PETE). Upon entering school contexts, beginning teachers encounter evaluative rules that may conflict with their PETE preparation. Material constraints (facilities, equipment), institutional expectations (assessment regimes), and pupil characteristics function as regulatory factors. In this context, self-efficacy regulates the extent to which PETE pedagogic discourse is sustained or transformed. High self-efficacy supports the negotiation of restrictive evaluative rules, fostering persistence in game-based approaches and SRL strategies. Conversely, low self-efficacy is associated with a reversion to traditional approaches when the reproductive field imposes conflicting demands. This recursive relationship positions self-efficacy as both a product of initial recontextualization and a driver of subsequent pedagogic reproduction.

METHODS

Research design

The study employed a sequential explanatory mixed-methods design (Creswell & Plano Clark, 2018), consisting of an initial quantitative phase followed by a qualitative phase designed to explain and elaborate quantitative findings. The quantitative phase examined the distribution of self-efficacy beliefs for implementing SRL strategies among beginning PE teachers and explored relationships between self-efficacy dimensions and background variables. The qualitative phase investigated how beginning teachers recontextualise ball games pedagogic discourse from PETE to school practice, and how self-efficacy beliefs intersect with evaluative rules and contextual factors in shaping pedagogic practice.

Participants

Participants were 287 beginning PE teachers (172 female, 115 male) currently employed in Italian state schools. Inclusion criteria required participants to: (1) hold a teaching qualification in PE (laurea magistrale in scienze motorie or equivalent); (2) have graduated within the previous five years (2019–2024); and (3) have between one and four years of PE teaching experience. Participants were recruited using convenience sampling through professional associations, social media networks, and direct contact with schools across 12 Italian regions. Due to the open nature of the recruitment, an exact response rate could not be calculated. It is acknowledged that this strategy limits representativeness and may introduce self-selection bias, as teachers with a pre-existing interest in SRL or innovative pedagogies might have been more likely to participate. Despite these limitations, the sample represented diverse geographical contexts (Northern regions: 45%; Central regions: 33%; Southern regions and islands: 22%) and school levels (lower secondary: 58%; upper secondary: 42%). Mean age was 28.7 years (SD = 3.2), and mean teaching experience was 2.1 years (SD = 0.9). Demographic characteristics are summarised in Table 1.

Following quantitative data analysis, purposeful sampling (Patton, 2015) was employed to select 18 participants for semi-structured interviews. Sampling criteria ensured diversity in self-efficacy profiles (high, moderate, low across AI-AA dimensions), PETE institution attended (six institutions represented), geographical region, school level, and gender (9 female, 9 male). Recruitment continued until data saturation was reached, defined as the point where no new themes or dimensions emerged (Saunders et al., 2018).

Furthermore, negative case analysis was actively pursued by specifically inviting teachers with low self-efficacy scores ($n = 6$). However, consistent with avoidance behaviours often associated with low self-efficacy, only three of these teachers agreed to participate, ensuring that disconfirming evidence regarding the recontextualization process was represented. Selected participants had graduated from PETE programmes between 2020 and 2023 and possessed 1–3 years of teaching experience. Most participants ($n = 16$) had received instruction in game-based approaches during PETE; two participants (Francesca and Roberto) reported traditional technique-focused preparation, as confirmed through institutional curriculum documents (see Table 2.)

Table 1. Demographic characteristics of participants in quantitative phase ($N = 287$).

Characteristic	Category	n	%
Gender	Female	172	59.9
	Male	115	40.1
Age (years)	24–26	89	31.0
	27–29	121	42.2
	30–35	77	26.8
	M (SD)	28.7 (3.2)	
	1	78	27.2
Teaching experience (years)	2	102	35.5
	3	71	24.7
	4	36	12.5
	M (SD)	2.1 (0.9)	
Geographical region	Northern Italy	129	44.9
	Central Italy	95	33.1
	Southern Italy and islands	63	22.0
School level	Lower secondary (grades 6–8)	166	57.8
	Upper secondary (grades 9–13)	121	42.2
Contract type	Permanent position	118	41.1
	Fixed-term contract	169	58.9
PETE institution type	Traditional university programme	198	69.0
	Post-graduate qualification pathway	89	31.0
GBA instruction in PETE	Yes	214	74.6
	No	73	25.4

Note. GBA = Game-based approaches (including TGfU, TGM, or similar pedagogical models). PETE = Physical education teacher education.

Instruments and data collection

The *quantitative instrument* AI-AA scale (Scierri, 2024) was administered via online survey. The validated Italian instrument comprises 17 items distributed across four dimensions: Direct Instruction of SRL Strategies (6 items; e.g., “*Insegnare ai suoi studenti come usare e applicare le diverse strategie di apprendimento autoregolato*” [Teach your students how to use and apply different self-regulated learning strategies]); Providing Choices (4 items; e.g., “*Scegliere insieme ai suoi studenti quando svolgere determinate attività di apprendimento*” [Choose together with your students when to carry out certain learning activities]); Providing Challenges and Complex Tasks (4 items; e.g., “*Dare esercizi impegnativi che possono essere svolti in modi diversi e non prestabiliti*” [Give challenging exercises that can be carried out in different and not predetermined ways]); and Co-constructing Assessment (3 items; e.g., “*Coinvolgere i suoi studenti nella preparazione dei criteri di valutazione dei compiti*” [Involve your students in preparing task assessment criteria]). Responses were recorded on a six-point Likert scale ranging from 1 (per niente in grado [not at all capable]) to 6 (del tutto in grado [completely capable]).

For this study, an additional contextualising prompt was included: “*Consider specifically your teaching of ball games (e.g., invasion games such as basketball, handball, football; net/wall games such as volleyball;*

striking/fielding games) when responding to the following items.” This modification ensured that self-efficacy assessments were anchored to ball games contexts rather than general PE instruction.

Table 2. Characteristics of interview participants (n = 18)

Pseudonym	Gender	Age	Years teaching	School level	Region	PETE institution	AI-AA profile	Sporting biography
Alessia	F	26	2	Lower sec.	Central	University A	Moderate-high	Volleyball, swimming
Marco	M	29	3	Upper sec.	Northern	University B	High	Basketball, athletics
Giulia	F	27	2	Lower sec.	Southern	University C	Moderate	Volleyball club player
Davide	M	30	3	Upper sec.	Northern	University B	High	Football, tennis
Elena	F	28	2	Lower sec.	Central	University D	Moderate-high	Rhythmic gymnastics
Simone	M	29	2	Lower sec.	Northern	University A	Moderate-high	Handball, swimming
Francesca	F	25	1	Lower sec.	Southern	University E	Low-moderate	Dance, fitness activities
Luca	M	31	3	Upper sec.	Central	University F	High	Football coach, athletics
Chiara	F	27	2	Lower sec.	Central	University A	Moderate	Basketball, skiing
Andrea	M	30	3	Upper sec.	Northern	University B	High	Volleyball, football
Matteo	M	26	2	Lower sec.	Southern	University C	Moderate	Football, martial arts
Sofia	F	28	2	Upper sec.	Central	University D	Moderate	Tennis, fitness
Roberto	M	27	2	Lower sec.	Southern	University E	Low-moderate	Football, basketball
Valentina	F	29	2	Lower sec.	Central	University A	Moderate	Volleyball, swimming
Federica	F	25	1	Lower sec.	Northern	University F	Low-moderate	Dance, gymnastics
Giuseppe	M	32	3	Upper sec.	Southern	University C	Moderate-high	Football, handball
Laura	F	28	2	Upper sec.	Northern	University B	High	Basketball coach
Paolo	M	29	3	Upper sec.	Central	University D	Moderate-high	Athletics, swimming

Note. F = Female; M = Male; sec. = secondary. AI-AA profile categories based on total scale scores: Low-moderate ($M = 2.50\text{--}3.49$); Moderate ($M = 3.50\text{--}4.19$); Moderate-high ($M = 4.20\text{--}4.79$); High ($M \geq 4.80$). PETE institutions anonymised (A–F represent six different Italian universities offering physical education teacher education programmes). All participants had received instruction in game-based approaches during PETE except Francesca and Roberto.

For *qualitative data collection* semi-structured interviews were conducted between October 2024 and February 2025 via secure video conferencing platform. The interview protocol, adapted from Mustell's (2024) framework, included questions addressing: participants' sporting biographies and motivations for PE teaching; experiences of ball games pedagogy within PETE (content, pedagogical approaches, practicum experiences); current ball games teaching practices (content selection, instructional strategies, use of SRL approaches); perceived usefulness and relevance of PETE preparation; challenges encountered in implementing ball games and SRL strategies; and perceptions of contextual factors influencing practice.

Interviews lasted 42–68 minutes (mean = 54 minutes) and were audio-recorded with informed consent. While a specific Confirmatory Factor Analysis (CFA) for the ball games context was not conducted, the internal consistency reliability coefficients (Cronbach's α) for this administration were high (0.81–0.89), suggesting that construct stability was maintained under the contextualised conditions.

Data analysis

Quantitative data were analysed using SPSS 28.0. Descriptive statistics (means, standard deviations) were calculated for each AI-AA dimension and the total scale. Internal consistency reliability was assessed using Cronbach's alpha. Correlations between AI-AA dimensions and between AI-AA scores and background variables (teaching experience, PETE institution type, school level) were examined using Pearson correlation coefficients. To examine the predictive contribution of PETE preparation, hierarchical multiple regression analysis was performed on the total AI-AA score, controlling for teaching experience and gender. Independent samples t-tests compared self-efficacy scores across categorical variables (gender, geographical region). Statistical significance was set at $p < .05$.

For qualitative analysis, interview recordings were transcribed verbatim, yielding 412 pages of transcript data. Data were imported into NVivo 12 (QSR International) to facilitate data management and coding organization. Thematic analysis following Braun and Clarke's (2021) reflexive approach was conducted. To enhance rigour and trustworthiness, the research team engaged in reflexive journaling throughout the process, documenting analytic decisions and personal assumptions stemming from their backgrounds as teacher educators. While the initial coding was inductive, subsequent theme development was guided by theoretical questioning derived from Bernstein's framework: Which elements of instructional and regulative discourse from PETE are reproduced? How are these transformed in response to evaluative rules? Which institutional and material conditions regulate practice? Analytic disagreements were resolved through deliberative dialogue rather than consensus-seeking, preserving interpretive nuance. An audit trail of coding decisions and theme evolution was maintained. Preliminary themes were reviewed against coded extracts and entire transcripts to ensure coherence. Participant quotations were translated into English, preserving semantic and pragmatic meaning.

Ethical considerations

The study received approval from the University Territorial Ethics Committee (UTEC 2024-67). All participants provided written informed consent prior to data collection, with confidentiality assured through pseudonymisation. Participants were informed of their right to withdraw at any stage without consequence.

RESULTS

Predictors of self-efficacy

Hierarchical regression analysis was conducted to determine the variance in total self-efficacy explained by GBA exposure after controlling for background variables. In Step 1, gender and teaching experience explained a small but significant proportion of variance ($R^2 = .05$, $p < .05$), with experience emerging as a positive predictor ($\beta = .19$, $p = .003$). In Step 2, explicit instruction in Game-Based Approaches (GBA) during PETE was added. This model explained significantly more variance ($R^2 = .18$, $\Delta R^2 = .13$, $p < .001$). GBA exposure was a significant positive predictor ($\beta = .37$, $p < .001$), indicating that specific pedagogical preparation in PETE is a stronger predictor of self-efficacy for SRL in ball games than teaching experience alone. No significant interaction effects were found between experience and GBA exposure.

Quantitative findings

Table 3 presents descriptive statistics for the AI-AA scale dimensions. Overall, beginning teachers reported moderate levels of self-efficacy for implementing SRL strategies in ball games contexts (total scale $M = 3.82$, $SD = 0.71$). Across dimensions, self-efficacy for providing challenges and complex tasks was highest ($M = 4.21$, $SD = 0.68$), followed by providing choices ($M = 3.94$, $SD = 0.79$) and co-constructing assessment ($M = 3.61$, $SD = 0.88$). Self-efficacy for direct instruction of SRL strategies was lowest ($M = 3.52$, $SD = 0.83$). Internal consistency reliability was acceptable to excellent across all dimensions (Cronbach's α ranging from 0.81 to 0.89) and for the total scale ($\alpha = 0.91$).

Table 3. Descriptive statistics and internal consistency reliability for AI-AA scale dimensions (N = 287)

AI-AA Dimension	No. of items	M	SD	Range	Skewness	Kurtosis	Cronbach's α
Direct Instruction of SRL Strategies (ID)	6	3.52	0.83	1.00–6.00	–0.28	–0.41	0.89
Providing Choices (II-S)	4	3.94	0.79	1.25–6.00	–0.35	–0.18	0.84
Providing Challenges and Complex Tasks (II-SC)	4	4.21	0.68	2.00–6.00	–0.47	0.22	0.86
Co-constructing Assessment (II-V)	3	3.61	0.88	1.00–6.00	–0.19	–0.52	0.81
Total AI-AA Scale	17	3.82	0.71	1.71–5.88	–0.31	–0.15	0.91

Note. AI-AA = Autoefficacia dell'Insegnante per l'implementazione dell'Apprendimento Autoregolato (Teacher Self-Efficacy Scale to Implement Self-Regulated Learning). SRL = Self-regulated learning. Responses were recorded on a six-point Likert scale: 1 (per niente in grado [not at all capable]) to 6 (del tutto in grado [completely capable]). All skewness and kurtosis values fall within acceptable ranges for normal distribution ($|\text{skewness}| < 2.0$; $|\text{kurtosis}| < 7.0$).

As shown in Table 4, all four AI-AA dimensions demonstrated significant positive intercorrelations ($r = 0.59$ – 0.74 , $p < .001$), indicating coherent but distinguishable facets of SRL implementation self-efficacy. Teaching experience (years) correlated weakly but significantly with Direct Instruction ($r = 0.19$, $p = .002$) and Co-constructing Assessment ($r = 0.16$, $p = .007$), suggesting modest increases in these efficacy beliefs over the initial years of practice. No significant correlations emerged between AI-AA dimensions and school level or geographical region.

Table 4. Intercorrelations Among AI-AA Dimensions and Correlations with Teaching Experience (N = 287)

Variable	1	2	3	4	5
1. Direct Instruction (ID)	—				
2. Providing Choices (II-S)	.68***	—			
3. Providing Challenges (II-SC)	.61***	.72***	—		
4. Co-constructing Assessment (II-V)	.59***	.74***	.66***	—	
5. Total AI-AA	.84***	.90***	.85***	.86***	—
6. Teaching experience (years)	.19**	.11	.08	.16**	.14*

Note. AI-AA = Autoefficacia dell'Insegnante per l'implementazione dell'Apprendimento Autoregolato. * $p < .05$. ** $p < .01$. *** $p < .001$ (two-tailed).

Table 5. Comparison of AI-AA dimension scores by exposure to game-based approaches in PETE.

AI-AA Dimension	GBA instruction (n = 214)	No GBA instruction (n = 73)	t	df	p	Cohen's d
	M (SD)	M (SD)				
Direct Instruction (ID)	3.57 (0.81)	3.42 (0.87)	1.48	285	.140	0.18
Providing Choices (II-S)	4.08 (0.74)	3.64 (0.82)	4.32	285	<.001	0.56
Providing Challenges (II-SC)	4.32 (0.63)	3.97 (0.74)	3.89	285	<.001	0.51
Co-constructing Assessment (II-V)	3.75 (0.84)	3.32 (0.94)	3.71	285	<.001	0.48
Total AI-AA	3.93 (0.67)	3.59 (0.76)	3.79	285	<.001	0.48

Note. AI-AA = Autoefficacia dell'Insegnante per l'implementazione dell'Apprendimento Autoregolato. GBA = Game-based approaches (TGfU, TGM, or similar pedagogical models) during physical education teacher education (PETE). Independent samples t-tests with equal variances assumed (Levene's test $p > .05$ for all comparisons). Cohen's d values: 0.20 = small effect, 0.50 = medium effect, 0.80 = large effect.

Independent samples t-tests (Table 5) revealed no significant gender differences across AI-AA dimensions. However, teachers who reported having received explicit instruction in game-based approaches (TGfU, TGM, or similar models) during PETE ($n = 214$) demonstrated significantly higher self-efficacy for Providing Choices ($M = 4.08$, $SD = 0.74$) compared to those without such instruction ($n = 73$; $M = 3.64$, $SD = 0.82$; $t(285) = 4.32$, $p < .001$, $d = 0.56$). This pattern was also evident for providing challenges ($M = 4.32$ vs. $M = 3.97$; $t(285) = 3.89$, $p < .001$, $d = 0.51$) and co-constructing assessment ($M = 3.75$ vs. $M = 3.32$; $t(285) = 3.71$, $p < .001$, $d = 0.48$), but not for direct instruction ($p = .14$).

Qualitative findings

Qualitative analysis identified three primary themes representing recontextualised elements of pedagogic discourse, alongside two themes illuminating constraints on recontextualization. These are presented below.

Theme 1: fostering pupil autonomy through modified game forms

A central element of pedagogic discourse from PETE that beginning teachers sought to recontextualise involved using modified games to foster pupil autonomy in tactical decision-making and strategic thinking. This element reflects both instructional discourse (tactical understanding as content) and regulative discourse (pupils as active, autonomous learners). Nearly all participants described exposure to game-based approaches during PETE and articulated their educational value:

During university, we learned about Teaching Games for Understanding and the Tactical Games Model. The philosophy is that you don't just teach techniques in isolation—you put students into game situations where they have to make decisions and solve tactical problems. This prepares them to be intelligent players, not just robots who can execute skills. (Alessia, 2 years' experience, Central Italy).

The idea from university was that games should be adapted to the learners' level, not the other way around. You modify rules, space, equipment, so everyone can participate and learn tactical concepts. For example, in basketball you might use smaller courts, allow more dribbles, or require a minimum number of passes before shooting. This makes the game more inclusive and focuses on understanding, not just skill. (Marco, 3 years' experience, Northern Italy)

However, participants' descriptions of actual practice revealed a complex negotiation rather than simple reproduction, characterised by selective and partial implementation of game-based pedagogies. Most teachers reported using modified games occasionally, particularly for unfamiliar sports where pupils lacked established expectations. Modified games for popular sports (football, volleyball) were described as more challenging to implement:

With volleyball, I try to use smaller courts and lighter balls, sometimes allowing one bounce. But many students, especially the boys who play in clubs, resist these modifications. They say, 'this isn't real volleyball.' They want to play six-on-six with normal rules. It's frustrating because I know the modifications help the beginners, but I also need to keep the experienced players engaged. (Giulia, 2 years' experience, Southern Italy)

The recontextualization of game-based pedagogies thus involved negotiation between PETE discourse emphasising inclusive, educative game forms and evaluative rules shaped by pupils' sporting expectations and competitive orientations. Consistent with the quantitative profile of high efficacy (AI-AA > 4.20), teachers like Davide described greater persistence in implementing modified games despite resistance:

I'm confident now in adapting games to different levels. At first I was worried about managing the class if I changed too many rules, but I've learned that if you explain the learning purpose clearly and

give students some choices—like choosing positions or deciding how to score—they accept the modifications better. My professor at university said, ‘never apologise for modifying the game; explain why it serves learning,’ and that advice has really helped me. (Davide, 3 years’ experience, Northern Italy)

Theme 2: adapting ball games to heterogeneous groups

A second recontextualised element concerned inclusive pedagogy oriented towards heterogeneous pupil populations. This element was universally present in participants’ accounts of PETE pedagogic discourse and strongly aligned with Italian curricular emphasis on personalised learning and inclusion. The regulative discourse here positioned all pupils as legitimate learners regardless of prior sporting experience:

In university, we constantly discussed how to include everyone in ball games. Not just students with disabilities, but also those who are shy, who have negative experiences from sports clubs, who come from families that don’t value sport. We learned strategies like small-sided games, cooperative rules, rotating positions, assessment based on improvement not absolute performance. (Elena, 2 years’ experience, Central Italy)

In school practice, participants described heterogeneity as simultaneously a core teaching value and a major practical challenge. Strategies for adaptation included differentiating task complexity, forming flexible groupings, and selecting fewer familiar sports:

I have classes where some students play in elite youth academies and others have never touched a basketball. Traditional drills don’t work—the advanced students are bored and the beginners are lost. So I use games with multiple objectives. For example, in handball I might have different scoring zones worth different points and require teams to use certain tactical movements. Advanced students focus on complex tactics; beginners focus on basic positioning and passing. Everyone participates, but at their level. (Simone, 2 years’ experience, Northern Italy)

However, several participants with lower self-efficacy for Providing Challenges expressed difficulty sustaining differentiated instruction:

I know I should adapt activities to different levels, but honestly it’s exhausting. You plan three different versions of the same game, and it still doesn’t work because the groups are so diverse. Sometimes I just let them play and circulate to give individual feedback. I know this isn’t ideal, but I don’t have enough confidence yet to manage truly differentiated lessons. (Francesca, 1 year experience, Southern Italy)

This theme illustrates how recontextualization of inclusive pedagogic discourse was mediated by self-efficacy, particularly confidence in designing and managing complex, differentiated tasks.

Theme 3: Connecting tactical learning to curricular goals

The third recontextualised element involved framing ball games instruction around curricular competencies rather than sport-specific performance outcomes. This element reflects transformation in the regulative discourse, positioning ball games as means to broader educational ends rather than intrinsically valuable content:

The most important thing I learned at university was that ball games in school are not about training athletes. They are about developing competencies: motor competence, yes, but also strategic thinking, cooperation, emotional regulation. So when I teach football, the goal is not to produce good football players but to use football to develop these broader competencies. This completely changed how I think about assessment—I’m not assessing how well they kick but how they solve tactical

problems, how they cooperate, whether they can reflect on their learning. (Luca, 3 years' experience, Central Italy)

Most participants described this curricular orientation as central to their pedagogic identity, distinguishing them from sport coaches. However, translating this orientation into coherent assessment practices proved challenging:

I believe in assessing learning, not just performance, but I struggle with how to do this practically. The curriculum says to assess problem-solving and strategic thinking, but when you have 25 students playing a game, how do you assess that systematically? I observe and take notes, but it feels subjective. I wish university had given us more concrete assessment tools. (Chiara, 2 years' experience, Central Italy)

Reinforcing the quantitative link between specific efficacy beliefs and practice, teachers with higher self-efficacy for Co-constructing Assessment (e.g., Andrea, high efficacy profile) more frequently described involving pupils in defining success criteria and self-assessment:

After we play a tactical game, I have students discuss what tactical principles they used—like creating space or supporting the ball carrier—and we develop criteria together for what successful application looks like. Then in the next lesson they assess themselves and peers against those criteria. This makes the learning visible and helps them regulate their own progress. (Andrea, 3 years' experience, Northern Italy)

This practice reflects sophisticated recontextualization of both game-based pedagogies (focusing on tactical concepts) and SRL strategies (involving pupils in assessment), yet it was described by a minority of participants with high domain-specific self-efficacy.

Theme 4: Absence of explicit SRL instruction in ball games

Despite moderate self-efficacy scores for Direct Instruction of SRL strategies, qualitative data revealed minimal explicit teaching of SRL strategies within ball games contexts. When probed about teaching pupils how to plan tactical approaches, monitor their decision-making, or reflect on performance, most participants acknowledged this occurred only implicitly:

I encourage students to think tactically and sometimes I ask questions like 'why did you choose to pass instead of shoot?' but I don't explicitly teach them metacognitive strategies. Like, I don't teach them how to set goals for a game or how to evaluate their tactical decisions systematically. Maybe I should, but I don't really know how to do that in a ball games context. (Matteo, 2 years' experience, Southern Italy)

Several participants indicated that SRL strategies were addressed in other PE content areas (e.g., fitness, gymnastics) but not systematically integrated into ball games units:

We did learn about self-regulated learning in university, but mainly in theoretical courses, not specifically applied to ball games. I use SRL strategies when I teach health and fitness topics—students set goals, plan workout programmes, self-monitor. But in ball games I focus more on tactical concepts and teamwork. I hadn't really thought about connecting them until this interview. (Sofia, 2 years' experience, Central Italy)

This finding suggests incomplete recontextualization of SRL pedagogic discourse into ball games practice, despite its presence in PETE curriculum. Lower self-efficacy for Direct Instruction may partly explain this gap, alongside insufficient modelling of SRL integration within game-based pedagogies during PETE.

Theme 5: evaluative rules and material constraints regulating practice

Participants consistently identified contextual factors that constrained recontextualization of PETE pedagogic discourse. These factors functioned as evaluative rules shaping what pedagogic practices were feasible and legitimate.

Facilities and equipment. Material conditions directly regulated pedagogic possibilities. Many schools lacked adequate indoor facilities, forcing outdoor instruction in adverse weather or shared use of multipurpose spaces. Limited equipment (e.g., insufficient balls for small-sided games) prevented implementation of pedagogies requiring simultaneous active participation:

My university had excellent facilities—multiple gymnasiums, abundant equipment. When I started teaching, I discovered my school has one small gym shared with three other teachers, and we have twelve basketballs for classes of 25 students. The game-based approaches I learned require small-sided games with everyone actively involved, but I physically cannot do that. So I fall back on drills where students take turns. (Roberto, low efficacy profile, 2 years' experience, Southern Italy)

Institutional assessment requirements. While Italian PE assessment formally emphasises competency development, participants described institutional pressures towards quantifiable performance outcomes:

The school wants me to give numeric grades based on clear criteria. I can't just write 'this student demonstrates good tactical understanding.' So I end up assessing technical skills because they're easier to measure—how many successful passes, shooting accuracy, things like that. This contradicts what I learned about assessing strategic thinking, but I need defensible grades. (Valentina, 2 years' experience, Central Italy)

Pupil expectations and resistance. As noted earlier, pupils' prior sporting experiences generated expectations that sometimes conflicted with educational approaches:

Many students, especially those in sports clubs, view PE as a chance to play competitively, not to work on tactics or self-regulation. They get frustrated when I stop the game for reflection or when I modify rules for learning purposes. Some boys say, 'coach doesn't do it this way in the club, why are we doing it differently here?' Managing these expectations is draining. (Federica, 1 year experience, Northern Italy)

Table 6. Integration of quantitative profiles and qualitative themes: a joint display of recontextualization patterns.

AI-AA profile	Quantitative characteristic (mean range)	Dominant recontextualization strategy (qualitative theme)	Key pedagogic action	Typical quote (participant)
High Efficacy	Total Score > 4.20	Negotiated Enactment (Theme 1 & 3)	Persists with modified games; explains rationale to resistant pupils; uses formative assessment.	"I explain why it serves learning... and that advice has really helped me." (Davide)
Moderate Efficacy	Total Score 3.50–4.19	Selective Reproduction (Theme 2)	Uses game modifications only when convenient; struggles with differentiation in heterogeneous groups.	"I try to use smaller courts... but many students resist." (Giulia)
Low Efficacy	Total Score < 3.50	Strategic Retreat (Theme 5)	Abandons innovative practices in face of constraints; reverts to drills/traditional instruction.	"I physically cannot do that. So I fall back on drills." (Roberto)

Note. This joint display illustrates how domain-specific self-efficacy functions as an interpretive filter regulating the recontextualization of PETE discourse.

These evaluative rules and contextual constraints illuminated the reproductive field's regulatory power over pedagogic practice, often overriding the innovative pedagogic discourse encountered in PETE. To synthesise these findings, Table 6 provides a joint display integrating the quantitative self-efficacy profiles with the qualitative recontextualization patterns identified.

DISCUSSION

This study examined how beginning PE teachers in Italy recontextualise ball games pedagogic discourse and SRL strategies in the transition from PETE to school practice, and how self-efficacy beliefs intersect with this process. The integration of Bernstein's sociological framework with Bandura's concept of self-efficacy provided theoretical leverage for understanding recontextualization as a negotiated act shaped by both structural-institutional forces and individual agentic resources.

Self-efficacy profiles and their pedagogical significance

Quantitative findings revealed moderate levels of self-efficacy for implementing SRL strategies in ball games, with notable variation across dimensions. Self-efficacy for Direct Instruction was lowest, consistent with qualitative evidence that explicit SRL teaching remained largely absent from ball games practice. This pattern suggests incomplete recontextualization of SRL pedagogic discourse from PETE to the reproduction field. While PETE programmes may introduce SRL theoretical frameworks and generic strategies, recent analyses of Italian kinesiology curricula reveal significant fragmentation and a predominant biomedical bias that often marginalises pedagogical training (Cereda, 2025b). This structural imbalance results in insufficient contextualisation within ball games teaching and appears to limit beginning teachers' confidence and competence in this domain. This finding resonates with broader research indicating that SRL implementation in PE contexts remains limited (Dignath-van Ewijk, 2013; Kistner et al., 2010) and extends this understanding by identifying domain-specific self-efficacy as a contributing mechanism.

Conversely, higher self-efficacy for Providing Challenges and Complex Tasks aligns with widespread exposure to game-based approaches during PETE. Italian PETE programmes have increasingly incorporated TGfU, TGM, and related models (Valentini & Rudisill, 2004), and participants in this study generally received instruction in these pedagogies. The significant association between explicit PETE instruction in game-based approaches and self-efficacy for indirect SRL strategies (Providing Choices, Providing Challenges, Co-constructing Assessment) suggests that modelling these pedagogies within PETE contributes to efficacy development. This finding supports recent advocacy for instructional alignment in PETE, wherein pre-service teachers experience pedagogical approaches they are expected to subsequently implement (MacPhail et al., 2023).

However, the correlation between teaching experience and self-efficacy dimensions was modest, indicating that experience alone does not substantially enhance efficacy beliefs during the initial years of practice. This suggests that efficacy development requires more than accumulated time; structured professional development, mentoring, and opportunities for reflective practice may be necessary to strengthen efficacy through mastery experiences and social persuasion (Bandura, 1997).

Recontextualised elements and selective reproduction

Qualitative analysis identified three elements of PETE pedagogic discourse that beginning teachers sought to recontextualise: fostering pupil autonomy through modified games, adapting ball games to heterogeneous groups, and connecting tactical learning to curricular goals. These elements reflect both instructional discourse (tactical understanding, motor competence) and regulative discourse (pupils as autonomous,

capable learners deserving inclusive pedagogy). Their prominence in participants' accounts suggests that Italian PETE graduates internalise educational perspectives on ball games that distinguish PE from sport club contexts (Larsson & Karlefors, 2015; Lundvall & Meckbach, 2008). This represents a significant achievement given persistent concerns about sportisation of PE and the powerful influence of teachers' sporting biographies (Dowling, 2011; Mordal-Moen & Green, 2014).

However, recontextualization involved selective and partial reproduction rather than faithful enactment. Participants frequently described tensions between educational ideals and practical realities, resulting in hybrid pedagogies that combined elements of game-based approaches with more traditional practices. This hybridity aligns with findings from previous research on GBA implementation and recent scholarship within the Italian context (Cereda, 2023), which documents teachers' tendency to blend constructivist and behaviourist approaches (Díaz-Cueto et al., 2010; O'Leary, 2014). From a Bernsteinian perspective, this hybridity reflects the ideological work inherent in recontextualization: beginning teachers must reconcile competing principles (educational vs. sporting, inclusive vs. performance-oriented, learner-centred vs. teacher-directed) within specific institutional contexts.

Self-efficacy functioned as an agentic mechanism in this process. Teachers with higher self-efficacy for Providing Choices and Providing Challenges demonstrated greater persistence in implementing modified games despite pupil resistance and institutional pressures. They described confidence in explaining pedagogical rationales, negotiating with pupils, and adapting approaches based on contextual feedback. Conversely, teachers with lower self-efficacy more readily abandoned innovative practices when confronted with challenges, reverting to familiar drill-based or play-focused formats. This pattern illustrates how self-efficacy is intertwined with the agentic dimension of recontextualization: while structural constraints impose real limits, self-efficacy beliefs appear to moderate whether teachers perceive these constraints as surmountable obstacles or insurmountable barriers.

Evaluative rules and the regulation of pedagogic practice

Bernstein's concept of evaluative rules proved analytically productive in understanding how contextual factors regulate the reproduction of pedagogic discourse. Material conditions (facilities, equipment), institutional assessment regimes, and pupil expectations functioned as evaluative rules that defined legitimate pedagogic practice within the reproduction field. These rules frequently contradicted the pedagogic discourse promoted in PETE, creating disjunctures that beginning teachers struggled to navigate.

Material constraints represented the most concrete form of regulation. Inadequate facilities and insufficient equipment directly limited the enactment of small-sided game formats central to game-based pedagogies. This finding corroborates previous research identifying material conditions as significant barriers to innovative PE practice (MacPhail et al., 2008) and extends this by demonstrating how such conditions function as evaluative rules that determine what pedagogic practices are thinkable and achievable. From a Bernsteinian perspective, these material conditions are not neutral obstacles but rather embody institutional priorities and resource allocation decisions that regulate which forms of PE are valued and supported.

Institutional assessment requirements constituted a second evaluative force. Despite Italian curricular rhetoric emphasising competency-based assessment, participants described pressures towards quantifiable performance metrics. This tension reflects broader contradictions within educational policy between progressive pedagogical ideals and accountability mechanisms privileging measurable outcomes (Evans, 2014). Teachers with higher self-efficacy for Co-constructing Assessment partially resisted these pressures by involving pupils in defining success criteria and using qualitative self-assessment, thereby maintaining

alignment with PETE pedagogic discourse. However, such practices remained exceptional rather than normative.

Pupil expectations and resistance represented a third regulatory force, one insufficiently addressed in PETE preparation. Participants frequently described surprise at the intensity of pupils' sporting identities and competitive orientations, which contradicted assumptions about learner receptivity to educational game forms. This finding echoes Mustell's (2024) identification of pupil expectations as a significant contextual factor regulating ball games pedagogy. The regulatory power of pupil expectations illuminates how social subjects within the reproduction field actively shape pedagogic discourse, rather than passively receiving transmitted knowledge. Beginning teachers with robust self-efficacy demonstrated greater capacity to negotiate these expectations through pedagogical persuasion, but many felt underprepared for this aspect of practice.

Implications for PETE and professional development

Findings suggest several implications for enhancing the impact of PETE and supporting beginning teachers' professional development, particularly within the broader goal of fostering physical literacy (Cereda, in press). First, PETE curricula should provide explicit instruction in integrating SRL strategies within ball games teaching. Rather than addressing SRL as generic content separate from subject matter, PETE should model how tactical game-based pedagogies can incorporate SRL principles such as goal-setting for tactical improvement, peer assessment of strategic decisions, and metacognitive reflection on game performance. This recommendation aligns with calls for instructional alignment in PETE (MacPhail et al., 2023) and addresses the identified gap in beginning teachers' self-efficacy for Direct Instruction of SRL strategies.

Second, PETE should prepare pre-service teachers more effectively for the evaluative rules and material constraints they will encounter in schools. While PETE programmes cannot replicate all school contexts, they can: (1) incorporate authentic school-based scenarios into coursework, requiring pre-service teachers to plan ball games units with realistic constraints; (2) foster critical analysis of institutional forces that shape pedagogic practice; and (3) develop adaptive expertise in modifying game-based pedagogies for varied contexts. School placement experiences should be strategically designed to expose pre-service teachers to diverse material conditions and institutional cultures, with structured reflection on how these contexts enable or constrain innovative practice.

Third, pre-service teachers would benefit from explicit preparation in managing pupil resistance and negotiating expectations. This involves developing communication skills for explaining pedagogical rationales, strategies for building legitimacy for modified game forms, and resilience when innovative practices meet initial resistance. Role-play activities, case study discussions, and mentoring from experienced teachers who successfully implement game-based approaches could support this development.

Fourth, findings underscore the importance of ongoing professional development for beginning teachers. The modest correlation between teaching experience and self-efficacy suggests that experience alone is insufficient for efficacy enhancement. Structured induction programmes, collaborative learning communities focused on ball games pedagogy, and opportunities to observe and discuss innovative practices with peers could provide the mastery experiences, vicarious learning, and social persuasion necessary for strengthening self-efficacy (Bandura, 1997). Such support structures would help beginning teachers sustain the pedagogic discourse acquired in PETE rather than abandoning it under institutional pressures.

Theoretical contributions

This study advances the international debate on teacher socialisation by moving beyond the binary of 'washing out' versus 'impact'. By integrating Bernstein's sociological framework with Bandura's concept of self-efficacy, the research illuminates recontextualization as a negotiated act rather than a passive transmission or inevitable failure. While Bernstein's framework maps the structural determinants (evaluative rules) that regulate pedagogic discourse, Bandura's construct explains the variation in individual response. Crucially, the study identifies domain-specific self-efficacy not merely as a trait, but as a context-sensitive filter that regulates the 'gap' between PETE ideals and school realities. Bernstein's framework explains how institutional forces, material conditions, and evaluative rules regulate the reproduction of pedagogic discourse, while Bandura's concept of self-efficacy explains variation in individual teachers' responses to these structural constraints. Self-efficacy thus operates as an agentic mechanism linking structural contexts to individual practice.

Methodologically, the study illustrates the value of mixed-methods approaches in educational research. Quantitative data established the distribution and correlates of self-efficacy beliefs across a substantial sample, while qualitative data provided rich accounts of how self-efficacy intersects with recontextualization processes in specific contexts. The explanatory sequential design enabled qualitative findings to illuminate patterns identified quantitatively, enhancing interpretive depth.

Limitations and future research

Several limitations warrant acknowledgement. First, the study relied on self-reported data for both quantitative (self-efficacy) and qualitative (descriptions of practice) components. Observational research examining actual pedagogic practices would provide complementary evidence and potentially reveal discrepancies between reported and enacted practices. Second, the cross-sectional design precludes causal inferences regarding relationships between self-efficacy and pedagogic practice. Longitudinal research tracking beginning teachers across multiple years could illuminate trajectories of efficacy development and identify critical periods for intervention. Third, while the study examined beginning teachers' perspectives on PETE preparation, it did not directly investigate PETE curricula or pedagogies. Comparative research examining recontextualization processes across different PETE institutions with varied curricular approaches would enhance understanding of how PETE design influences beginning teacher practice.

Future research should investigate interventions designed to strengthen beginning teachers' self-efficacy for implementing SRL strategies in ball games contexts. Experimental or quasi-experimental designs comparing induction programmes incorporating targeted professional development with standard induction could provide evidence for effective support structures. Additionally, research examining pupils' experiences of SRL-oriented ball games pedagogy would illuminate the learner perspective currently absent from this investigation. Finally, extending this line of inquiry to other PE content domains (e.g., dance, outdoor education, fitness) would clarify whether patterns identified in ball games teaching generalise or whether domain-specific factors shape recontextualization processes differently across PE curriculum areas.

CONCLUSION

This investigation underscores the complexity of transforming physical education practice. While beginning teachers in Italy demonstrate a theoretical alignment with educational ball games pedagogies, the translation of this discourse into practice is selectively regulated by institutional constraints and mediated by self-efficacy beliefs. Self-efficacy functioned as a mediating mechanism influencing persistence in implementing innovative practices when confronted with evaluative rules and contextual constraints.

The integration of Bernstein's theory of the pedagogic device with Bandura's concept of self-efficacy provided a robust theoretical framework for understanding recontextualization as shaped by both structural-institutional forces and individual agency. Evaluative rules operating within schools—including material conditions, assessment regimes, and pupil expectations—exerted powerful regulatory effects on pedagogic practice, frequently overriding the innovative pedagogic discourse promoted in PETE. However, beginning teachers with higher domain-specific self-efficacy demonstrated greater capacity to navigate these constraints and sustain educationally oriented ball games pedagogies.

The study contributes empirical evidence regarding the impact of PETE on professional practice, a question that has proven persistently challenging across decades of research (Backman et al., 2023; Schempp & Graber, 1992). While PETE successfully transmits educational perspectives that distinguish PE from sport club contexts, substantial work remains in preparing beginning teachers for the practical implementation of innovative pedagogies within authentic school constraints. Strengthening instructional alignment between PETE coursework and school-based experiences, explicitly modelling SRL integration within ball games teaching, and providing sustained professional development during the induction period represent promising directions for enhancing PETE's influence on the field of reproduction.

Ultimately, this investigation underscores the complexity of pedagogical change in PE. Transforming established practices requires not only innovative pedagogic discourse within PETE but also attention to the evaluative rules governing school contexts, the development of beginning teachers' self-efficacy through carefully designed preparation and support and sustained critical engagement with the ideological and material conditions shaping the reproduction field. Only through such multifaceted efforts can the promise of educationally oriented, SRL-informed ball games pedagogy be realised in school PE practice.

SUPPORTING AGENCIES

No funding agencies were reported by the author.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

ACKNOWLEDGMENTS

The author wishes to express sincere gratitude to Massimo Lattoniere for his essential technical assistance with survey administration and the logistics of quantitative data collection, and to Alessandra Corvaglia for her support in performing statistical computations and data processing in accordance with the analytical framework established for this study. The author is also grateful to the network of school administrators and professional associations that facilitated access to the research sites and participants. Finally, deep appreciation is extended to the beginning physical education teachers who took part in this investigation, generously sharing their time, professional experiences, and challenges. While technical and logistical support was provided for specific operational tasks, the author confirms that the conceptualization, study design, interpretation of findings, and manuscript preparation were undertaken solely by the author.

REFERENCES

Backman, E., Tidén, A., Wiorek, D., Svanström, F., & Pihl, L. (2021). "Things that are taken from one culture don't necessarily work well in another culture." Investigating epistemological tensions through preservice

- teachers' views on the assessment of a games course in Swedish PETE. *Cogent Education*, 8(1). <https://doi.org/10.1080/2331186X.2021.1940636>
- Backman, E., Tolgfors, B., Nyberg, G., & Quennerstedt, M. (2023). How does physical education teacher education matter? A methodological approach to understanding transitions from PETE to school physical education. *Physical Education and Sport Pedagogy*, 28(4), 411-424. <https://doi.org/10.1080/17408989.2021.1990248>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Information Age Publishing.
- Bernstein, B. (2000). *Pedagogy, symbolic control and identity: Theory, research, critique* (Rev. ed.). Rowman & Littlefield. <https://doi.org/10.5040/9798216401735>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328-352. <https://doi.org/10.1080/14780887.2020.1769238>
- Casey, A., & Kirk, D. (2020). *Models-based practice in physical education*. Routledge. <https://doi.org/10.4324/9780429319259>
- Cereda, F. (2023). Models in physical education: From pedagogical perspective to practice. *tab edizioni*.
- Cereda, F. (2025a). Epistemological beliefs and professional identity formation in exercise science education: A mixed-methods investigation of developmental trajectories. *Quest*, 77(4), 606-624. <https://doi.org/10.1080/00336297.2025.2525116>
- Cereda, F. (2025b). The shifting landscape of kinesiology in Italy: A call for epistemological re-evaluation and curricular reform. *Formazione & Insegnamento*, 23(1), 157-165. https://doi.org/10.7346/fei-XXIII-01-25_19
- Cereda, F. (2026). Physical literacy development through interdisciplinary approaches in educational contexts: A critical review of theory, research, and practice. *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education*. Advance online publication.
- Colella, D. (2017). Teaching styles, motor learning and educational process. *Formazione & Insegnamento*, 14(1 Suppl.), 25-34. Retrieved from [Accessed 2026, 17 March]: <https://ojs.pensamultimedia.it/index.php/siref/article/view/1874>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education*, 24(3), 751-766. <https://doi.org/10.1016/j.tate.2007.02.010>
- Díaz-Cueto, M., Hernández-Álvarez, J. L., & Castejón, F. J. (2010). Teaching games for understanding to in-service physical education teachers: Rewards and barriers regarding the changing model of teaching sport. *Journal of Teaching in Physical Education*, 29(4), 378-398. <https://doi.org/10.1123/jtpe.29.4.378>
- Dignath-van Ewijk, C. (2013). Which components of teacher competence determine whether teachers enhance self-regulated learning? Predicting teachers' self-reported promotion of self-regulated learning by means of teacher beliefs, knowledge, and self-efficacy. *Frontline Learning Research*, 1(2), 83-105. <https://doi.org/10.14786/flr.v4i5.247>
- Dowling, F. (2011). 'Are PE teacher identities fit for postmodern schools or are they clinging to modernist notions of professionalism?' A case study of Norwegian PE teacher students' emerging professional identities. *Sport, Education and Society*, 16(2), 201-222. <https://doi.org/10.1080/13573322.2011.540425>
- Ennis, C. D. (1999). Creating a culturally relevant curriculum for disengaged girls. *Sport, Education and Society*, 4(1), 31-49. <https://doi.org/10.1080/1357332990040103>
- Evans, J. (2014). Equity and inclusion in physical education PLC. *European Physical Education Review*, 20(3), 319-334. <https://doi.org/10.1177/1356336X14524854>
- García-López, L. M., Gutiérrez, D., Sánchez-Mora, D., & Evangelio, C. (2019). Teachers' use of teaching games for understanding in Central Spain. *Physical Education and Sport Pedagogy*, 24(5), 463-477. <https://doi.org/10.1080/17408989.2019.1628931>

- Hardman, K. (2013). Global issues in physical education: Worldwide physical education survey III findings. *International Journal of Physical Education*, 50(3), 15-28. <https://doi.org/10.5771/2747-6073-2013-3-15>
- Harvey, S., & Jarrett, K. (2014). A review of the game-centred approaches to teaching and coaching literature since 2006. *Physical Education and Sport Pedagogy*, 19(3), 278-300. <https://doi.org/10.1080/17408989.2012.754005>
- Harvey, S., & Pill, S. (2016). Comparisons of academic researchers' and physical education teachers' perspectives on the utilization of the Tactical Games Model. *Journal of Teaching in Physical Education*, 35(4), 313-323. <https://doi.org/10.1123/jtpe.2016-0085>
- Kirk, D. (2010). *Physical education futures*. Routledge. <https://doi.org/10.4324/9780203874622>
- Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Büttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating frequency, quality, and consequences for student performance. *Metacognition and Learning*, 5(2), 157-171. <https://doi.org/10.1007/s11409-010-9055-3>
- Larsson, H., & Karlefors, I. (2015). Physical education cultures in Sweden: Fitness, sports, dancing ... learning? *Sport, Education and Society*, 20(5), 573-587. <https://doi.org/10.1080/13573322.2014.979143>
- Larsson, H., & Quennerstedt, M. (2016). Same, same but different: (Re)understanding the place of context in physical education practice. *Recherches & Éducatives*, 15, 69-86. <https://doi.org/10.4000/rechercheseducations.3067>
- Lundvall, S., & Meckbach, J. (2008). Mind the gap: Physical education and health and the frame factor theory as a tool for analysing educational settings. *Physical Education and Sport Pedagogy*, 13(4), 345-364. <https://doi.org/10.1080/17408980802353362>
- MacPhail, A., Kirk, D., & Griffin, L. (2008). Throwing and catching as relational skills in game play: Situated learning in a modified game unit. *Journal of Teaching in Physical Education*, 27(1), 100-115. <https://doi.org/10.1123/jtpe.27.1.100>
- MacPhail, A., Tannehill, D., Leirhaug, P. E., & Borghouts, L. (2023). Promoting instructional alignment in physical education teacher education. *Physical Education and Sport Pedagogy*, 28(2), 153-164. <https://doi.org/10.1080/17408989.2021.1958177>
- Mitchell, S. A., Oslin, J. L., & Griffin, L. L. (2020). *Teaching sport concepts and skills: A tactical games approach* (4th ed.). Human Kinetics.
- Ministry of Education, University and Research. (2012). Regulation on the national curriculum guidelines for pre-primary school and the first cycle of education (Ministerial Decree No. 254 of 16 November 2012). *Official Gazette of the Italian Republic, General Series*, No. 30, 5 February 2013. Retrieved from [Accessed 2026, 17 March]: <https://www.gazzettaufficiale.it/eli/gu/2013/02/05/30/sq/pdf>
- Ministry of Education and Merit. (2025). Regulation on the national curriculum guidelines for pre-primary school and the first cycle of education (Ministerial Decree No. 221 of 9 December 2025). *Official Gazette of the Italian Republic, General Series*, No. 21, 27 January 2026. Retrieved from [Accessed 2026, 17 March]: <https://www.gazzettaufficiale.it/eli/gu/2026/01/27/21/sq/pdf>
- Mordal-Moen, K., & Green, K. (2014). Physical education teacher education in Norway: the perceptions of student teachers. *Sport, Education and Society*, 19(6), 806-823. <https://doi.org/10.1080/13573322.2012.719867>
- Mustell, J. (2024). Beginning teachers' descriptions of ball games as pedagogic practice in Swedish physical education. *European Physical Education Review*, 31(3), 426-443. <https://doi.org/10.1177/1356336X241280841>
- O'Leary, N. (2014). Learning informally to use teaching games for understanding: The experiences of a recently qualified teacher: The experiences of a recently qualified teacher. *European Physical Education Review*, 20(3), 367-384. <https://doi.org/10.1177/1356336X14534359>
- OECD. (2020). *What students learn matters: Towards a 21st century curriculum*. OECD Publishing. <https://doi.org/10.1787/d86d4d9a-en>
- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Sage.

- Peeters, J., De Backer, F., Reina, V. R., Kindekens, A., Buffel, T., & Lombaerts, K. (2014). The role of teachers' self-regulatory capacities in the implementation of self-regulated learning practices. *Procedia - Social and Behavioral Sciences*, 116, 1963-1970. <https://doi.org/10.1016/j.sbspro.2014.01.504>
- Schempp, P. G., & Graber, K. C. (1992). Teacher socialization from a dialectical perspective: Pretraining through induction. *Journal of Teaching in Physical Education*, 11(4), 329-348. <https://doi.org/10.1123/jtpe.11.4.329>
- Schempp, P. G., Sparkes, A. C., & Templin, T. J. (1993). The micropolitics of teacher induction. *American Educational Research Journal*, 30(3), 447-472. <https://doi.org/10.3102/00028312030003447>
- Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36(1-2), 111-139. <https://doi.org/10.1007/s11165-005-3917-8>
- Scierri, I. D. M. (2024). Italian validation of the Teacher Self-Efficacy scale to implement Self-Regulated Learning (AI-AA). *Italian Journal of Educational Research*, 32, 32-46. <https://doi.org/10.7346/sird-012024-p32>
- Singh, P. (2002). Pedagogising knowledge: Bernstein's theory of the pedagogic device. *British Journal of Sociology of Education*, 23(4), 571-582. <https://doi.org/10.1080/0142569022000038422>
- SISMeS (2024). L'Educazione fisica in Italia: Scenari, sfide, prospettive [Physical education in Italy: Scenarios, challenges, perspectives]. Calzetti & Mariucci.
- Standal, Ø. (2016). *Phenomenology and pedagogy in physical education*. Routledge.
- Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. *Contemporary Educational Psychology*, 66, 101976. <https://doi.org/10.1016/j.cedpsych.2021.101976>
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783-805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Valentini, N. C., & Rudisill, M. E. (2004). Effectiveness of an inclusive mastery climate intervention on the motor skill development of children with and without disabilities. *Adapted Physical Activity Quarterly*, 21(4), 330-347. <https://doi.org/10.1123/apaq.21.4.330>
- Woods, A. M., Gentry, C., & Graber, K. C. (2016). Research on physical education teachers' career stages and socialization. In K. A. R. Richards & K. L. Gaudreault (Eds.), *Teacher socialization in physical education: New perspectives* (pp. 81-97). Routledge.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2



Effect of hurdle height on neuromuscular activity during bilateral jumps in female athletes and physical education students

-  **Emmanuel Nicolás Ignacio Morales**  . University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.
-  **Dario Fernando Cappa**. University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.
-  **Ezequiel Aquistapace**. University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.
-  **Leandro Nodari**. University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.
-  **Marco Ramos**. University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.

ABSTRACT

This study aimed to evaluate the neuromuscular performance of hurdle jumps at different obstacle heights in women to identify the hurdle height eliciting the highest neuromuscular demands. Participants included two groups of regional-level athletes, field hockey players (HOC) and track and field athletes (TAF), and a group of physical education students (PE). Assessments consisted of the countermovement jump with arm swing (CMJA) and hurdle jumps at 60%, 80%, and 100% of CMJA height performed on a force platform. Electromyographic (EMG) activity of the medial gastrocnemius (GM) and tibialis anterior (TA) was recorded, and force variables were analyzed. Rate of force development (RFD) and vertical stiffness (Kvert) were calculated. Results indicated that RFD was significantly greater at 100% of CMJA height, regardless of the group analyzed. EMG activity of the GM was also higher at 100% CMJA height, while no significant differences were observed for the TA. Kvert was not sensitive to hurdle height. In conclusion, 100% of CMJA height is the hurdle height eliciting the greatest neuromuscular and force–time demands in women, independent of sport background.

Keywords: Physical education, Plyometrics, Hurdle jumps, Neuromuscular activity, Force–time curve, Stretch–shortening cycle.

Cite this article as:

Morales, E. N. I., Cappa, D. F., Aquistapace, E., Nodari, L., & Ramos, M. (2026). Effect of hurdle height on neuromuscular activity during bilateral jumps in female athletes and physical education students. *Scientific Journal of Sport and Performance*, 5(3), 482-492. <https://doi.org/10.55860/KEMR9640>

 **Corresponding author.** University Institute of Physical Activity and Sport «Dr. Jorge E. Coll». Mendoza, Argentina.

E-mail: enig_mor_91@hotmail.com

Submitted for publication March 11, 2026.

Accepted for publication April 06, 2026.

Published May 07, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/KEMR9640>

INTRODUCTION

Plyometric training is of vital importance to optimally enhance overall muscle power and specific competitive actions. Jumps, throws, strikes, and sprints are considered short stretch-shortening cycle sports actions (Schmidtbleicher, 1992). These exercises are characterized by the development of high velocity at the end of the concentric phase of the movement (Newton, Robert U. ; McEvoy, 1994).

Loturco et al. (Loturco et al., 2023) and Healy et al. (Healy et al., 2021) surveyed sports coaches and found that 89% and 93%, respectively, employed hurdle jumps, positioning this exercise as one of the most important in the plyometric method to improve sprint power. Hurdles are considered an obstacle, and their height is thus an important variable that influences the intensity of the exercise.

Several studies have described the neuromuscular characteristics of two-legged hurdle jumps (Cappa & Behm, 2013; Healy et al., 2020; Janikov et al., 2023; Kossow & Ebben, 2018). However, many of these studies used male samples. Heick et al. (Heick et al., 2020) used a mixed sample but did not identify sex-based differences. Kossow et al. (Kossow & Ebben, 2018) suggested there are no differences between men and women when analysing single hurdle jumps at fixed heights of 45 cm. However, in this research, the vertical and horizontal strength levels observed were very low, suggesting that the jumps were not continuous, as they are typically performed in training.

Regarding hurdle height, no studies have analysed multiple obstacle heights in women. A previous study on hurdle jumps investigated men using obstacles at 100%, 120%, 140%, and 160% of the maximum height reached in a countermovement jump with arm swing (CMJA). The rate of force development showed that the optimal height was 100%, with higher heights not adding intensity. This was evident when analysing two key variables: contact time increased by 18% when hurdle height rose from 100% to 160% (from 177 to 209 ms, respectively), while force decreased by 6.9% (from 4305 to 4007 N) (Cappa & Behm, 2013). This indicates that excessively increasing hurdle height above the maximum CMJA performance does not enhance the intensity of the motor action.

During plyometric activities, the lower limbs behave like a spring (Farley et al., 1991). Depending on the movement type and intensity used, the neuromuscular system adjusts limb stiffness to increase power. For instance, Farley found that greater execution speed increases muscle stiffness. High execution speed also ensures short ground contact times (Farley & Morgenroth, 1999).

To achieve high levels of stiffness, ranges of motion with small flexion-extension degrees at the hip, knee, and ankle should be used, as shown in sprint analysis at different speeds (Kuitunen et al., 2002). This facilitates the proper use of elastic energy stored during the contact phase (Wilson et al., 1991). Kuitunen demonstrated that vertical stiffness increases significantly (+40%) when sprint speed rises from 70% to 90% of maximum speed.

Arampatzis (Arampatzis et al., 2006) analysed drop jumps from various heights and confirmed that maximum power is achieved with high lower limb stiffness. In this study, athletes were asked to voluntarily modify ground contact times, decreasing and increasing them twice to create five groups. Leg stiffness increased when ground contact time was reduced but worsened when it was increased. However, maximum power at all drop heights was observed when athletes gave their best effort. This suggests there is an optimal stiffness level for the lower limbs to maximise mechanical power, which is self-regulated and decreases when modified voluntarily.

In this sense, when hurdle height exceeds maximum CMJA performance, the subject must increase lower limb flexion, which conditions jump execution by raising the centre of gravity. Therefore, hurdle height plays a critical role in generating stiffness.

Although recommendations for determining hurdle height exist, these are based on male data. In highly explosive actions such as sprinting, women exhibit slightly longer contact times than men (~+7%). Consequently, it is possible that hurdle heights for jumps may differ from those recommended for men. Therefore, the objective of this study was to analyse neuromuscular performance at different hurdle heights in active women.

METHODS

Thirty-eight female participants categorized into three distinct groups based on their training background: Physical Education students (PE, $n = 14$, age 23.67 ± 2.67 , mass 60.98 ± 9.75 kg, height 159.17 ± 5.42 cm and BMI 24.01 ± 3.28), Field Hockey (HOC, $n = 12$, age 22.42 ± 1.77 , mass 61.38 ± 8.91 kg, height 160.84 ± 6.39 cm and BMI 23.70 ± 2.94) and Track and Field (TAF, $n = 12$, age 21.18 ± 2.97 , mass 60.07 ± 9.10 kg, height 162.64 ± 6.80 cm and BMI 22.65 ± 2.60) athletes. Sports volunteers had a minimum of 3 years of sports experience and a strength training frequency of at least 3 times per week for the past year. Inclusion criteria for non-sporting subjects were being physically active and not participating in competitive sports. All subjects were familiar with plyometric and explosive jump training. The exclusion criterion was having disabling injuries in the last 6 months or at the time of testing. Written informed consent was obtained from the subjects before study participation. This study was conducted in accordance with the principles of the Declaration of Helsinki (World Medical Association, 2013). This study was approved by the University Institute of Physical Education from Mendoza (Argentina) committee of research.

Procedures

Vertical jump testing

The warm-up consisted of 4 minutes on a bicycle ergometer (50 watts of power), lumbar and abdominal hyperextension exercises (3 sets of 10 repetitions), a series of squats of 10 repetitions with 20 kg of overload, sub maximum vertical jump without arm swing (2 sets of 5 repetitions), sub maximum rebounds (2 sets of 10 repetitions), maximum rebounds (2 sets of 10 repetitions) and hurdle jumps at a height between 10 and 20 cm (2 sets of 4 repetitions). For the evaluation of the vertical jump, the subjects were placed on 2 force platforms (one for each foot) brand PASCO Scientific 2-axis PS-2142, with sampling frequency of 1000 Hz. They were then asked to perform a maximal vertical jump with swing of arms (CMJA). The height reached was calculated using the formula: $\text{CMJA height (cm)} = (t^2 \cdot g) / 8$, where t is flight time and g is acceleration due the gravity. Two maximum attempts were made, and the highest value was selected. Then, was calculated 60 and 80 percent of the height of CMJA to use with hurdles jump evaluation.

Hurdle jumping testing

After the CMJA evaluation, the subjects were instructed on the hurdles test. Four hurdles were placed on a slide, aligned and separated from each other according to the participants' preference. The force platforms were located between the 2nd and 3rd hurdles as in Cappa and Behm (Cappa & Behm, 2011) (see Figure 1). During both jumps, neuromuscular activation was collected through EMG data using a Wireless electromyograph (biosignalsplux Explorer Biosignals S.A.). Electrodes (Medi-Trace MT200 Kendall™, Technical products, Toronto, Ontario, Canada) were placed on the GM and TA of the subjects' right leg according to SENIAM guidelines (Stegeman & Hermens, 2007). Raw data collection was performed with OpenSignals software (version 2.0.0). Processing of this data was performed with MatLab R2018b (The

MathWorks, Inc.) and a band-pass Butterworth filter (20–450 Hz, 6th order), with a 50-Hz notch filter applied to remove power line noise. Subjects were instructed to perform the jumps as quickly as possible, minimizing contact time on the platforms. No technique was specified to overcome the fences. Subjects were only verbally encouraged to take the test as quickly as possible. Two attempts were performed with a 1-minute rest between each to avoid fatigue. Contact and take-off with the force platforms were strictly monitored, and if in any attempt a delay in the shortening-stretching cycle was observed or if the subjects fell off the platforms, the attempt was repeated. If a difference in performance of more than 10% was found, a third attempt was made. The EMG signal was analysed at 2 moments: during pre-activation, 100 milliseconds before ground contact (PRE) and during ground contact time (CON). The protocol was repeated at 60, 80 and 100% of the maximum height of the CMJA, in random order between trials for each hurdle height. The selected hurdle heights were chosen to represent progressively increasing relative intensities commonly used in plyometric training. Heights below 60% of CMJA were not considered, as pilot observations and previous literature suggest that such low obstacles do not elicit sufficient neuromuscular demand to differentiate force–time or EMG responses from low-intensity rebound jumps.



Figure 1. Hurdles jump testing.

Statistical analysis

During the CMJA test and Hurdles jumps, the following raw data were collected: peak vertical force (PF, N), average vertical force (AF, N), contact time (CT, s), and flight time (FT, ms). Derived variables included: normalized peak vertical (nPF, N·N⁻¹), normalized average vertical force (nMF, N·N⁻¹), rate of force development (RFD, N·s⁻¹), normalized rate of force development (nRFD, N·s⁻¹·N⁻¹), mechanical power (POT, W), normalized power (nPOW, W·N⁻¹) and vertical stiffness (Kvert, kN·m⁻¹). Mechanical power was computed using the Sayers equation (Sayers et al., 1999). Normalized variables were calculated using the body weight. Peak rate of force development (RFD) was computed as the maximal slope of the force–time curve during ground contact ($\Delta F/\Delta t$), expressed in N·s⁻¹. Normalized RFD (nRFD) was obtained by dividing RFD by body mass (N·s⁻¹·kg⁻¹). Kvert was estimated using the spring–mass model: $K_{vert} = m(2\pi f)^2$, where m is body mass and f is the natural frequency of oscillation estimated from contact/flight times (Mc Mahon, Thomas, Cheng, 1990). EMG data collected during hurdle jumps were normalized to the corresponding CMJA EMG activity (set at 100%). EMG signals obtained during hurdle jumps were normalized to the corresponding CMJA EMG activity. This approach was selected because CMJA represents a maximal dynamic task with similar neuromuscular and stretch–shortening cycle characteristics to hurdle jumps.

Statistical treatment

Statistical analyses were performed using JASP v0.19 and MS Excel 365. Normality was checked with Shapiro–Wilk. Normally distributed variables were analysed with one-way ANOVA (between groups) and repeated-measures ANOVA (within subjects); Bonferroni adjustments were used for post-hoc pairwise comparisons. Non-normal variables were tested with Kruskal–Wallis (with Dunn post-hoc) or Mann–Whitney U as appropriate. Effect sizes are reported as eta-squared (η^2) for ANOVA and Cohen's d for pairwise contrasts; for non-parametric tests report eta-squared (η^2) or Kendall's W. The level of significance was set at $*p < .05$; $**p < .01$; $***p < .001$.

RESULTS

A Kruskal-Wallis test revealed a significant difference in age between the PE and TAF groups ($H = 6.578$, rank $\eta^2 = 0.131$, 95% CI: 0.000, 0.512, $p = .037$). Post-hoc analysis using the Mann-Whitney U test indicated that participants in the PE group were significantly older than those in the TAF group ($z = 2.56$, $p = .031$). No other significant differences were found between the groups.

Table 1 shows the mean ($\pm$ SD) of the variables of the force-time curve of the CMJA test for each group. The statistical analysis revealed significant differences in CMJA height ($F = 15.40$, $\eta^2 = 0.468$, 95% CI: 0.209, 0.635, $p < .001$). Post-hoc analysis indicated that the TAF group had significantly higher CMJA height compared to both the PE ($t = -5.496$, $d = -2.162$, 95% CI: -3.345, -0.978, $p_{\text{bonf}} < .001$) and HOC ($t = -3.517$, $d = -1.436$, 95% CI: -2.549, -0.322, $p_{\text{bonf}} = .004$) groups.

Average force normalized by body weight also differed significantly between groups ($F = 5.328$, $\eta^2 = 0.239$, 95% CI: 0.019, 0.446, $p = .01$). Post-hoc tests showed that the HOC group had significantly higher values than the TAF group ($t = 3.157$, 95% CI: 0.021, 0.19, $p_{\text{bonf}} = .01$).

Power normalized by body weight significantly differed among groups ($H = 15.49$, rank $\eta^2 = 0.397$, 95% CI: 0.236, 0.618, $p < .001$). Post-hoc Mann-Whitney U tests indicated that the TAF group had higher values compared to both PE ($z = -3.927$ [$W_j = 11.786$, $W_i = 28.909$], $p_{\text{bonf}} < .001$) and HOC ($z = -2.341$ [$W_j = 18.333$, $W_i = 28.909$], $p = .019$) groups.

No other significant differences were found between the groups. A significant interaction between hurdle height and force-time curve variables was observed when hurdles were set at 100% vertical jump height (As shown in Table 2).

Peak vertical force was significantly higher at 100% hurdle height compared to both 60 and 80% ($X^2 = 10.865$, $p = .004$). ANOVA post hoc analysis showed that 100% > 60% (T-stat = 3.352, [$W_i = 63.00$, $W_j = 90.00$], $p_{\text{bonf}} = .004$) and 100% > 80% (T-stat = 2.607, [$W_i = 69.00$, $W_j = 90.00$], $p_{\text{bonf}} = .033$). Similarly, normalized peak vertical force was significantly greater at 100% ($F = 7.813$, $p < .001$, $\eta^2 = 0.187$) compared to 60% ($t = -3.901$, $d = -0.749$, 95% CI: [-0.774, -0.167], $p_{\text{bonf}} = .001$).

Mean vertical force also showed significant differences ($F = 6.970$, $p = .002$, $\eta^2 = 0.162$), with 100% exceeding both 60% ($t = -3.045$, $d = -0.416$, 95% CI: [-171.191, -16.463], $p_{\text{bonf}} = 0.013$) and 80% ($t = -3.440$, $d = -0.381$, 95% CI: [-148.645, -23.203], $p_{\text{bonf}} = 0.004$). Consistent patterns were observed for normalized mean vertical force ($F = 5.961$, $p = .004$, $\eta^2 = 0.139$) where 100% was greater than 60% ($t = -3.084$, $d = -0.450$, 95% CI: [-0.294, -0.030], $p_{\text{bonf}} = 0.012$) and 80% ($t = -2.720$, $d = -0.363$, 95% CI: [-0.252, -0.010], $p_{\text{bonf}} = 0.03$).

Rate of force development was significantly higher at 100% ($X^2 = 10.343$, $p = .006$) compared to both 60% (T-stat = 2.425, [Wi = 66.00, Wj = 85.00], $p_{bonf} = .036$) and 80% (T-stat = 3.318, [Wi = 59.00, Wj = 85.00], $p_{bonf} = .004$) with similar results for normalized RFD ($X^2 = 10.343$, $p = .006$) where 100% was over 60% (T-stat = 2.425, [Wi = 66.00, Wj = 85.00], $p = .018$) and 80% (T-stat = 3.318, [Wi = 59.00, Wj = 85.00], $p_{bonf} = .004$). Vertical stiffness (K_{vert}) did not differ significantly across hurdle heights ($p > .05$).

Table 1. CMJA variables (Mean $\pm$ SD).

	PE	HOC	TAF
CMJA height (cm)	27.71 $\pm$ 3.3	30.75 $\pm$ 4.7	36.75 $\pm$ 4.5 ^{a*** b**}
Contact time (s)	0.94 $\pm$ 0.12	0.93 $\pm$ 0.18	1.00 $\pm$ 0.14
Average force (N)	731 $\pm$ 89	750 $\pm$ 80	718.4 $\pm$ 97
Normalized average force (N·N ⁻¹)	1.27 $\pm$ 0.07	1.30 $\pm$ 0.07 ^{c*}	1.19 $\pm$ 0.10
Peak force (N)	1348 $\pm$ 214	1431 $\pm$ 236	1403 $\pm$ 257
RFD (N·s ⁻¹)	1686 $\pm$ 497	1805 $\pm$ 582	1633 $\pm$ 384
Normalized RFD (N·s ⁻¹ ·N ⁻¹)	2.80 $\pm$ 0.60	3.13 $\pm$ 1.03	2.83 $\pm$ 0.88
K_{vert} (N·m ⁻¹)	2.41 $\pm$ 0.89	2.60 $\pm$ 0.93	2.47 $\pm$ 0.98
Power (W)	2419 $\pm$ 410	2599 $\pm$ 553	2803 $\pm$ 459 ^{d*** e*}
Normalized power (W·N ⁻¹)	4.05 $\pm$ 0.22	4.29 $\pm$ 0.45	4.69 $\pm$ 0.27

Note. CMJA = countermovement jump with arms swing. TAF = Track and Field; HOC = Field Hockey; PE = Physical Education. a: TAF vs PE; b: TAF vs HOC; c: HOC vs TAF; d: TAF vs PE; e: TAF vs HOC. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 2. Analysis of hurdles jump (Mean $\pm$ SD).

	Hurdles heights as a percentage of maximum CMJA		
	60%	80%	100%
Contact Time (s)	0.164 (0.02)	0.166 (0.02)	0.165 (0.02)
Peak Vertical Force (N)	3122 (459)	3101 (467)	3335 (364) ^{a** b*}
Normalized Peak Vertical Force (N·N ⁻¹)	5.15 (0.55)	5.34 (0.83)	5.73 (0.68) ^{c**}
Mean Vertical Force (N)	1654 (262)	1645 (233)	1740 (203) ^{d**}
Normalized Mean Vertical Force (N·N ⁻¹)	2.79 (0.34)	2.82 (0.37)	2.95 (0.37) ^{e*}
RFD (N·s ⁻¹)	51652 (15773)	51090 (18496)	56259 (13356) ^{f** g*}
Normalized RFD (N·s ⁻¹ ·N ⁻¹)	87.37 (27.63)	87.18 (33.01)	95.57 (24.32) ^{h* i***}
K_{vert} (N·m ⁻¹)	38.37 (8.29)	39.42 (13.76)	36.08 (9.03)

Note. RFD = Rate Force Development. K_{vert} = Vertical Stiffness. a: 100% > 60%; b: 100% > 80%; c: 100% > 60%; d: 100% > 60 and 80%; e: 100% > 60 and 80%; f: 100% > 60%; g: 100% > 80%; h: 100% > 60%; i: 100% > 80%. * $p < .05$; ** $p < .01$; *** $p < .001$.

No statistical differences were found in force-time variables when comparing between groups, nor when comparing athletes and students ($p > .05$).

A significant difference was observed between CMJA and hurdle jumps in the PRE activity of the TA ($p < .001$) and CON activity of the GM ($p < .001$) for all heights. No statistically significant differences were found between CMJA and hurdle jumps in the CON activity of the TA or the PRE activity of the GM.

Repeated measures ANOVA (see Figure 2) revealed significant differences between hurdle heights for the pre-contact activity of the GM ($\chi^2 = 9.842$, $p = .02$) and the ground contact activity of the GM ($\chi^2 = 65.612$, $p < .001$).

Specifically, for the PRE activity of the GM, the 100% hurdle height condition showed significantly greater activity compared to the 60% condition ($T\text{-stat} = 2.928$, $W_i = 73.0$, $W_j = 103.5$; $p_{\text{Bonf}} = .025$). For the CON activity of the GM, the 100% hurdle height condition showed significantly greater activity compared to both the 60% condition ($T\text{-stat} = 4.183$, $W_i = 90.0$, $W_j = 117.0$; $p_{\text{Bonf}} < .001$) and the 80% condition ($T\text{-stat} = 2.944$, $W_i = 98.0$, $W_j = 117.0$; $p_{\text{Bonf}} = .024$).

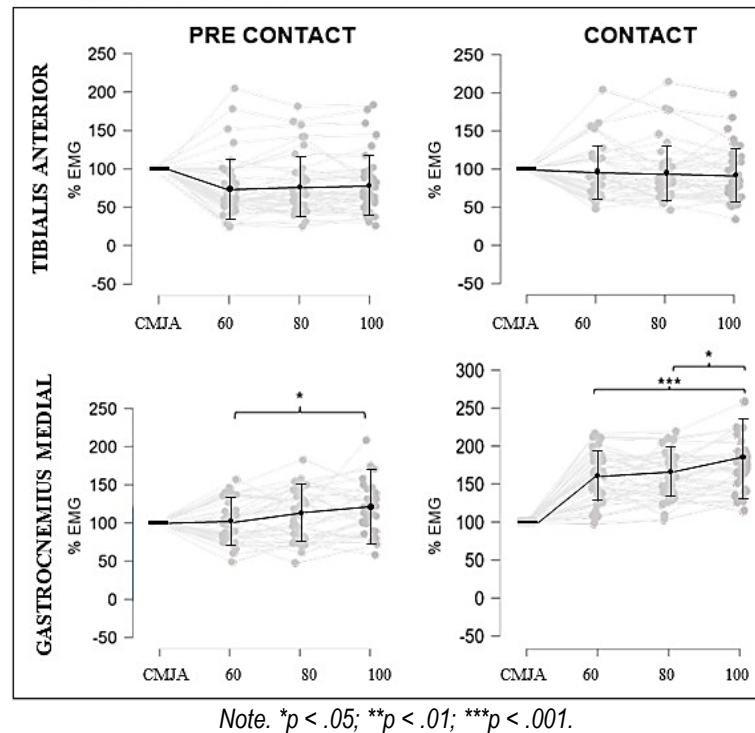


Figure 2. EMG of hurdle jumps.

Analyses of EMG by groups revealed a significant difference in the contact phase of GM at 100% (Kruskal-Wallis = 8.051, $p = .018$, $\eta^2 = 0.178$). Post hoc pairwise comparisons indicated that the TAF subgroup exhibited a significantly higher activation of GM in contact phase than the HOC group ($z = -2.803$, $p = .015$). No other significant differences were found in the analysis.

DISCUSSION

The manipulation of training intensity is fundamental for programming and undulating workload distribution. In hurdle jumps, hurdle height is a key variable to establish exercise intensity (Cappa & Behm, 2013; Jensen & Ebben, 2007). The main finding of this study was that rate of force development (RFD) was significantly greater at 100% of CMJA height compared with 60% and 80%, when analysing the overall sample. However, when groups were analysed separately, no significant differences were found. This is likely explained by the fact that all groups maintained low ground contact times as hurdle height increased, while relative force production increased slightly. Notably, at 100% CMJA height, all groups exhibited the highest relative force and the lowest contact times (Table 1). Thus, RFD at 100% CMJA height was consistently greater, although in some cases it did not reach statistical significance.

Although hurdle jumps are widely used by coaches to develop lower-limb power (Healy et al., 2021), the most appropriate method to determine training intensity and movement specificity remains unclear. In male

athletes, Cappa and Behm (Cappa & Behm, 2011) reported that the optimal hurdle height corresponded to the maximum height achieved in the countermovement jump with arm swing (CMJA). Higher hurdle heights required increased flexion to clear the obstacle, which raised the centre of mass, led to longer ground contact times, and reduced peak force. Consequently, RFD decreased when hurdle height exceeded 100% CMJA. For example, RFD decreased by approximately 21% when comparing 100% with 160% CMJA (40,981 vs. 32,153 N·s⁻¹).

In contrast, the present study analysed hurdle heights below maximal CMJA performance. Our results showed that reducing hurdle height below 100% CMJA led to significant decreases in RFD both in pooled analyses and when groups were compared independently (Figure 2, Table 1).

Regarding muscle stiffness (Kvert), no significant differences were observed across hurdle heights. This result was unexpected, given that increases in RFD are typically associated with increases in stiffness. Arampatzis et al. (Arampatzis et al., 2001), however, demonstrated in drop jumps that increases in stiffness may occur simultaneously with decreases in power output. In their study, when athletes were instructed to minimize ground contact time, stiffness remained elevated but power decreased. These findings suggest that stiffness adaptations may not always translate into enhanced performance. Furthermore, Kvert has limitations as a performance indicator, particularly for horizontal displacements (Maloney & Fletcher, 2021), and it does not account for the summation of joint-specific stiffness.

The calculation method also influences stiffness outcomes. Kvert, based on a spring-mass model (Mc Mahon, Thomas, Cheng, 1990), reflects whole-body stiffness relative to centre-of-mass displacement. However, ankle and knee joints may adopt distinct strategies to modulate stiffness during hurdle jumps. Farley and Morgenroth (Farley & Morgenroth, 1999) reported that ankle modulation is the primary mechanism for adjusting stiffness in vertical hopping, while knee contribution is more limited. When an external obstacle requires additional flexion to clear, stiffness regulation may differ from vertical hopping strategies.

It should be noted that Kvert, as derived from the spring-mass model, represents a global measure of lower-limb stiffness and does not capture joint-specific adaptations. During hurdle jumps, ankle and knee joints may adopt different stiffness regulation strategies depending on obstacle height, which may not be reflected in whole-body Kvert values. Consequently, the lack of sensitivity observed for Kvert across hurdle heights may be partly explained by methodological limitations inherent to this model.

Peak force values during hurdle jumps were significantly higher than during CMJA (3,400 vs. 1,400 N, respectively), likely due to the high pre activation and force requirements of hurdle clearance. Similar findings have been reported in box jumps in female handball players (Koefoed et al., 2022), where increasing box height above ~70% of CMJA did not significantly increase power output, as athletes compensated primarily by flexing the legs rather than producing higher propulsive forces.

From a training perspective, while maximal strength development is typically prescribed at 80–100% of 1RM in exercises such as squats, ballistic multi-jump actions such as hurdle jumps provide a more physiologically maximal expression of force. For instance, Perttunen et al. (Perttunen et al., 2000) reported peak forces exceeding 10,000 N in male triple jumpers, values that far surpass those observed in closed-chain resistance training. Although sex differences exist, female athletes achieve ~85% of male world record performance in this discipline, reinforcing the high neuromuscular demands of jumping-based exercises. In the present study, hurdle jumps produced forces that were double or even triple those observed in CMJA.

Electromyographic analyses revealed that GM activity increased with hurdle height, particularly during ground contact (+10%) and concentric phase (+24%), coinciding with an 8% increase in concentric force. TA activity remained relatively constant across hurdle heights, consistent with prior reports (Cappa & Behm, 2013). When analysed by training background, no group differences were found, suggesting that 100% CMJA hurdle height optimizes neuromuscular output independently of sport specialization.

These findings have practical implications. Many training programs employ fixed commercial hurdle heights (15–20 cm) for group sessions. However, such standardization may underestimate or misrepresent the optimal intensity for individual athletes. Using individualized hurdle heights based on CMJA performance ensures that athletes train at an intensity that maximizes neuromuscular output.

A limitation of the present study is the modest sample size in subgroup analyses ($n = 12–14$), which may have reduced statistical power to detect interaction effects. Additionally, EMG signals were recorded unilaterally (right leg), and kinematic measures were not collected; future studies should include bilateral EMG and motion capture to elucidate joint-specific strategies. Additionally, vertical stiffness was estimated using a spring–mass model, which does not account for joint-specific stiffness regulation; therefore, potential adaptations at the ankle or knee level may not have been detected.

CONCLUSION

In conclusion, RFD was greatest at hurdle heights equivalent to 100% of CMJA performance. Lower hurdle heights were associated with reduced neuromuscular variables. Vertical stiffness (K_{vert}) was not sensitive to changes in hurdle height, suggesting that alternative methods may be required to assess stiffness during hurdle jumps. Practically, prescribing hurdle height relative to each athlete's CMJA rather than using fixed commercial standards appears to be the eliciting the highest neuromuscular demands strategy for maximizing neuromuscular performance in women.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. Contributions are specified according to the CRediT (Contributor Roles Taxonomy) as follows:

Conceptualisation: Emmanuel N. I. Morales. Methodology: Emmanuel N. I. Morales; Dario F. Cappa. Formal analysis: Emmanuel N. I. Morales. Investigation: Emmanuel N. I. Morales; Dario F. Cappa; Ezequiel Aquistapace; Leandro Nodari; Marco Ramos. Data curation: Emmanuel N. I. Morales; Ezequiel Aquistapace. Writing – original draft: Emmanuel N. I. Morales. Writing – review & editing: Dario F. Cappa; Emmanuel N. I. Morales. Supervision: Dario F. Cappa. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

- Arampatzis, A., De Monte, G., Karamanidis, K., Morey-Klapsing, G., Stafilidis, S., & Brüggemann, G. P. (2006). Influence of the muscle-tendon unit's mechanical and morphological properties on running economy. *Journal of Experimental Biology*, 209(17), 3345-3357. <https://doi.org/10.1242/jeb.02340>
- Arampatzis, A., Schade, F., Walsh, M., & Brüggemann, G. P. (2001). Influence of leg stiffness and its effect on myodynamic jumping performance. *Journal of Electromyography and Kinesiology*, 11(5), 355-364. [https://doi.org/10.1016/S1050-6411\(01\)00009-8](https://doi.org/10.1016/S1050-6411(01)00009-8)
- Cappa, D. F., & Behm, D. G. (2011). Training specificity of hurdle vs. countermovement jump training. *Journal of Strength and Conditioning Research*, 25(10), 2715-2720. <https://doi.org/10.1519/JSC.0b013e318208d43c>
- Cappa, D. F., & Behm, D. G. (2013). Neuromuscular characteristics of drop and hurdle jumps with different types of landings. *Journal of Strength and Conditioning Research*, 27(11), 3011-3020. <https://doi.org/10.1519/JSC.0b013e31828c28b3>
- Farley, C. T., Blickhan, R., Saito, J., & Taylor, C. R. (1991). Hopping frequency in humans: A test of how springs set stride frequency in bouncing gaits. *Journal of Applied Physiology*, 71(6), 2127-2132. <https://doi.org/10.1152/jappl.1991.71.6.2127>
- Farley, C. T., & Morgenroth, D. C. (1999). Leg stiffness primarily depends on ankle stiffness during human hopping. *Journal of Biomechanics*, 32(3), 267-273. [https://doi.org/10.1016/S0021-9290\(98\)00170-5](https://doi.org/10.1016/S0021-9290(98)00170-5)
- Healy, R., Bolger, R., Kenny, I. C., & Harrison, A. J. (2020). Kinematic Differences Between Sprinting and the Hurdle Jump Exercise - a Preliminary Analysis. 38th International Society of Biomechanics in Sport Conference, 520-523.
- Healy, R., Kenny, I. C., & Harrison, A. J. (2021). Resistance training practices of sprint coaches. *Journal of Strength and Conditioning Research*, 35(7), 1939-1948. <https://doi.org/10.1519/JSC.0000000000002992>
- Heick, J. D., Talkington, M., & Jain, T. (2020). Electromyographic Analysis of Gluteal Recruitment: an Exploration of Activation During Jumping Tasks. *International Journal of Sports Physical Therapy*, 15(6), 1019-1028. <https://doi.org/10.26603/ijsp20201019>
- Janikov, M. T., Pádecký, J., Doguet, V., & Tufano, J. J. (2023). Countermovement, Hurdle, and Box Jumps: Data-Driven Exercise Selection. *Journal of Functional Morphology and Kinesiology*, 8(2). <https://doi.org/10.3390/jfmk8020061>
- Jensen, R. L., & Ebben, W. P. (2007). Quantifying plyometric intensity via rate of force development, knee joint, and ground reaction forces. *Journal of Strength and Conditioning Research*, 21(3), 763-767. <https://doi.org/10.1519/R-18735.1>
- Koefoed, N., Dam, S., & Kersting, U. G. (2022). Effect of Box Height on Box Jump Performance in Elite Female Handball Players. *Journal of Strength and Conditioning Research*, 36(2), 508-512. <https://doi.org/10.1519/JSC.0000000000003481>

- Kossow, A. J., & Ebben, W. P. (2018). Kinetic analysis of horizontal plyometric exercise intensity. *Journal of Strength and Conditioning Research*, 32(5), 1222-1229. <https://doi.org/10.1519/JSC.0000000000002096>
- Kuitunen, S., Komi, P. V., & Kyröläinen, H. (2002). Knee and ankle joint stiffness in sprint running. *Medicine and Science in Sports and Exercise*, 34(1), 166-173. <https://doi.org/10.1097/00005768-200201000-00025>
- Loturco, I., Pereira, L. A., Freitas, T. T., Moura, T. B. M. A., Mercer, V. P., Fernandes, V., Moura, N. S. A., Moura, N. A., Zajac, A. A., & Bishop, C. (2023). Plyometric Training Practices of Brazilian Olympic Sprint and Jump Coaches: Toward a Deeper Understanding of Their Choices and Insights. *Journal of Human Kinetics*, 88(July), 131-150. <https://doi.org/10.5114/jhk/169167>
- Maloney, S. J., & Fletcher, I. M. (2021). Lower limb stiffness testing in athletic performance: a critical review. *Sports Biomechanics*, 20(1), 109-130. <https://doi.org/10.1080/14763141.2018.1460395>
- Mc Mahon, Thomas, Cheng, G. (1990). The Mechanics of Running: How does Stiffness Couple With Speed? In *J Biomechanics* (Vol. 23, pp. 65-78). [https://doi.org/10.1016/0021-9290\(90\)90042-2](https://doi.org/10.1016/0021-9290(90)90042-2)
- Newton, Robert U. ; McEvoy, K. P. (1994). Baseball Throwing Velocity: A Comparison of Medicine Ball Training and Weight Training. *Journal of Strength and Conditioning Research*, 8(3), 198-203. [https://doi.org/10.1519/1533-4287\(1994\)008<0198:BTVACO>2.3.CO;2](https://doi.org/10.1519/1533-4287(1994)008<0198:BTVACO>2.3.CO;2)
- Perttunen, J., Kyröläinen, H., Komi, P. V., & Heinonen, A. (2000). Biomechanical loading in the triple jump. *Journal of Sports Sciences*, 18(5), 363-370. <https://doi.org/10.1080/026404100402421>
- Sayers, S. P., Harackiewicz, D. V., Harman, E. A., Frykman, P. N., & Rosenstein, M. T. (1999). Cross-validation of three jump power equations. *Medicine and Science in Sports and Exercise*, 31(4), 572-577. <https://doi.org/10.1097/00005768-199904000-00013>
- Schmidtbleicher, D. (1992). Training for power event. In: Komi PV (ed) *Strength and power in sport*. Blackwell Scientific, London, pp 381-395. Retrieved from: The School Review.
- Wilson, G. J., Elliott, B. C., & Wood, G. A. (1991). The effect on performance of imposing a delay during a stretch-shorten cycle movement. In *Medicine and Science in Sports and Exercise* (Vol. 23, Issue 3, pp. 364-370). <https://doi.org/10.1249/00005768-199103000-00016>
- World Medical Association. (2013). Declaración de Helsinki de la AMM - Principios éticos para las investigaciones médicas en seres humanos. *World Medical Association, Inc*, 81(3), 1-8. <https://doi.org/10.1001/jama.2013.281053>



This work is licensed under a [Attribution-NonCommercial-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0).

Effects of cause-based and effect-based motor instruction on six-minute walk and cooper test performance in adolescent female volleyball athletes



Arianna Fogliata . Department of Mental and Physical Health and Preventive Medicine. University of Campania "Luigi Vanvitelli". Caserta, Italy.

Stefano Tardini. eLAB. University of Italian Switzerland (USI). Lugano, Switzerland.

Ivan Pavic. eLAB. University of Italian Switzerland (USI). Lugano, Switzerland.

Lorenzo Cantoni. Institute of Digital Technologies for Communication. University of Italian Switzerland (USI). Lugano, Switzerland.

ABSTRACT

Background: Instructional strategies influence motor learning during adolescence, but it remains unclear whether instruction focused on biomechanical causes of movement yields different functional outcomes from instruction focused on observable movement effects. **Methods:** A total of 184 adolescent female volleyball athletes (mean age 14.44 ± 1.71 years) were stratified by age category and randomly assigned to a cause-based instruction group or an effect-based instruction group. Participants completed the Six-Minute Walk Test (6MWT) and 12-minute Cooper test before and after a 7–10 day intervention period, during which training load remained unchanged. **Results:** The cause-based group showed a significant pre–post improvement in 6MWT distance. Change-score analysis revealed a greater improvement in the cause-based group than in the effect-based group for the 6MWT ($p < .001$, $r = .223$), whereas no between-group difference emerged for the Cooper test ($p = .134$, $r = .112$). **Conclusions:** Cause-based instruction was associated with a small-to-moderate short-term improvement in submaximal walking performance, with no clear transfer to endurance field performance. These findings suggest that directing attention to biomechanical causes of movement may support functional motor organisation in adolescent athletes.

Keywords: Physical education, Motor learning, Attentional focus, Biomechanics, Sports performance.

Cite this article as:

Fogliata, A., Tardini, S., Pavic, I., & Cantoni, L. (2026). Effects of cause-based and effect-based motor instruction on six-minute walk and cooper test performance in adolescent female volleyball athletes. *Scientific Journal of Sport and Performance*, 5(3), 493-501. <https://doi.org/10.55860/JVST6069>

Corresponding author. Department of Mental and Physical Health and Preventive Medicine. University of Campania "Luigi Vanvitelli". viale Abriamo Lincon, 5, Caserta, Italy.

E-mail: fogliataarianna@gmail.com

Submitted for publication March 25, 2026.

Accepted for publication April 30, 2026.

Published May 07, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/JVST6069>

INTRODUCTION

The quality of the explanation provided by the teacher represents an important factor in motor learning, particularly during adolescence, when movement control processes remain sensitive to instructional approaches (Haibach-Beach et al., 2024). In educational contexts, a common teaching approach is based on the description of the visible effects of movement, whereby learners are guided to observe and reproduce the external form of the gesture (De Stefani et al., 2020; Wulf, 2012). This model, referred to here as effect-based instruction, relies primarily on the reproduction of observable movement features (Bach et al., 2024). Building on these considerations, the OPTIMAL Theory proposed by Wulf and Lewthwaite (2016) suggests that the way instruction directs learners' attention directly influences movement quality, learning stability and transferability across contexts (Song, 2019). Consistent with this perspective, research in motor learning has shown that skill acquisition is enhanced when attention is directed towards deeper and functionally relevant components of movement, such as its dynamics, generative forces and underlying muscular sequences (Pascua et al., 2015; Riemann & Lephart, 2002). This theoretical framework suggests that orienting learners' attention towards the components that generate movement, rather than solely towards its visible effects, may facilitate a more conscious and efficient construction of motor skills. At an educational level, this perspective also finds empirical support. De Oliveira et al. (2016) demonstrated that the quality and structure of information provided to learners influence perceived motor competence and movement organisation (Chambal et al., 2024). This indicates that not only what is taught, but also how motor content is presented, significantly contributes to the quality of learning. Similarly, Riskowski et al. (2013) highlighted that making cause–effect relationships explicit facilitates the identification of functional elements of movement and improves interpretative processes during observation. These considerations are particularly relevant when working with adolescent female athletes (Balyi & Hamilton, 2004; Lloyd et al., 2015).

Literature reports specific developmental trajectories in aerobic capacity, respiratory regulation and body perception in this population, often characterised by a relative slowdown compared with male peers after the onset of puberty (Malina et al., 2004; Armstrong & Welsman, 1997; Haapala et al., 2022). Furthermore, during the initial phase of the competitive season, movement organisation and functional efficiency may be especially sensitive to instructional modalities (Jelonek et al., 2017). In young athletes, physiological adaptations are responsive to variations in training load and to early preparatory phases, during which fluctuations in cardiorespiratory parameters may occur (Baxter-Jones & Maffulli, 2003; Migliaccio et al., 2023). However, it has not yet been investigated whether different instructional approaches, namely cause-based versus effect-based instruction, may influence standardised field-based functional indicators commonly used in physical education settings, such as walking performance tests (Troosters et al., 1999; Bohannon, 2018). Within this framework, growing interest has emerged in instructional approaches centred on the biomechanical generators of movement. In the present study, cause-based instruction was delivered through an experimental system designed to provide motor instruction based on explicit cause–effect relationships through guided visual and verbal explanations.

In light of these considerations, it appears relevant to examine whether an instructional approach centred on the generative biomechanical components of movement can produce measurable functional effects compared with an approach focused solely on observable outcomes. In the context of adolescent physical education, understanding the impact of instructional structure may help clarify the role of explanation in shaping movement organisation and functional performance quality. Therefore, the present study aimed to compare the effects of cause-based instruction and effect-based instruction on field-based functional indicators (Six-Minute Walk Test and Cooper test) in adolescent female athletes during the early phase of the competitive season.

METHODS

A total of 184 adolescent female volleyball athletes (mean age = 14.44 ± 1.71 years) participated in the study. All participants had been engaged in organised competitive training for at least three years within the organised youth volleyball system. They trained approximately four sessions per week (120–150 minutes per session) and followed comparable pre-season training schedules. All athletes were clinically healthy, free from musculoskeletal injuries at the time of testing and during the previous six months and were not taking any medication during data collection. In order to reduce potential hormonal influences, where feasible, participants were not assessed during the luteal phase of their menstrual cycle (Tenan et al., 2014; Sims & Heather, 2018).

The sample consisted exclusively of female adolescents to ensure homogeneity and reduce variability related to sex-based differences in physical development, strength and body composition (Armstrong et al., 2011). Body mass index (BMI) was recorded to verify baseline comparability between groups (overall mean BMI = 22.5 ± 2.5). Participants were allocated to two groups: Group 1 (experimental; $n = 100$; mean age = 14.41 ± 1.75 years), which received cause-based motor instruction, and Group 2 (control; $n = 84$; mean age = 14.48 ± 1.68 years), which received effect-based instruction. Stratified randomisation by age category (e.g., U14, U16) was performed using a computer-generated allocation sequence to ensure balanced distribution across age, training load and competitive level (Kernan et al., 1999).

Assessments were conducted following a two-month summer break, during the initial stage of the pre-season cycle. All athletes followed the same training programme under identical coaching staff and environmental conditions. The only difference between groups was the instructional modality applied during three instructional sessions delivered over a 7–10 day period. Data from four athletes were excluded due to absence during testing ($n = 2$) or injury ($n = 2$). Final analyses included only participants with complete T1 and T2 measurements. Pre-test (T1) assessments were conducted at the resumption of training, prior to the instructional intervention. Post-test (T2) assessments were conducted 7–10 days later, following three differentiated instructional sessions (cause-based vs. effect-based) while maintaining identical physical training loads. In Group 1, cause-based instruction was delivered through an experimental instructional system designed to provide biomechanically grounded explanations of movement through guided visual and verbal models.

The system emphasises generative forces, muscular sequencing and internal movement mechanics, encouraging deeper cognitive and proprioceptive engagement (see Table 1). In Group 2, effect-based instruction was delivered verbally by a single trained instructor using standardised technical cues focused on observable movement characteristics, such as segmental trajectories, limb positioning and rhythm (see Table 2). No explicit reference to internal biomechanical mechanisms was provided. Two validated field-based tests were selected: the Six-Minute Walk Test (6MWT) and the 12-minute Cooper test. The 6MWT was chosen to assess submaximal functional walking capacity and potential changes in functional movement organisation (Li et al., 2005; Geiger et al., 2007).

The Cooper test was used as a broader indicator of endurance field performance and functional endurance (Cooper, 1968). Both tests were administered on the same day under identical environmental conditions. The 6MWT was performed first, followed by the Cooper test after sufficient recovery, defined as the return of heart rate to individual baseline values (Bandyopadhyay, 2015).

Table 1. Operational summary of cause-based instruction on walking (Group 1).

Instructional phase	Delivered content	Perceptual focus	Biomechanical objective
Introduction to walking	Walking is a multisystemic and complex motor action involving muscles, joints, nervous system, and senses.	Global awareness of the body in motion	Understand walking as a coordinated outcome of multiple systems
Focus on movement initiation	The first step does not start from the leg swinging forward, but from the backward push of the rear limb.	Awareness of the posterior chain (glutes, hamstrings)	Highlight the invisible propulsive forces initiating movement
Posterior activation cue	Imagine your feet 'rowing' the floor backward to move forward	Focus on the rearward ground contact and propulsion	Engage glutes and hamstrings to generate the propulsion vector
Correction of common misconceptions	The movement doesn't start from the ankle or from pulling the leg forward	Eliminate perceptual illusions	Correct false motor representations
Guided practical experience	Walking uphill to feel clearer activation of the posterior chain	Distinction between what is seen and what is felt	Reinforce cause-effect connection through embodied learning
Symbolic anchoring	Like oars of a boat, your legs push backward to propel your body forward	Use of metaphor to guide motor focus	Integrate cognitive, perceptual and biomechanical understanding

Table 2. Operational summary of the effect-based instructional protocol (Group 2).

Instructional phase	Delivered content	Perceptual focus	Instructional objective
Movement form emphasis	Observe the rhythm and symmetry of the steps; maintain consistent limb movement.	Visual imitation of external form	Reinforce movement aesthetics and timing consistency
Foot rolling technique	Roll the foot from heel to toe in a fluid motion.	Awareness of foot contact pattern	Enhance step fluidity and walking efficiency through visible correction
Upper body posture	Keep the trunk upright and shoulders aligned while walking.	Postural alignment via observation	Guide external postural form
Step length and cadence	Try to maintain regular step length and tempo throughout the test.	Step symmetry and pacing	Promote rhythmic and coordinated gait
Correction by demonstration	Watch and repeat the correct walking model; mirror the instructor.	Visual mimicry	Facilitate learning via external modelling
Feedback based on outcome	You walked too slowly focus on increasing your step length.	Performance outcome	Provide correction through external results rather than internal cues

RESULTS

Statistical analyses were conducted using IBM SPSS Statistics (version 29.0). Normality of distributions was assessed using the Shapiro–Wilk test. Significant violations of normality were observed for most variables in both groups ($p < .05$), with the exception of the 6MWT in Group 1 ($p = .086$) and the Cooper test in Group 2 ($p = .242$). Consequently, non-parametric tests were applied for all subsequent analyses. To assess pre–post differences within groups, the Wilcoxon signed-rank test was applied. In Group 1, a statistically significant improvement was observed in the 6MWT, with a mean increase of +7.95 m from T1 ($M = 723.65$ m) to T2 ($M = 731.60$ m), $p < .001$. No significant changes were detected in BMI or Cooper test performance ($p > .05$). In Group 2, the Wilcoxon test revealed no significant differences between T1 and T2 in any of the measured variables (6MWT, Cooper test, BMI). Between-group comparisons at T1 and T2 were conducted using the Mann–Whitney U test. No significant differences were found at baseline (T1) or post-test (T2), confirming the initial comparability of the groups. To evaluate the intervention effect, individual change scores ($\Delta = T2 - T1$) were calculated for the 6MWT and the Cooper test. Mann–Whitney U analyses demonstrated a significantly greater increase in Group 1 compared with Group 2 for the 6MWT ($U = 5287.50$, $Z = 3.022$, p

$< .001$, $r = .223$), indicating a small-to-moderate effect size. No significant between-group difference emerged for the Cooper test ($U = 4748.00$, $Z = 1.50$, $p = .134$, $r = .112$). A descriptive analysis of individual changes in 6MWT performance showed that 32% of participants in Group 1 increased their walking distance from T1 to T2, compared with 12% in Group 2. A decline in performance was observed in 2% of participants in Group 1 and 11% in Group 2 (see Table 3).

Table 3. Distribution of individual performance changes from T1 to T2.

Test	Group	Improved n (%)	Declined n (%)
6MWT	G1 (n = 102)	32 (31.4%)	2 (2.0%)
6MWT	G2 (n = 82)	10 (12.2%)	9 (11.0%)
Cooper	G1 (n = 102)	16 (15.7%)	4 (3.9%)
Cooper	G2 (n = 82)	9 (11.0%)	8 (9.8%)

DISCUSSION

The present study examined whether a cause-based instructional approach, centred on the biomechanical causes of movement, would produce different functional outcomes compared with a traditional effect-based approach focused on observable movement form. The findings indicate that cause-based instruction led to a significantly greater improvement in submaximal walking performance, as measured by the 6MWT, compared with effect-based instruction. Although the effect size was small to moderate ($r = .223$), the between-group difference in change scores suggests that the instructional modality may have contributed to the observed improvement. In relative terms, the improvement corresponded to approximately a 1.1% increase in walking distance over a short time frame without modifications to training load. In trained adolescent athletes, even small short-term performance changes may reflect alterations in motor organisation rather than conditioning effects. Therefore, the magnitude of change should be interpreted within the context of a brief, instruction-only intervention. Within-group analysis showed a significant increase in 6MWT distance only in the cause-based group, while performance remained stable in the control group. More importantly, the comparison of individual change scores confirmed that the improvement was significantly greater in the experimental group.

The descriptive analysis further supported this pattern, with a higher number of participants in the cause-based group showing performance improvements and fewer declines compared with the control group. This distribution strengthens the interpretation that the instructional approach influenced the direction of change at the individual level. The results align with the OPTIMAL theory (Wulf & Lewthwaite, 2016), which emphasises the role of attentional focus and motivational mechanisms in optimising motor learning (Wulf et al., 2012). By explicitly highlighting the internal causal mechanisms underlying movement, the cause-based approach may have facilitated a deeper cognitive representation of gait mechanics, leading to measurable functional gains in a low-intensity task such as walking.

Regarding the Cooper test, a small, non-significant between-group difference was observed ($r = .112$). The absence of significant within-group changes indicates that short-term instructional interventions may exert more immediate effects on submaximal functional tasks than on endurance performance requiring greater physiological adaptation. This distinction likely reflects the different determinants underlying the two tests. While the 6MWT is sensitive to subtle changes in movement economy and mechanical coordination, the Cooper test is predominantly influenced by cardiorespiratory capacity and metabolic adaptations. No significant changes were detected in BMI in either group, which was expected given the short duration of the intervention and the absence of modifications to training load.

The stability of these variables strengthens the interpretation that the observed improvements in walking performance were attributable primarily to instructional rather than physiological factors. The relatively short duration of the intervention was a deliberate methodological choice. The study was intentionally structured to isolate the instructional variable while minimising physiological adaptations related to training load. By maintaining identical physical workloads and limiting exposure to three instructional sessions, the aim was to examine whether a “pure” cognitive-motor effect of instructional framing could be detected independently of conditioning changes. Given that training load remained unchanged and the intervention was instructional rather than physiological, the absence of significant effects on the Cooper test is consistent with the study design.

These findings support the interpretation that cause-based instruction primarily modulates motor organisation rather than endurance conditioning within a short time frame. From a pedagogical perspective, emphasising the causal and biomechanical foundations of movement may enhance functional motor organisation in adolescent athletes. Walking, as a relatively low cognitive-load activity, may be particularly sensitive to subtle improvements in internal coordination and movement awareness, whereby small adjustments in movement organisation may translate into measurable performance gains.

Limitations

Several methodological choices, while appropriate for the study context, should be acknowledged as limitations. The intervention was limited to three instructional sessions delivered over approximately one week. This design allowed isolation of the instructional variable without interfering with pre-season training demands; however, it likely constrained the magnitude and stability of observable adaptations. Previous research suggests that educational and motor learning interventions require prolonged exposure to produce consolidated and transferable effects, particularly in youth sport contexts (Granacher et al., 2011; Behringer et al., 2011). Future studies should therefore consider longer intervention periods, structured follow-up assessments, and integration of instructional approaches within regular training cycles to evaluate retention and long-term transfer. Additionally, although statistically significant differences were detected between groups, the observed effect sizes were small to moderate. Caution is therefore warranted in generalising the findings, and replication across different sports and developmental stages would strengthen external validity.

CONCLUSIONS

The present study indicates that a cause-based instructional approach, centred on the biomechanical causes of movement, was associated with a statistically significant improvement in submaximal walking performance (6MWT) in adolescent female athletes compared with a traditional effect-based approach. The observed between-group differences in change scores suggest that directing attention towards the causal mechanisms underlying movement may enhance functional motor organisation, even in the absence of modifications to training load. However, the modest magnitude of the effects and the short duration of the intervention suggest that instructional modality alone may not be sufficient to generate broader physiological adaptations within a brief time frame.

In practical educational contexts, the present findings support considering the inclusion of biomechanically grounded explanations within motor teaching practices, particularly during adolescence, when cognitive engagement and movement representation remain highly adaptable. Future research should investigate longer instructional cycles, integration with practical training, and longitudinal follow-up to determine the stability and transferability of these effects across different tasks and developmental stages.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. A. F. contributed to the conceptualization of the study, literature review, methodological design and drafting of the manuscript. S. T. contributed to the methodological supervision, research design, interpretation of findings, and critical revision of the manuscript. I. P. contributed to the technical and digital aspects of the study, including support for tool development and implementation and data analysis. L. C. contributed to the scientific supervision of the project, theoretical framing, interpretation of findings, and critical revision of the manuscript. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

We would like to thank the Cal State LA Office of Research, Scholarship, and Creative Activities for the mini-grant approved in support of this project.

CONFLICT OF INTEREST

Participation was voluntary and preceded by signed informed consent from athletes and their guardians. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

ETHICAL CONSIDERATIONS

The study procedures adhered to internationally recognised ethical standards for research involving human participants and were conducted in accordance with the Declaration of Helsinki. Written consent was obtained from both the athletes and their legal guardians before enrolment. In line with national regulations, formal institutional ethics committee approval was not required, as the intervention was non-invasive and conducted within routine training activities.

REFERENCES

- Armstrong, N., & Welsman, J. R. (1997). *Young People and Physical Activity*. Oxford University Press.
- Armstrong, N., Tomkinson, G., & Ekelund, U. (2011). Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *British journal of sports medicine*, 45(11), 849-858. <https://doi.org/10.1136/bjsports-2011-090200>
- Bach, P., Frank, C., & Kunde, W. (2024). Why motor imagery is not really motoric: towards a reconceptualization in terms of effect-based action control. *Psychological research*, 88(6), 1790-1804. <https://doi.org/10.1007/s00426-022-01773-w>

- Balyi, I., & Hamilton, A. (2004). Long-term athlete development: trainability in childhood and adolescence. *Olympic Coach*, 16(1), 4-9.
- Bandyopadhyay A. (2015). Validity of Cooper's 12-minute run test for estimation of maximum oxygen uptake in male university students. *Biology of sport*, 32(1), 59-63. <https://doi.org/10.5604/20831862.1127283>
- Baxter-Jones, A. D., & Maffulli, N. (2003). Endurance in young athletes: it can be trained. *British journal of sports medicine*, 37(2), 96-97. <https://doi.org/10.1136/bjsm.37.2.96>
- Behringer, M., Vom Heede, A., Yue, Z., & Mester, J. (2011). Effects of resistance training in children and adolescents: a meta-analysis. *Pediatrics*, 126(5), e1199-e1210. <https://doi.org/10.1542/peds.2010-0445>
- Bohannon, R. W. (2018). Six-minute walk test: A meta-analysis of data from apparently healthy elders. *Topics in Geriatric Rehabilitation*, 34(1), 30-36.
- Troosters, T., Gosselink, R., & Decramer, M. (1999). Six minute walking distance in healthy elderly subjects. *The European respiratory journal*, 14(2), 270-274. <https://doi.org/10.1034/j.1399-3003.1999.14b06.x>
- Chambal, E. J., Nhamussusa, D. M., Pacheco, M. M., Drews, R., & Tani, G. (2024). Read or listen? Effects of different kinds of instruction on the learning of a sport motor skill. *Motriz*, 30, e10230146. <https://doi.org/10.5016/s1980-6574e10230146>
- Cooper, K. H. (1968). A Means of Assessing Maximal Oxygen Intake: Correlation Between Field and Treadmill Testing. *JAMA*, 203(3), 201-204. <https://doi.org/10.1001/jama.1968.03140030033008>
- De Oliveira, I. S., Da Silva Oliveira, D., & Cattuzzo, M. T. (2016). The effect of different instructions in general motor competence and perceived athletic competence of children. *Journal of Physical Education and Sports Management*, 3(1), 108-126.
- De Stefani, E., Rodà, F., Volta, E., Pincolini, V., Farnese, A., Rossetti, S., Pedretti, F., & Ferrari, P. F. (2020). Learning new sport actions: Pilot study to investigate the imitative and the verbal instructive teaching methods in motor education. *PloS one*, 15(8), e0237697. <https://doi.org/10.1371/journal.pone.0237697>
- Geiger, R., Strasak, A., Trembl, B., Gasser, K., Kleinsasser, A., Fischer, V., Geiger, H., Loeckinger, A., & Stein, J. I. (2007). Six-minute walk test in children and adolescents. *The Journal of pediatrics*, 150(4), 395-399.e3992. <https://doi.org/10.1016/j.jpeds.2006.12.052>
- Granacher, U., Muehlbauer, T., Doerflinger, B., Strohmeier, R., & Gollhofer, A. (2011). Promoting strength and balance in adolescents during physical education: effects of a short-term resistance training. *Journal of Strength and Conditioning Research*, 25(4), 940-949. <https://doi.org/10.1519/JSC.0b013e3181c7bb1e>
- Haapala, E. A., et al. (2022). Associations of cardiorespiratory fitness, body composition and energy availability in adolescent female athletes. *Scientific Reports*, 12, 25795. <https://doi.org/10.1038/s41598-022-25795-x>
- Haibach-Beach, P. S., Perreault, M., Brian, A., & Collier, D. H. (2024). Motor learning and development. *Human Kinetics*.
- Jelonek, J., Pilis, W., Świat, M., Michalski, C., & Stec, K. (2017). Quality of sports training and the biological adaptation of athletes to race walking. *Physical Activity Review*, 5, 212-221. <https://doi.org/10.16926/par.2017.05.26>
- Kernan, W. N., Viscoli, C. M., Makuch, R. W., Brass, L. M., & Horwitz, R. I. (1999). Stratified randomization for clinical trials. *Journal of Clinical Epidemiology*, 52(1), 19-26. [https://doi.org/10.1016/S0895-4356\(98\)00138-3](https://doi.org/10.1016/S0895-4356(98)00138-3)
- Li, A. M., Yin, J., Yu, C. C., Tsang, T., So, H. K., Wong, E., Chan, D., Hon, E. K., & Sung, R. (2005). The six-minute walk test in healthy children: reliability and validity. *The European respiratory journal*, 25(6), 1057-1060. <https://doi.org/10.1183/09031936.05.00134904>

- Lloyd, R. S., Oliver, J. L., Faigenbaum, A. D., Howard, R., De Ste Croix, M. B., Williams, C. A., & Myer, G. D. (2015). Long-term athletic development-part 1: a pathway for all youth. *Journal of Strength and Conditioning Research*, 29(5), 1439-1450. <https://doi.org/10.1519/JSC.0000000000000756>
- Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, maturation, and physical activity (2nd ed.). *Human Kinetics*. <https://doi.org/10.5040/9781492596837>
- Migliaccio, G. M., Russo, L., Maric, M., & Padulo, J. (2023). Sports Performance and Breathing Rate: What Is the Connection? A Narrative Review on Breathing Strategies. *Sports (Basel, Switzerland)*, 11(5), 103. <https://doi.org/10.3390/sports11050103>
- Pascua, L. A., Wulf, G., & Lewthwaite, R. (2015). Additive benefits of external focus and enhanced performance expectancy for motor learning. *Journal of sports sciences*, 33(1), 58-66. <https://doi.org/10.1080/02640414.2014.922693>
- Riemann, B. L., & Lephart, S. M. (2002). The Sensorimotor System, Part II: The Role of Proprioception in Motor Control and Functional Joint Stability. *Journal of athletic training*, 37(1), 80-84. PMID: PMC164312.
- Riskowski, J. L., Dufour, A. B., Hagedorn, T. J., Hillstrom, H. J., Casey, V. A., & Hannan, M. T. (2013). Associations of foot posture and function to lower extremity pain: results from a population-based foot study. *Arthritis care & research*, 65(11), 1804-1812. <https://doi.org/10.1002/acr.22049>
- Sims, S. T., & Heather, A. K. (2018). Myths and Methodologies: Reducing scientific design ambiguity in studies comparing sexes and/or menstrual cycle phases. *Experimental physiology*, 103(10), 1309-1317. <https://doi.org/10.1113/EP086797>
- Song J. H. (2019). The role of attention in motor control and learning. *Current opinion in psychology*, 29, 261-265. <https://doi.org/10.1016/j.copsyc.2019.08.002>
- Tenan, M. S., Brothers, R. M., Tweedell, A. J., Hackney, A. C., & Griffin, L. (2014). Changes in resting heart rate variability across the menstrual cycle. *Psychophysiology*, 51(10), 996-1004. <https://doi.org/10.1111/psyp.12250>
- Wulf, G. (2012). Attentional focus and motor learning: a review of 15 years. *International Review of Sport and Exercise Psychology*, 6(1), 77-104. <https://doi.org/10.1080/1750984X.2012.723728>
- Wulf, G., Chiviacowsky, S., & Lewthwaite, R. (2012). Altering mindset can enhance motor learning in older adults. *Psychology and aging*, 27(1), 14-21. <https://doi.org/10.1037/a0025718>
- Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic Bulletin & Review*, 23, 1382-1414. <https://doi.org/10.3758/s13423-015-0999-9>



Diego Armando Maradona: Between secular sainthood and the redemption of Southern Italy

 Elia Fiorenza  . University of Calabria. Italia.

ABSTRACT

This contribution examines the figure of Diego Armando Maradona as a symbolic construction and an identity-forming device in 1980s Naples. The champion's arrival at SSC Napoli in 1984 occurred during a phase of profound socio-economic crisis marked by the aftermath of the earthquake and ongoing deindustrialisation, thereby acting as a catalyst for a process of collective identification. Through a historical and sociological analysis of the urban and footballing context, the essay interprets the "secular sainthood" attributed to Maradona as an expression of Southern redemption within the framework of the persistent North–South divide. The first league title in 1987 thus acquires a metapolitical significance, configuring itself as a symbolic victory against power structures perceived as hegemonic. The player's charismatic dimension transcends the sporting sphere and consolidates into a global myth capable of fostering pride, belonging, and shared memory, notwithstanding an awareness of his biographical contradictions.

Keywords: Sport history, Maradona, Naples, Collective identity, Southern redemption, Sporting myth.

Cite this article as:

Fiorenza, E. (2026). Diego Armando Maradona: Between secular sainthood and the redemption of Southern Italy. *Scientific Journal of Sport and Performance*, 5(3), 502-515. <https://doi.org/10.55860/OIXY3541>

 **Corresponding author.** University of Calabria. Italia.

E-mail: elia.fiorenza@unical.it

Submitted for publication March 23, 2026.

Accepted for publication April 30, 2026.

Published May 07, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/OIXY3541>

INTRODUCTION

5 July 1984: this was the day on which Diego Armando Maradona first set foot in Naples' Stadio San Paolo. At that time, the city was suffering the disastrous socio-economic consequences of the Irpinia earthquake—which, only four years earlier, had claimed approximately three thousand lives—and was simultaneously grappling with the sporting fortunes of a team that had avoided relegation to the second division by a single point just two months before. Two forms of distress, seemingly irreconcilable, nonetheless constituted fertile ground for investing the Argentine footballer's arrival with profound meaning. Two domains—the political and the sporting—merged and overlapped, fostering a shared representation of sanctity intelligible only considering the broader historical, socio-economic, and political context. The sanctification of Maradona by the Neapolitan people cannot, in fact, be understood apart from the long-standing rivalry between North and South. In this regard, the Southern Question should be interpreted not so much in terms of the post-Unification marginalisation of Southern demands, but rather as the rhetorical and mortifying prejudice imposed by the North upon the South and upon Naples in particular. It was a discriminatory phenomenon directed at a city which, in a quasi-messianic organization in the days preceding his arrival, designated Maradona as the myth capable of healing and redeeming an entire people in the face of Northern dominance.

From the outset, the footballer embodied the spirit of a prophetic leader, endowed with the capacity to “*read the hearts of men and [...] foretell what is to come*” (Vauchez, 1981, p. 445). By intuitively grasping the community's deepest aspirations, Maradona—employing a markedly political rhetoric—promised his supporters the conquest of the league title against the wealthy North. Before choosing the team, he chose the city; and he sought to become its charismatic leader.

On 10 May 1987, the miracle was fulfilled: Maradona delivered the first league title to the city of Naples. A victory secured on the football pitch soon assumed the symbolic significance of long-awaited social redemption. Supporters and citizens alike, united in a shared emotional horizon, interpreted and represented the scudetto as a triumph over the industrial, political, and sporting wealth of the North, and as the culmination of a “*battle against the powerful [that] could only be fought with the ball*” (Brancato, 2014, p. 37). A victory with one undisputed protagonist: Diego Armando Maradona.

MARADONA IS NAPLES. NAPLES IS MARADONA

The reflections that follow are the result of a field investigation. Interviews were conducted with the generation of Neapolitans who were able to witness Maradona's performances in person, to understand the mechanisms underlying the fans' identification with the Argentine player. At the same time, the study sought to assess the extent to which the transmission of collective memory contributed to the deification of the Argentine footballer's figure.

Even today, the negotiations surrounding Maradona's transfer—at the time a player for FC Barcelona—remain one of the most controversial (and protracted) episodes in the history of football transfers. Many have sought, and continue to seek, to claim authorship of the insight that led to the removal of the world's greatest footballer from the blaugrana. The truth, however, is that a convergence of circumstances—perhaps even of fate—created the conditions whereby Maradona, who as a very young man had been welcomed to Barcelona as a hero, came to be portrayed as the sole culprit behind two deeply disappointing seasons for the entire team.

The climate of resentment toward the Argentine champion was further exacerbated by the blaugrana president himself, Josep Lluís Núñez, who repeatedly accused him in public of leading an excessive lifestyle, effectively offering him up as a scapegoat upon whom to place the blame for the failed 1983–84 season.

Thus emerged a golden opportunity, one that few immediately recognised. Among those few was Ricardo Fajic, an intermediary who contacted several Italian clubs, conveying Maradona's desire to move to Serie A. It was, indeed, a golden opportunity, notwithstanding the reputation that preceded Diego: that of an unruly, extravagantly paid player, allegedly ill-suited—both professionally and tactically—to European football. Several clubs dismissed the proposal, but SSC Napoli did not. Represented by its president Corrado Ferlaino and its long-serving executive Antonio Juliano, the club met Maradona's manager on a late-May afternoon and reached a preliminary financial agreement. What followed was more than a month of announcements (and subsequent denials), newspaper headlines, internal clashes within Barcelona's board of directors, and public statements by the player himself—on radio and in print—affirming his firm intention to join Napoli, thereby rendering his reintegration into the Spanish side virtually impossible.

In the late afternoon of 30 June 1984, in room 1715 of the Hotel Princesa Sofía, Ferlaino and Juliano shook hands with Barcelona's vice-president Joan Gaspart, sealing an overall agreement worth thirteen billion lire. On the other side of the world, in his opulent villa in the shadow of the Cross of Pedralbes, Diego followed the unfolding events until he received confirmation of his transfer from Barcelona to Napoli (Piaci, 2025).

In the city, a murmur began to circulate—soon swelling, almost taking shape in the alleyways and squares—until it became a newspaper headline: Diego Armando Maradona is a new Napoli player.

Aboard the Alitalia Boeing 727 departing from Barcelona and bound for Fiumicino, where it would land a couple of hours later, Maradona granted his first interview to a RAI journalist, stating:

"I was fortunate that Napoli took an interest in Maradona and succeeded in signing him. Naples was in my destiny as early as 1978 Gianni Di Marzio tried to bring me to you. [...] I expect to find the tranquillity here that I did not have in Barcelona [...]. I know that football is lived here with passion, and I like that—Argentina is the same. They want to make me into a myth [...]. I will do everything I can to help Napoli win. Let us hope that God remembers Napoli and Naples [...]" (Piaci, 2025, p. 31).

It may appear paradoxical to seek tranquillity in Naples, especially when one considers the social challenges that characterised the city in the 1980s: crippled by unemployment and organised crime, worn down by degrading stereotypes, overwhelmed by mass tourism, and diminished by speculative interests. Nevertheless, Naples—whose rhythms of sleep and awakening followed the sporting fortunes of its team—and Maradona were united in a shared promise to free the city from the role to which the media had consigned it.

On 5 July 1984, a day remembered as Naples' own Woodstock, Maradona encountered the Neapolitan people for the first time, receiving the kind of welcome ordinarily reserved for deities. The city was covered with banners of greeting. At the packed Stadio San Paolo, more than sixty thousand spectators held two-thousand-lire tickets to witness the arrival of the *Pibe de Oro*, scheduled for 6:30 p.m. At that hour, television cameras from across the globe were trained on the staircase leading from the dressing rooms to the pitch. Dressed in a sky-blue tracksuit and white T-shirt, Maradona ascended from darkness into light, step by step, before a throng of photographers jostling to capture the scene. Nowhere else in the world would he have

received such a reception, and Diego seemed to grasp this as he emerged onto the field, moving with difficulty—during his half lap of honour, accompanied by a brief display of juggling—amid photographers, club officials, and awestruck witnesses. The San Paolo erupted in an uninterrupted roar lasting two minutes and forty seconds, welcoming its new idol as though at a birth (Piacci, 2025, p. 41). Those present in the stands sensed that they were living through a decisive moment; they perceived that they themselves—the city, its alleyways and neighbourhoods, every single stone of Naples and its province—stood at a turning point. Thus began Naples' footballing *belle époque*.

In the numerous accounts of Maradona's arrival in Naples, one repeatedly encounters testimony to a widespread sense of expectation surrounding a figure whom many attributed with quasi-divine qualities, as the following passage suggests:

"5 July 1984, a historic date. Maradona lands in Naples. It is the arrival of the Messiah of football, the one who will bring the league title to the Gulf. An entire stadium waiting for him in trembling anticipation. [...] To journalists arriving from afar, perhaps from Europe or Japan (and there was more than one), [...] the city offers itself without reserve, as always ready to display its fertile womb as mother of men and things unmatched anywhere in the world. [...] From the ever-fertile belly of a continent bounded by Vesuvius and the thousand ripples that lap in the sea of imagination, there emerged [...] in a few convulsive and festive days, a thousand keys to the authentic soul of a people capable, like no other in the world, of winning the daily wager of survival with nothing: Maradona posters, lighters bearing Maradona's likeness, T-shirts emblazoned with his face [...], caps and scarves praising Maradona, sky-blue flags with the face of the Pibe de Oro, [...] children's games named after the champion. [...] Someone even thought to imprint the unmistakable features of the Argentine star on the ankles of a new line of socks, to be launched with the slogan 'Maradona at the feet of the Neapolitans'" (Pecchinenda, 2023, pp. 35–54).

The destinies of Naples and Maradona converge into an indivisible representation. The Argentine champion is experienced as someone who fought against the establishment, deconstructing the system and overturning the balance of power in football and in society, in the name of a vindication of southern claims. Maradona thus becomes the mirror in which Neapolitan supporters rediscover their own identity; he is the guide who delivers to the community the first triumph in its history, subverting the existing order through a winning mentality and through traits that Neapolitans themselves often attribute to their self-representation: passion, rebellious spirit, anarchic genius, and the capacity to redeem humble origins through skill and impulsiveness.

In the — evidently idealised — narrative shared by Neapolitan supporters, Maradona arrives in Naples to assume the burden of restoring the community's pride; that same community does not hesitate to extend success on the pitch to the broader territorial arena, identifying the Argentine champion as a leader capable of providing the material necessary to oppose the tangible clichés that afflict Neapolitan identity, thereby rehabilitating and redeeming the city in its socio-economic dimension.

Indeed, Diego was always distinguished by his capacity for synthesis and response. As he grew older, he increasingly took on social causes, beginning to express – at times harsh – judgments against governments, both sporting and political, and performing radical gestures while struggling with his own personal torments. His continuous metamorphosis had rebellion as its common denominator: Maradona was a rebel both on and off the field, and for this reason his story bears the marks of uniqueness.

In short, it was a phenomenon which—although complex—managed to reconcile genders and generations across time. In his relentless forays across the pitch, and in the way he expanded the horizons of the footballing imagination—until then known—through unthinkable plays and goals, his fans nurtured the hope of being able to reclaim almost everything, to redeem humiliations and insults. From this perspective, one may agree that Maradona represented— from Buenos Aires to Naples—*"dream and social raison d'être"* (Ciriello, 2018, p. 51).

10 May 1987: the day of redemption. The day of the first championship title in the history of Napoli and of Naples.

In the jubilant dressing room, Maradona declared to the press:

"This is the most important celebration of my life. I have won everything through suffering. I won the World Cup far from my homeland, whereas in Naples I felt at home, so this is the most important victory" (Piaci, 2025, p. 95).

These might seem like words dictated by euphoria and exaltation. Nothing of the sort; years later, in his autobiography, he added:

"Winning the first Scudetto with Napoli in sixty years of history was, for me, a victory without comparison. Different from any other, including the title with the 1986 Selección. Because we built Napoli from the ground up, like workers. [...] It was a Scudetto for the entire city" (Maradona, 2012, p. 87).

Through the match on the pitch, the cohesion of a community solidifies, a community that feels it belongs to the same destiny and is represented by the same idol. Whether they witnessed him live or learned of his feats through mediated memory, supporters — beyond recognising Maradona as the protagonist of victories at national and international level — found in his authenticity and sporting merits the ground on which to cultivate collective pride and self-esteem. Thus matches, which represent confrontations between irreconcilable worlds and different ways (not only of playing football but also) of living, are elevated to the status of:

"Social rituals that foster belonging, solidarity and the social identity of those who feel they participate in a communal destiny, who share happiness and suffering around their devotion to their heroes". (Bifulco, 2021, p. 1).

The substance of the ideas through which new generations perceive Maradona therefore does not differ from that of those who had the privilege of experiencing him live; this highlights how the Argentine still maintains his status as a symbolic, emotional and ritual centre, catalysing processes of shared identification and serving as a mirror of Neapolitan self-representation.

The first championship, as *La Repubblica* wrote, was *"like stepping into a theatre for those who until then had had to content themselves with watching others enter it. A moment of happiness granted to many Neapolitans: both those living in their homeland and those forced to travel the world, who for a day reclaim their shared identity."* And it had been delivered by someone who liked to define himself as the son of Naples, the son of a city in which — during the years of his stay — more than five hundred children were registered with the name that more than any other had to be bestowed: Diego (Carotenuto, 2017).

The title was celebrated as a miracle: crowds flooded the streets, car horns blared incessantly, the balconies of houses were painted in blue, and the image of Maradona was carried in procession as if he were a saint. A striking sign of the extent to which the victory was perceived in a transcendent manner was the installation, at the cemetery, of a banner that read: “*You do not know what you missed!*”, implying that the magnificence of that triumph was such that even the dead had been unfortunate not to witness it (Nardi, 2025, pp. 38–39).

When the city exploded with joy for its first national title, it certainly did not enjoy the tourist and cultural momentum that characterises it today; quite the opposite. At the time, Naples had reached the nadir of socio-economic, administrative and cultural decline and was considered among the most degraded metropolises in Europe, also due to the misleading media narrative of bloody clashes between Camorra clans, the misallocation of funds intended for post-earthquake reconstruction, and the dismantling both of the steel plant in Bagnoli and of the Alfa Romeo factory in Pomigliano d’Arco.

The discontent of the 1980s was alleviated by Maradona, who acted as a social remedy capable of restoring joy, distraction and lightness to a restless people and of repositioning a deindustrialised and stigmatised city before Europe’s other capitals. In a league that, at the time, was considered the most important in the world, Napoli was made champion by the Argentine superstar: from that moment on, Naples and Diego became inseparable.

From the moment he left the city in 1991, Maradona has been everywhere. It is impossible to walk through Naples without seeing his presence on the walls or on the shirts of children playing football in the squares; those shirts bearing the number 10 on their backs transcend the sporting sphere and represent the desire to keep history alive, with all its contradictions, revanchism and claims. His effigy now fills the city, from the Historic Centre to the Centro Direzionale, from the Sanità district to the Spanish Quarters: here, in the 1990s, a short distance from the Diego Armando Maradona Museum, a mural was created that has become one of the most visited tourist attractions in the world, celebrating every day the one who sided with the indomitable side of the world, in defence of the marginalised.

His death, on 25 November 2020, in addition to triggering collective mourning (also facilitated by new media), further propelled the magnification of a figure who will forever symbolise and embody a football-territorial community — that of those who continually strive for affirmation.

THE MARADONA PHENOMENON

For a long time, attempts have been made — and continue to be made — to narrate the Maradona phenomenon through art, cinema, poetry, literature, journalism, and even faith. None of these numerous efforts, however, has succeeded in formulating a definition capable of encompassing Diego Armando Maradona in his entirety. This is not due to any lack of commitment on the part of those who have tried; rather, it is because he represents a figure of such irreducible complexity that any attempt to synthesise the breadth of meanings he assumed — and continues to assume postmortem — proves inadequate. Diego embodies an ideal — that of redemption — in which ordinary people, the popular classes, have been able to recognise themselves. He represents one who made it, and perhaps it is precisely this aspect — even more than his performances on the pitch — that projected him into the dimension of a phenomenon capable of generating collective aggregation despite his personal fragilities and contradictions.

His death completed that process of collective identification that had begun on the day of his presentation at the stadium, the place where his creative genius would long find expression. From 5 July 1984 onward, it

became impossible to separate Maradona from a team of which he became a genuine intangible asset, and from a city whose people he identified with to such an extent that today he is often said to possess “*the same gaze found in Neapolitan Street children*” (Bromberger et al., 1995, p. 137).

Among all the champions in the global history of sport, it is difficult to identify a personality who has received the same degree of media attention as Maradona, both in his private and professional spheres. Constantly in the spotlight of success, deeply embedded in the commercial dynamics of modern sport, frequently unconventional, and yet an idol for several generations, he offered — well beyond the playing field — points of identification so powerful that he may be described less as a celebrity of the popular classes and more as a universal cultural phenomenon.

Maradona was loved and continues to live in the memories and narratives of supporters, because he changed the image of Naples, demonstrating to Neapolitans that it was possible to compete with — and even triumph over — the wealthy North. That triumph, beginning with the first national title, triggered a chain reaction in which the city of Naples became, simultaneously and in all its dimensions, both subject and object of exaltation and glorification.

El pibe de oro, by writing unforgettable pages in the history of football, succeeded in weaving intense bonds with both the Neapolitan and the Argentine people. It is therefore worth asking how a figure — albeit a genius — frequently criticised for his friendships and for his drug addiction, managed to penetrate so viscerally the heart and soul of Argentines. The answer may lie in the conditions that characterised Argentina in the first half of the twentieth century, marked by profound socio-economic imbalances. While Buenos Aires enjoyed development of an almost European character, moving away from the capital revealed degraded contexts dominated by precariousness, poverty, and social inequality. The country was literally divided between two contrasting scenarios, separated by a deep fracture within which Maradona positioned himself as a unifying force, restoring — through football — hope to the oppressed and confidence in social redemption.

Although thousands of kilometres distant from South America, Naples presented certain parallels. The city, with all its contradictions, was likewise afflicted by socio-economic and political difficulties that almost constituted fertile ground for the construction of an indissoluble bond with Diego. Maradona thus became a son of Naples because he elevated a reality often denigrated and brought it to the summit of Italian football, breaking the established pattern whereby clubs from Northern Italy contested the national championship each year.

This perspective helps explain how Maradona became a social phenomenon both in Argentina and in Naples. His humble origins — never denied and frequently emphasised with pride — were invoked as testimony to emancipation from oppression and suffering. This condition perhaps provided the assist, to remain within football terminology, for the construction of a myth into which collective dreams and hopes could be projected. In certain respects, this Third-World hero represented the South — both of America and of Italy — and, through sport, he generated emotions that alleviated the pain of precariousness for that part of the world.

For many, he represented a god who, despite his weaknesses, rendered proud realities such as Naples and Argentina, accustomed to living in the shadows. That pride bore the unmistakable imprint of redemption.

A footballing god is such insofar as his sporting virtues exceed ordinary standards, rendering him *de facto* worthy of devotion. Yet mastery within the playing field must also be combined with the capacity to “*trigger powerful emotional gratifications nourished by a genuine form of vicarious glory*” (Bifulco & Pirone, 2014, pp.

36–53). For Neapolitan supporters, Maradona remains the charismatic footballer whose extraordinary performances continue to sustain their self-esteem. In this case, one clearly observes the psychological mechanism known as “*basking in reflected glory*” (Cialdini et al., 1976), whereby the champion’s triumphs are absorbed and experienced by supporters as if they were their own.

His name still resonates in conversations among supporters, in the streets and across social media. This oral and digital circulation reinforces the exaltation of those who recognise themselves in their leader and who, through the memory of his exploits, cement a cohesive and solidaristic community. Within the narrative of the myth, Maradona acquires traits of inimitable perfection that constitute a fundamental resource for sustaining the self-regard of a community that felt chosen by the greatest footballer of all time and guided by him toward the awareness that the city could — and should — reclaim its own destiny, since no preordained system existed capable of preventing its social ascent.

His undisputed qualities, the widespread belief that he single-handedly won the 1986 World Cup, and the attraction he exerted on players willing to join Napoli simply to play alongside him all contributed to the typification of his figure. To this must be added the rhetoric of social redemption, coloured by multiple nuances: opposition to the consolidated powers of football, reversal of social hierarchies, and the redemption of the South against the North. Such elements allow Maradona to be identified both as a “*reinforcing hero*,” insofar as he embodied attitudes¹ — anarchy and passion — through which Neapolitans describe themselves (Bromberger, 1999, pp. 87–105; Bifulco & Pirone, 2014, p. 48), and as a “*seductive hero*,” one who, by breaking with established rules, indicated new possible horizons (Bifulco & Dini, 2014, p. 117).

Once it has been established that many of the reinforcing and captivating features of Maradona’s figure function as a mirror for certain idealized traits regarded as constitutive of identity by his community of supporters, one must avoid the mistake of assuming that Diego was also adopted as a behavioural model in everyday life. On the contrary, his problems with drug addiction, his less than irreproachable management of certain family dynamics, and his borderline friendships are by no means acknowledged as beyond reproach². This is clearly perceptible in the words of G., who states:

“[...] certainly what is remembered of Maradona are his sporting feats and, to some extent, also his decline on a human level, because one must not conceal that [...] from this point of view he disappointed his supporters somewhat, especially the more demanding ones, that is, the more educated ones, those who would have wanted a Maradona who remained a permanently positive symbol of the city of Naples. Instead, Maradona was great in his profession but less so from a human standpoint, because he gave in to the pressures of the environment in which he lived” (Bifulco & Dini, 2014, p. 117).

A. expresses a similar view, maintaining:

“[...] Maradona made many mistakes... he used drugs, he cheated on his wife, he did not acknowledge a son and always treated him badly. [...] Ultimately, his character on the pitch — rebellious and combative — later created problems for him in life. [...] he did not always know how to choose the right company, he often surrounded himself with opportunists, and he was unable to

¹ Recurring is the attribution to Maradona of the image of the *scugnizzo*, that is, the young Neapolitan who — rebellious by nature — makes his way through the hardships of everyday life, often resorting to cunning stratagems.

² One may consider the recognition of Diego Armando Maradona Junior, the son born in Naples from an extramarital relationship.

manage his own distress. One can understand him to some extent, even if not justify him. [...] He is certainly not a model man... at most, a reference point as a player" (Bifulco & Dini, 2014, p. 118).

These negative traits, however, do not constitute the core of identification between Maradona and his fans. This is also because many of them, when abstractly describing the prototype of a "*champion*," emphasize the importance of ethical conduct beyond the sporting arena. What is striking, nevertheless, is the extent to which Maradona generates pride within his community to the point that, in practice, his private life is entirely separated from his sporting life. In his case, reprehensible behaviour is considered strictly personal and therefore beyond judgment, whereas — from sport — those aspects most suitable for self-representation (charisma, tenacity, generosity, leadership) are selectively retained.

POLITICS, WAR, AND FOOTBALL

In Argentina, during the 1920s and 1930s, art—together with other cultural forms (above all cinema and tango)—actively participated in the construction of an *Argentinity* that had previously been unknown. In this process of identity formation, football also played a significant role, as a sport capable of channelling and containing intense popular passions.

During those years, the so-called *criollo* style emerged on football fields. This style was celebrated as typical of the Argentine people and was characterized by features that were manifestly opposed to those of British football—the game introduced to Argentina together with the Union Jack. Unlike the British style, which was taught in sports clubs and emphasized athletic discipline, the *criollo* style was characteristic of the *pibe*, the boy accustomed to playing on the *potreros*, uncultivated lands used for free cattle grazing under the watch of the *gaucho*. On the *potreros*, the *pibe* learns and refines cunning and ingenuity, expressing them through dribbling—a gesture which, more than any other, represents an act of indiscipline grounded in virtuosic individual effort (Serra, 2015, p. 14).

Within this perspective, "*el fútbol criollo resulta por tanto estrechamente asociado con la espontaneidad y la libertad de la infancia*" (Serra, 2015, p. 15), so much so that great players continue to be called *pibes* even in adulthood; ultimately, the finest plays are nothing other than the actualization of a skill acquired on the *potrero* and untainted by the tactical formalism of the mature game.

Among all players, Maradona was the *pibe de oro*, the perfect prototype of *criollo* football, whose greatness derived from incomparably extraordinary gestures—so extraordinary that they transcended the purely sporting dimension and entered the symbolic realm. That greatness reached one of its peaks on 22 June 1986 at the Estadio Azteca in Mexico City, during the World Cup quarter-final against England, a country against which Argentina had—just four years earlier—lost a war over the Falkland Islands.

To fully understand the nature of the media coverage Maradona received from the British press, it is necessary to examine how politics intertwined with narratives concerning both the player and his country. The 1982 war over the disputed Falkland Islands (Las Malvinas, according to the Argentine denomination), which lasted seventy-four days and resulted in Argentine casualties more than double those of the British, was accompanied by a strong mobilization of the popular press in both countries, contributing to the hardening of public opinion.

The conflict shaped the way in which subsequent sporting encounters between the two nations would be represented. Its legacy was such that, in the lead-up to the 1986 World Cup quarter-final, British tabloids

expressed concerns that a fringe of supporters might exploit the sporting event as an opportunity for politically motivated violence. The *Mirror* reported that “troops and riot police (were being) called in for England v Argentina game” (Boyle, 2023, pp. 6–7).

Despite repeated calls for calm from Maradona—who had also been appointed captain of the national team—he later admitted in his autobiography:

“In the pre-match interviews we all had said that football and politics shouldn’t be confused, but that was a lie. We did nothing but think about that. Bollocks was it just another match!” (Boyle, 2023, p. 7).

None of the British media believed that the match between Argentina and England was merely a football game: it was evident that the pitch had become an echo chamber for the political tensions inherited from the still-recent conflict over the Falklands. The tabloids, moreover, played their own nationalistic game by employing language reminiscent of military conflict, stirring patriotic sentiment while simultaneously condemning potential violence.

The match was won 2–1 by Argentina thanks to two goals scored by Maradona, both destined to enter football history—albeit for opposite reasons.

The first, the famous handball goal in the 51st minute, was interpreted by the English press as a deliberate act of deception, with scathing headlines openly branding him a *cheat*. For Argentines, however, that offensive gesture against conventional reason represented the work of the *pibe* of the *potrero*, who outwits the stronger opponent through cunning and trickery.

The second goal, scored in the 56th minute, was universally recognized as an unprecedented technical masterpiece. Even British commentators—despite the highly charged emotional climate—described it as one of the greatest goals in football history. To understand how Argentine fans experienced the moment, it is useful to recall the words of commentator Víctor Hugo Morales:

“La va a tocar para Diego... [...] Gracias Dios, por el fútbol, por Maradona, por estas lágrimas, por este Argentina 2 - Inglaterra 0” (Serra, 2015, p. 19).

Only five minutes earlier Maradona had broken the rules, yet that transgression was redeemed by the “Goal of the Century,” through which Argentines felt avenged. Their avenger was a demigod: the first goal representing his earthly half, the second his transcendent one.

These two moments determined the ambivalent representation of Maradona in the British media until the day of his death: “For some he was cheat, for others a footballing genius and for many more he was in fact both of these things” (Boyle, 2023, p. 10).

From that match onward—at a time when television coverage of football was still limited—Maradona became the first player to experience the dimension of global celebrity constructed through the media. In subsequent years, however, attention gradually shifted from sporting admiration to off-field events, including drug addiction, fiscal controversies, and accusations of connections with organized crime. This trajectory culminated in the public fall of the genius during the 1994 World Cup in the United States, when he tested positive in an anti-doping test and was suspended from the tournament.

The episode consolidated the representation of a constant tension inherent in Maradona: between glory and scandal, myth and human fragility, talent and self-destruction. Yet rather than diminishing him in the eyes of the world, this tension further reinforced his status as a global cultural icon.

FROM FALL TO POSTHUMOUS CULT

Mediterranean cultures—crossroads of civilizations in which ancient and modern myths coexist—often blur the boundary between the sacred and the profane, blending religious devotion with shared symbolic constructs related to family, art, and sport. Naples represents an emblematic example. On the one hand stands the sacred (and official) cult of the city's patron saint, Saint Januarius, whose blood miraculously liquefies; on the other stands the secular cult of Diego Armando Maradona. Both coexist within a unique form of syncretism.

In Naples, therefore, manifestations of the sacred are not confined to religion alone: the football match becomes a liturgy, the playing field a secular temple, and supporters' chants totemic litanies. The veneration of Maradona quickly evolved into a socio-cultural phenomenon endowed with its own symbolism and rituals from the moment of his arrival in the city. His sanctification is reflected in countless objects, rituals, and images. A gigantic mural in the Quartieri Spagnoli has become a pilgrimage site where fans light candles and leave offerings; a bar in the historic centre has set up a shrine displaying a lock of Maradona's hair preserved as a sacred relic and labelled "*Capello miracoloso*" (Nardi, 2025, p. 39). Such practices, almost religious in nature, demonstrate how supporters have created rituals analogous to the cult of saints³.

However, like every hero and myth, Maradona's story also included a fall followed by a process of transfiguration. The god revealed his human weaknesses in the early 1990s with his suspension for doping and his departure from Naples amid scandals. A significant portion of supporters felt betrayed and abandoned, and only time was able to heal the wounds and open the way to a more indulgent narrative. Maradona thus began to be portrayed as a martyr who had sacrificed himself to glorify the Global South and as the victim of a conspiracy orchestrated by the powerful North to discredit Naples. Within such a narrative, the martyr who died for the aspirations of the people he had chosen to represent could only deserve full absolution and veneration.

In the Neapolitan collective imagination, Maradona reached apotheosis in three distinct moments: the bankruptcy of the football club in 2004, the player's death in 2020, and Napoli's victory in the third league title in 2023. After his departure, Napoli sank into the depths of the national league standings, reaching the lowest point in its history with a bankruptcy ruling caused by debts accumulated in previous years. This event further illuminated the Maradonian golden age, nourished by stories transmitted with pathos and folklore. His death, moreover, plunged the city of Naples into a grief comparable to that of a family bereavement: thousands gathered before the mural to recite prayers, leave flowers, and light candles; at the same time, the city council promptly decided to rename the San Paolo Stadium, relocating it within the civic pantheon of Naples as the Stadio Diego Armando Maradona⁴.

³ In this regard, the famous graffiti proclaiming "*Maradona is better than San Gennaro*" is an expression that—though steeped in irony—reflects the elevation of Maradona's status to the city's new patron saint. Cf. Nardi, "*Maradona, Secular God of the Mediterranean*," 39.

⁴ Furthermore, during the celebration of Holy Mass, a priest commemorated Maradona as the one who "*had brought joy [to] the city*," de facto elevating him to an intercessor of collective happiness. Moreover, following the victory of the Coppa Italia in 2020, the players carried a massive poster depicting his image around the pitch, symbolically including him in the celebrations and as proof of the conviction that he watched over the team's fortunes. Cf. Nardi, "*Maradona, Secular God of the Mediterranean*," 40, 42.

Finally, the third league title reactivated the cult of the *pibe de oro* among a generation that had never seen him play. The victory was explicitly linked—by the media themselves—to Maradona’s hand from the heavens: on that occasion his image was displayed throughout the city, the stadium was illuminated with the colours of Argentina, and supporters celebrated wearing the number 10 jersey, claiming that the triumph after thirty-three years was a true miracle, once again invoking religious language.

All this suggests that Maradona assumed the traits of a divinity venerated outside official religious canons and became an all-encompassing symbol of values and identity. His story provided the city of Naples with a triumphant narrative that redeemed it from the stigma to which it had long been relegated. Although Maradona remains a universal icon, it is in Naples that his cult found particularly fertile ground. The city’s cultural traits—strong communal ties and a deep reverence for its patron saint—led people to apply religious practices naturally to a modern figure.

In a city with a deeply rooted Catholic tradition, religious fervour was thus channelled toward an alternative outlet, enriching the symbolic landscape of Saint Januarius and placing beside him a civic saint: a charismatic patron-like figure capable of uniting believers and non-believers within a tangible cult. This does not imply that Neapolitans abandoned their faith but rather that “[the] football has partly taken over the role of religion as the chief occasion for mass celebration” and that football functions “as a parallel civil religion with its own calendar, rituals and saints” (Nardi, 2025, p. 43).

CONCLUSIONS

It is difficult to define all that Maradona was — and continues to be — for his countless fans. For Argentinians he was more than José Francisco de San Martín; for South America (with the exception of Brazil), more than Simón Bolívar; for Africa, a rock star scarcely less celebrated than Nelson Mandela; for the United States, part demon and part outlaw; and finally in Europe, where many tend to draw a sharp distinction between the man and the footballer (Ciriello, 2018, p. 106).

And then there is Naples, to which Diego gave glory while receiving love and compassion for every extra-footballing mistake. Each time the myth fell, there was always someone ready to lift him up again, “*rekindling that electricity which [...] he had brought to the city*” (Ciriello, 2018, p. 106).

At the mere utterance of his name, people grow emotional; memories burst forth in a triumph of feelings and sentiments that nourish the spiritual energy of the idol, the chosen one, the man who — in a collective sport — managed to push self-monumentalization to its extreme. That same process ultimately destroyed him, making him feel alone despite being constantly surrounded by his court — the court of tyrants.

In pursuing the absurd, Diego possessed the rare ability never to appear unsympathetic and, indeed, to transform the collective hostility of stadiums into admiration and consent, combining the authority of a king with the imagination of a jester. His present capacity to dominate the scene mirrors the ease with which, forty years ago, he would enter the goal with the ball at his feet — a simplicity that only reinforced his aura of the supernatural.

Maradona defies conventions: he exercises power while rejecting its structures and lives with the constant contradiction of someone who seeks targets among a long list of enemies and (false) friends, never withdrawing from judgment. That judgment highlights the missteps and anxieties of his private life—marked by illness, arrests, courts, psychiatric institutions, and hospitals—yet Diego mocks it all, for he has no interest

in securing anything other than the status that has always defined him: that of the child from Villa Fiorito on a football field, the only place where he truly exists.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

- Bifulco, L. (2021). Tra eroismo calcistico, identità e comunità: Maradona, Napoli e il tifo partenopeo. *Bibliomanie. Letterature, storiografie, semiotiche*, 52(1), 1.
- Bifulco, L., Bory, S., Diana, P., & Pecchinenda, G. (Eds.). (2023). *Global Maradona: Dall'uomo all'eroe sportivo. Dalla celebrità al mito* (Vol. 1). Novalogos.
- Bifulco, L., & Dini, V. (Eds.). (2014). *Maradona: Sociologia di un mito globale*. Associazione Ipermedium Libri.
- Bifulco, L., & Pirone, F. (2014). *A tutto campo: Il calcio da una prospettiva sociologica*. Guida.
- Boyle, R. (2023). Maradona and the British media. In P. Brescia & M. Paz (Eds.), *Diego Maradona: A socio-cultural study* (pp. 6-10). Routledge. <https://doi.org/10.4324/9781003196587-7>
- Brancato, S. (2014). Fenomenologia di Maradona: Il "briccone divino" e l'epica mediale del calcio. In L. Bifulco & V. Dini (Eds.), *Maradona: Sociologia di un mito globale* (p. 37). Associazione Ipermedium Libri.
- Bromberger, C. (1999). *La partita di calcio: Etnologia di una passione*. Editori Riuniti.
- Bromberger, C., Hayot, A., & Mariottini, J.-M. (1995). *Le match de football: Ethnologie d'une passion partisane à Marseille, Naples et Turin*. Éditions de la MSH. <https://doi.org/10.4000/books.editionsmsh.4077>
- Carotenuto, A. (2017, May 5). I figli del primo scudetto del Napoli si chiamano Diego. *La Repubblica*.
- Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, 34(3), 366-375. <https://doi.org/10.1037/0022-3514.34.3.366>
- Ciriello, M. (2018). *Maradona è amico mio*. 66th and 2nd.
- Maradona, D. A. (2012). *Io sono El Diego*. Fandango.
- Nardi, L. (2025). Maradona, secular god of the Mediterranean: Sporting legend and identity cohesion in post-championship Naples. *International Journal of Social Sciences & Education*, 15, 38-43.

- Pecchinenda, G. (2023). Essere Maradona: L'umana fragilità di un eroe. In L. Bifulco, S. Bory, P. Diana, & G. Pecchinenda (Eds.), *Global Maradona: Dall'uomo all'eroe sportivo. Dalla celebrità al mito* (pp. 35–54). Novalogos.
- Piacci, G. (2025). *A Napoli con Maradona: Non sarò mai una città comune*. Giulio Perrone Editore.
- Serra, M. (2015). Maradona entre la tierra y el cielo. *Cuadernos de Información y Comunicación*, 20, 14-19. https://doi.org/10.5209/rev_CIYC.2015.v20.49538
- Vauchez, A. (1981). Santità. In R. Romano (Ed.), *Enciclopedia Einaudi* (Vol. 12, Ricerca-Socializzazione, p. 445). Giulio Einaudi.



Tracking the athlete's eye: Comparing oculomotor metrics with non-athletes

-  **Clea P. Tucker**  . California State University. Los Angeles, California, United States of America.
Western University of Health Sciences. Pomona, California, United States of America.
-  **Beatriz Barragan**. A.T. Still University. Mesa, Arizona, United States of America.
-  **Arianna Soave**. NeuroFlex/Saccade Analytics, Inc. Canada.
- Nafiseh Ghoroghchian**. NeuroFlex/Saccade Analytics, Inc. Canada.
-  **William Pei**. NeuroFlex/Saccade Analytics, Inc. Canada.

ABSTRACT

Oculomotor measures provide objective, quantifiable markers of brain health that support the assessment and longitudinal monitoring of concussion outcomes. Athletes may perform differently than non-athletes on these tasks due to sport-specific training, yet existing protocols use normative ranges based on general populations. This study established baseline oculomotor metrics in college athletes using a virtual reality protocol and compared performance to general population normative data. A total of 163 university student-athletes (ages 17-30) completed an eight-subtest protocol assessing smooth pursuit, saccades, vestibular ocular reflex, optokinetic nystagmus, anti-saccades, and spontaneous nystagmus. Results revealed significant differences between athletes and the general population on 28 of 44 metrics, indicating distinct performance levels in head and eye movements. These findings emphasize the need for athlete-specific normative standards to ensure accurate concussion assessment. The normative ranges established provide a foundation for future concussion evaluation and management in athletes, underscoring the critical need to compare athletes to sport-specific rather than general population standards for improved diagnostic accuracy and clinical utility.

Keywords: Sport medicine, Assessment, Oculomotor, Vestibular, Virtual reality, Concussion.

Cite this article as:

Tucker, C. P., Barragan, B., Soave, A., Ghoroghchian, N., & Pei, W. (2026). Tracking the athlete's eye: Comparing oculomotor metrics with non-athletes. *Scientific Journal of Sport and Performance*, 5(3), 516-529. <https://doi.org/10.55860/TCXH2824>

 **Corresponding author.** Western University of Health Sciences. Pomona, California, United States of America.

E-mail: cleatucker@gmail.com

Submitted for publication February 23, 2026.

Accepted for publication April 06, 2026.

Published May 07, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/TCXH2824>

INTRODUCTION

Oculomotor measures provide objective brain health metrics that can be valuable for assessing and monitoring concussion progression (Hunfalvay et al., 2019). Advances in virtual reality technology now make it possible to administer these measures in non-medical environments, such as on the sports field, offering a practical tool for evaluating athletes in real time. However, previous research suggests that athletes may perform differently from non-athletes on oculomotor tasks, likely reflecting sport-specific training effects and performance demands (Kubitz et al., 2020; Garner et al., 2023; Laby & Appelbaum, 2021; Zwierko et al., 2019). Despite this, existing protocols and normative ranges are based primarily on non-athlete populations (Lunkova et al., 2024). Comparing athletes to these reference values is therefore essential to determine whether their performance diverges significantly from typical norms. If so, the development of athlete-specific normative ranges will be critical to improving the accuracy and clinical relevance of concussion diagnosis and management in this unique population.

An estimated 1.6 to 3.8 million sport-related concussions occur every year in the United States (Centers for Disease Control and Prevention, 2019; Choe, 2016), not including the large number of sub-concussive events that often go unreported (Nauman, Talavage, & Auerbach, 2020). Despite this high prevalence, there are still no universally accepted diagnostic criteria to reliably detect concussions or to characterize the severity of their pathophysiological and functional sequelae, nor are there rehabilitation protocols endorsed by most sport clinicians (Broglia, Sosnoff, & Ferrara, 2009). Traumatic brain injury is not a single event but rather a sequence in which the initial concussion acts as a catalyst for a cascade of physiological and functional changes in the brain and body, expressed with varying paces, severities, and durations of symptoms (Choe, 2016). Assessing concussions remains particularly challenging because current approaches rely heavily on self-report and subjective clinical evaluation, which are limited in accuracy and sensitivity (Dams-O'Connor et al., 2014; Dorman et al., 2015; Finnoff et al., 2009; Kontos et al., 2014; Kroshus et al., 2015; McLeod & Leach, 2012). Moreover, these methods are often not sensitive enough to detect subtle impairments that persist beyond the acute stage, with their effectiveness diminishing after four weeks to three months post-injury (Cavanaugh et al., 2005; Slobounov et al., 2011).

Concussion assessment in athletes presents unique challenges that add complexity to an already difficult task. In many cases, evaluations must be performed by clinicians in distracting and high-pressure environments, where grading and capturing all possible impairments is particularly challenging (Dams-O'Connor et al., 2014; Kroshus et al., 2015). Furthermore, while portable testing technologies hold promise, they are not equally inclusive or effective for all populations. For instance, functional near-infrared spectroscopy (fNIRS), a widely studied tool, often fails to collect valid data from individuals with darker complexions or coarse/thick hair, disproportionately affecting athletes from the African Diaspora and Latinx backgrounds (Tucker & Rivera, 2021), highlighting broader systemic disparities and biases in neuroscience (Webb, Etter, & Kwasa, 2022). Another concern is that although the percentage of reported concussions in sports is lower compared to other daily activities, athletes are uniquely pressured to return to play quickly, often without a clear understanding of the injury's severity or potential long-term consequences. Given the highly variable nature of concussion signs and symptoms, some injuries are mislabelled as "*mild*" or dismissed altogether as "*not real*," contributing to premature return-to-play decisions. Critically, prior research indicates that even athletes who appear asymptomatic may still show structural and functional brain changes (Broglia, Sosnoff, & Ferrara, 2009), underscoring the limitations of current assessment approaches and the risks faced by this population.

Post-concussion impairments in athletes vary by factors such as sport, gender, and type of head impact, and may not always present with obvious symptoms (Chen et al., 2019; Choe, 2016; Kontos et al., 2019; Lumba-Brown et al., 2020). Although many are believed to "fully recover" within one to two weeks, this definition is problematic, as subtle physiological, cognitive, and motor changes may persist undetected (Pearce, Tommerdahl, & King, 2019; Quintana et al., 2020). In fact, 15%–20% of individuals report lingering symptoms beyond three weeks, and up to 38% meet criteria for post-concussion syndrome, with measurable oculomotor deficits such as impaired antisaccades observed months after injury (Johnson, Zhang, Hallett, & Slobounov, 2015). These findings show how athletes may return to play while still experiencing subtle, unresolved impairments that escape current testing protocols. Thus, protocols including a combination of the prevalence history, self-report symptoms, clinical evaluation, and objective measures evaluating vestibular, cognitive, visual, and somatosensory integration, would have greater sensitivity to identify deficits at all stages after a concussion (Capó-Aponte et al., 2012).

In this context, oculomotor metrics offer a promising avenue, as they provide objective, quantifiable measures of brain function that may overcome many of the limitations of traditional concussion assessment methods (Hunfalvay et al., 2019). Visuomotor pathways assessment provide an objective and reliable perspective on brain functioning related to concussions (Hunfalvay et al., 2019; Johnson, Zhang, Hallett, & Slobounov, 2015; Kelly et al., 2019; Kontos et al., 2017). Neural pathways supporting oculomotor functions such as saccades, antisaccades, and optokinetic nystagmus span wide areas of the brain that are particularly vulnerable to concussive injury. Lesions in these circuits can be detected through measurable eye and gaze responses (Hunfalvay et al., 2019; Johnson, Zhang, Hallett, & Slobounov, 2015). For example, dysfunction in optokinetic nystagmus has been linked to patients with chronic concussion or post-concussive syndrome (Cochrane et al., 2019). Importantly, oculomotor testing can capture brain dysfunction even after symptoms are no longer reported and the injury is presumed resolved (Kelly et al., 2019). Thus, technologies assessing visuomotor pathways may support clinicians in building individualized concussion profiles and in tracking recovery across different stages of healing (Kontos et al., 2017; Lumba-Brown et al., 2020).

Oculomotor measures using virtual reality have been explored as tools to generate brain health-related metrics, with normative ranges calculated from general population samples. For example, the NeuroFlex protocol has normative ranges derived from a sample population of healthy individuals (Lunkova et al., 2024). The normative ranges were calculated as the 95% confidence interval for metrics with a distribution resembling normal. For non-normal metrics, the upper range was defined as $Q3 + 1.5IQR$ and the lower range was defined as $Q1 - 1.5IQR$ (Soave et al., 2023). However, athlete-specific normative values are essential, as athletes often perform differently from non-athletes. Research demonstrates that visual-perceptual and visual-cognitive abilities vary depending on the visual demands of each sport, the training of specific visual skills, vestibulo-ocular reflex (VOR) adaptations, and the athlete's level of expertise (Kubitz et al., 2020; Laby & Appelbaum, 2021; Millard et al., 2021; Slobounov et al., 2011). Evidence suggests that athletes frequently show superior performance on several vestibulo-oculomotor measures compared to non-athletes (Kubitz et al., 2020; Laby & Appelbaum, 2021; Lange et al., 2018; Millard et al., 2021; Quintana et al., 2020; Vera et al., 2020). For instance, Laby and Appelbaum (2021) reported that high-achieving athletes displayed better perceptual cue detection, more efficient eye movements, and faster processing than less-skilled athletes or non-athletes. Similarly, Kubitz et al. (2020) found that sport-specific training correlated with enhanced visual-oculomotor skills, noting that elite baseball players outperformed amateurs and non-athletes in measures such as cardinal gaze speed. This "*sport-specific training effect*" highlights how different sports impose unique visual and visuomotor demands that shape performance (Vera et al., 2020).

Although athletes often demonstrate superior visuomotor abilities, research shows they do not always outperform the general population. Performance differences depend on several factors, including the specific oculomotor skill being tested, whether that skill is routinely practiced in a given sport, and the timing of assessment. For example, some oculomotor functions may decline immediately following strenuous exercise (Garner et al., 2023; Zwierko et al., 2019). Importantly, a lack of differences between athletes and non-athletes should not automatically be interpreted as "*typical performance*." Instead, it may reflect underlying issues such as prior trauma, concussion history, or the stage of recovery after an injury. More broadly, variables such as age, handedness, fatigue, central nervous system function, exercise load, and level of expertise should be considered when interpreting oculomotor metrics.

Baseline measures, such as those provided by the NeuroFlex protocol (Lunkova et al., 2024; Soave et al., 2023), offer valuable insight but capture performance only at a single point in time and are shaped by an individual's training history. For athletes, these baselines already reflect sport-specific adaptations and enhancements, which may differ markedly from those of non-athletes. Without comparisons to non-athletes, it becomes difficult to determine whether performance reflects training-related advantages, typical population-level functioning, or deficits associated with injury. Establishing these comparisons is a critical first step toward identifying meaningful differences between populations and developing normative data that are specific to athletes. The subsequent ability to then have the capacity to properly compare athletes to athletes can complete a meaningful comparison with full consideration of contributing factors to results encountered. In concussion management—particularly when eye-tracking assessments are used for return-to-play decisions—athlete-specific benchmarks are essential. They ensure that the unique influences of training, trauma history, and sport demands are accounted for, thereby improving the accuracy, reliability, and clinical utility of oculomotor assessments.

The primary goal of this project is to compare oculomotor health-related metrics between college athletes and the typical adult population. By analysing differences across groups, this study will clarify whether athletic participation influences visuomotor performance and identify the extent to which athletes may demonstrate distinct patterns. A secondary goal is to establish baseline measures specific to athletes, which may serve as a valuable reference point for clinicians and researchers using eye-tracking technologies in the evaluation and monitoring of concussions. Such population-specific normative data are critical for developing accurate, non-invasive, and objective tools for return-to-play decision-making and broader concussion management protocols.

MATERIAL AND METHODS

Participants

A total of 163 university student-athletes between 17 and 30 years of age, 95 female and 68 male, were tested using objective oculomotor and related vestibular metrics protocol. Participants self-disclosed their racial/ethnicities: 54 (33%) Latinx, 27 (16.7%) from the African diaspora, 16 (9.87%) mixed heritage, 40 (24.7%) Caucasian, 9 (5.5%) Asian, 3 (1.9%) Pacific Islander, 1 (0.6%) "*other*", and 12 (7.4%) did not disclose this information.

The student-athlete status was determined per the National Collegiate Athletic Association (NCAA) and California State University, Los Angeles (CSULA) guidelines and included all student-athletes participating in either Fall or Spring sports. Data from two athletes were not included in the study due to data collection difficulties. All student-athletes, regardless of concussion history were recruited as participants. Written

informed consent was obtained prior to the study, and procedures were approved by the CSULA Institutional Review Board.

Measures

The NeuroFlex® software, introduced in 2016, employs virtual reality (VR) technology to capture head and eye responses generating crucial metrics of brain health (Soave et al., 2023). This tool was developed on the premise of overlap between eye movement circuitry and concussion pathophysiology effects on the brain, thus aiming to detect oculomotor impairments post-concussion. Previous research found consistency in concussion identification between NeuroFlex® metrics and functional magnetic resonance imaging (fMRI) (Lunkova et al., 2024). The protocol consists of eight subtests with accompanying software for data analysis on a dashboard with reports within 8-10 minutes. The dashboard normative ranges were generated from a sample of 700 male and female healthy adult participants with no reported history of head injuries (Lunkova et al., 2024). The technology built into the VR headset is non-invasive, including 2D cameras to record each eye position and 3D accelerometers to track the head.

Oculomotor measures protocol

1. *Smooth pursuit in 2D (Head-free)*: In this test, head movement is required to keep the target within visual range ($H \pm 30$ deg, $V \pm 30$ deg). The participant is instructed to keep a 2D moving target centred in the screen by coordinating responses from both the eyes and head.

2. *Smooth pursuit in 2D (Head-fixed)*: In this test, the target remains within ± 20 degrees. The head-free test is a functional representation of real-life movement while the head-fixed test isolates the oculomotor functions to a greater degree.

3. *2D Saccades*: This test evaluates the saccadic system which creates rapid jumps in the eye position to new targets (saccades) with the head fixed. The targets appear randomly within the field of vision (± 25 degrees). The test detects deficits particularly related to a specific target location (e.g., only on the left or right side).

4. *Active Visual Vestibular Ocular Reflex (AVVOR) Horizontal*

5. *Active Visual Vestibular Ocular Reflex (AVVOR) Vertical*

The AVVOR horizontal and vertical assesses the coordination of the eyes with head movement during brief and rapid head motion as the semicircular canals of the inner ear elicit compensatory eye movements to maintain the image on the fovea. This requires coordination of vestibular, visual, and neck proprioception. The presented target will jump left to right (for AVVOR-H protocol), or up and down (for AVVOR-V protocol). The subject will need to move their head rapidly to allow fixation of gaze over large target jumps.

6. *Optokinetic Nystagmus*: This is evaluated by asking the participant to fixate on a target while a visual field of dots provides visual slip stimuli in horizontal or vertical directions at variable velocities. The result is a nystagmus in the direction of the induced visual slip (field velocity) where the individual makes pursuits in the direction of the target movement, and corrective saccades in the opposite direction. The participant must keep the head fixed throughout this test.

7. *Anti-saccades*: Evaluates inhibition (cognitive aspect), or the ability to resist reflexively a stimulus and the ability to re-map peripheral vision. Participants are provided with a distracting target and instructed to look in

the opposite direction. An error is recorded if the saccade is performed in the direction of the target instead of the opposite direction. The target position is randomized between trials to avoid a training effect.

8. Spontaneous Nystagmus: To assess the involuntary movement of the eyes in absence of visual, vestibular, or cognitive stimuli, the participant looks at a flashed target that quickly disappears. Then, the participant is asked to try and preserve the orientation of the eyes in the dark for several seconds. The test identifies any imbalance in the gaze orientation system and whether it is related to a particular sector of the visual field.

Procedure

Athletes were assessed individually (Figure 1). They completed the informed consent approved by the California State University, Los Angeles IRB, and the virtual reality oculomotor protocol (8 subtests). The total time for the completion of the protocols was approximately 10 minutes.



Figure 1. Virtual reality set-up and recording.

Analysis

To quantify the differences between the athletes' baselines and the NeuroFlex normative dataset, two different analysis methods were selected: 1) an unpaired comparison via unpaired t-test or Mann-Whitney U (MWU) test depending on the normality of the metric, and 2) Kolmogorov-Smirnov two-sample test to compare the distribution between the two groups to determine the likelihood of the groups coming from the same distribution. The median sample size used for computing the NeuroFlex® norms across all protocol metrics was 216, whereas 163 for the athletes' metric norms.

Normality for each metric was assessed using the Kolmogorov-Smirnov test and Shapiro-Wilks test for the NeuroFlex athletes' normative groups comparison. Due to our data set containing negative values, the data was used as is without any transformations. Metrics that were not normally distributed used the MWU test for significance testing. Metrics that were normally distributed used the unpaired t-test for significance testing. A *p*-value less than .05 was associated with statistical significance. All statistical tests were performed with Python 3.11.5.

The Kallmann Method was used to compute the normative ranges for the VOMS test, where there are a limited number of possible outcomes for each measured metric. The method follows two main steps: 1) Outlier Removal: The Chauvenet criteria remove data points with a probability of occurrence less than $1/(2 \times \text{sample-size})$. 2) Quantile Selection was used to select 2.5% and 97.5% quantiles as the lower and upper norms to capture 95% of the dataset range.

The Hoffmann method was used to determine normative ranges that achieve higher accuracy and reproducibility than the ones that only depend on quantiles, especially where the outcome measurables are continuous or have many possibilities. The Hoffmann method obtains the 2.5% and 97.5% quantiles from an extrapolation of a linear regression approach and are referred to as the lower and upper thresholds of the normative range.

RESULTS

The virtual reality oculomotor software normative sample data was compared with 163 student-athletes from diverse backgrounds to determine the likelihood of the two groups coming from the same distribution. The unpaired t-tests MWU and the Kolmogorov-Smirnov revealed the metrics that were and were not significantly different between the athletes and the virtual reality oculomotor software normative sample. Overall, 28 out of 44 metrics were significantly different between the two samples (Table 1), indicating that a large proportion of athlete metrics differ from those of the non-athlete population. The result for most of the measures matched, except for two metrics, AVVOR-H gain right 75deg/s and smooth pursuit (head free) number of saccades, in which the unpaired test and Kolmogorov-Smirnov tests did not match. However, for these two metrics, although only one of the two tests was significant, the other was marginally significant.

Table 1. Comparison between NeuroFlex® normative sample distribution with the athletes' sample.

Measures	Unpaired test p-value	KS test p-value
Active Visual VOR (Horizontal) Mean Vergence	0.00 *	0.00 ^
Active Visual VOR (Horizontal) Vergence Standard Deviation	0.45 *	10.09 ^
Active Visual VOR (Horizontal) Gain Right 75deg/s	14.76 *	55.40
Active Visual VOR (Vertical) Mean Vergence	11.36 *	1.56 ^
Active Visual VOR (Vertical) Vergence Standard Deviation	4.34 *	0.10 ^
Active Visual VOR (Vertical) Gain Down Mean	0.00 *	0.02 ^
Active Visual VOR (Vertical) Gain Down 50 deg/s	0.52 *	1.70 ^
Active Visual VOR (Vertical) Gain Down 75 deg/s	0.00 *	0.00 ^
Active Visual VOR (Vertical) Gain Up Mean	0.00 *	0.00 ^
Active Visual VOR (Vertical) Gain Up 25 deg/s	1.33 *	7.90 ^
Active Visual VOR (Vertical) Gain Up 50 deg/s	0.10 *	6.18 ^
Active Visual VOR (Vertical) Gain Up 75 deg/s	0.00 *	0.00 ^
Smooth Pursuit (Head Free) Head Contribution Mean	26.18 *	0.54 ^
Smooth Pursuit (Head Free) Head Contribution Horizontal	7.43 *	5.35 ^
Smooth Pursuit (Head Free) Head Contribution Vertical	35.27 *	0.57 ^
Smooth Pursuit (Head Free) Number of Saccades	58.08	2.22 ^
Optokinetic Nystagmus Gain Left 25 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Left 50 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Right 25 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Right 50 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Down 25 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Down 50 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Up 25 deg/s	0.00 *	0.00 ^
Optokinetic Nystagmus Gain Up 50 deg/s	0.00 *	0.00 ^
Antisaccades Mean Latency	4.75 *	0.47 ^
Antisaccades Acquisition Error	6.24 *	22.71 ^
Antisaccades Mean Vergence	0.00 *	0.00 ^
Saccades Mean Latency	0.00 *	0.00 ^

Note. * = significant differences found with the MWU test. ^ = significant differences found with the KS test.

Table 2 .Descriptive statistics and normative ranges for protocol metrics for all student-athletes using NeuroFlex®.

Protocol name	Sample size	Mean	SD	Min	25%	Median	75%	Max
Active Visual VOR (Horizontal)_Gain Left 25deg/s	163	91.11	9.65	65.03	85.96	92.03	96.69	117.79
Active Visual VOR (Horizontal)_Gain Left 50deg/s	162	89.08	10.57	55.94	82.35	90.13	96.84	114.74
Active Visual VOR (Horizontal)_Gain Left 75deg/s	162	87.08	12.84	45.74	78.85	87.03	96.68	113.93
Active Visual VOR (Horizontal)_Gain Left Mean	163	88.07	10.04	54.8	81.24	88.62	95.72	113.62
Active Visual VOR (Horizontal)_Gain Right 25deg/s	163	92.23	8.84	64.9	87.51	92.47	98.9	118.11
Active Visual VOR (Horizontal)_Gain Right 50deg/s	163	90.2	9.61	52.55	85.24	90.01	96.94	118.02
Active Visual VOR (Horizontal)_Gain Right 75deg/s	161	86.86	11.5	41.45	80.85	87.52	94.21	115.75
Active Visual VOR (Horizontal)_Gain Right Mean	163	87.43	9.25	56	82.27	87.65	93.67	113.72
Active Visual VOR (Horizontal)_Mean Vergence	163	0.34	1.27	-6.92	-0.16	0.37	0.96	3.72
Active Visual VOR (Horizontal)_Vergence Standard Deviation	163	2.31	3.01	0.63	1.09	1.46	2.33	25.2
Active Visual VOR (Vertical)_Gain Down 25 deg/s	163	85.35	9.15	54.84	79.89	86.44	91.26	103.66
Active Visual VOR (Vertical)_Gain Down 50 deg/s	161	85	10.7	45.82	79.68	86.24	91.98	103.55
Active Visual VOR (Vertical)_Gain Down 75 deg/s	147	85.59	11.92	43.26	80.27	87.08	93.17	108.12
Active Visual VOR (Vertical)_Gain Down Mean	163	85.72	9.58	58.83	81.16	86.5	92.22	106.54
Active Visual VOR (Vertical)_Gain Up 25 deg/s	163	82.17	8.76	56.37	77.44	82.63	88.35	102.4
Active Visual VOR (Vertical)_Gain Up 50 deg/s	159	82.87	9.87	49.71	77.3	83.05	89.39	103.12
Active Visual VOR (Vertical)_Gain Up 75 deg/s	137	82.8	12.47	42.91	74.63	84.15	90.25	120.56
Active Visual VOR (Vertical)_Gain Up Mean	163	82.58	9.61	54	76.48	83.63	89.11	102.43
Active Visual VOR (Vertical)_Mean Vergence	163	0.31	1.19	-5.52	-0.27	0.28	0.71	5.9
Active Visual VOR (Vertical)_Vergence Standard Deviation	163	1.16	0.86	0.41	0.78	0.94	1.29	7.72
Antisaccades_Acquisition Error	163	6.29	2.24	2.2	4.74	5.65	7.67	13.21
Antisaccades_Directional Accuracy	163	71.11	22.53	8.33	58.33	75	83.33	100
Antisaccades_Mean Latency	163	451.02	93.02	299.48	389.06	435.67	486.28	772.11
Antisaccades_Mean Vergence	163	0.59	2.09	-6.55	-0.16	0.37	0.99	17.78
Antisaccades_Vergence Standard Deviation	163	2.25	3.01	0.53	1.09	1.43	2.04	22.95
Optokinetic Nystagmus_Gain Down 25 deg/s	163	76.32	13.24	36.47	69.96	79.18	83.84	124
Optokinetic Nystagmus_Gain Down 50 deg/s	143	58.41	16.31	20.14	48.49	58.14	68.04	132.81
Optokinetic Nystagmus_Gain Down Mean	163	70.52	13.21	31.05	63.85	71.49	78.46	119.01
Optokinetic Nystagmus_Gain Left 25 deg/s	162	86.77	9.32	38.71	82.36	87.68	93.01	111.41
Optokinetic Nystagmus_Gain Left 50 deg/s	128	67.09	12.72	19.78	59.14	65.72	76.62	99.56
Optokinetic Nystagmus_Gain Left Mean	162	79.13	10.85	33.69	73.76	80.3	86.89	102.42
Optokinetic Nystagmus_Gain Right 25 deg/s	163	87.47	8.39	43.72	85.07	88.74	92.35	105.31
Optokinetic Nystagmus_Gain Right 50 deg/s	124	70.16	16.89	25.22	61.57	70.2	77.47	180.24
Optokinetic Nystagmus_Gain Right Mean	163	80.72	10.38	41.15	74.33	82.53	87.51	107.2
Optokinetic Nystagmus_Gain Up 25 deg/s	163	76.37	12.21	31.36	70.23	78.48	84.9	101.64
Optokinetic Nystagmus_Gain Up 50 deg/s	134	60.19	14.76	18.14	51.16	61.47	71.88	87.86
Optokinetic Nystagmus_Gain Up Mean	163	70.36	12.57	30.82	63.66	71.85	78.68	103.53
Optokinetic Nystagmus_Mean Vergence	163	-0.88	2.8	-10.03	-2.21	-1.15	-0.07	15.82
Optokinetic Nystagmus_Vergence Standard Deviation	163	3.36	3.62	0.63	1.56	2.06	3.68	27.2
Saccades_Acquisition Error	163	2.37	1.89	1.23	1.68	2.08	2.5	23.98
Saccades_Mean Latency	163	253.03	27.07	201.7	235.8	249.59	265.65	365.03
Saccades_Mean Vergence	163	0.23	1.12	-3.14	-0.34	0.13	0.61	7.56
Saccades_Vergence Standard Deviation	163	1.62	1.67	0.54	0.87	1.2	1.64	13.83
Smooth Pursuit (Head Fixed)_Mean Error	163	2.4	1.54	1.3	1.77	2.06	2.44	18.02
Smooth Pursuit (Head Fixed)_Mean Vergence	163	0.72	1.51	-7.94	0.08	0.78	1.37	6.74
Smooth Pursuit (Head Fixed)_Number of Saccades	163	11.53	9.4	1	5	10	15	57
Smooth Pursuit (Head Fixed)_Vergence Standard Deviation	163	2.24	2.4	0.64	1.22	1.55	2.14	17.36
Smooth Pursuit (Head Free)_Head Contribution Horizontal	163	93.41	9.19	48.07	89.69	94.7	99.36	111.32
Smooth Pursuit (Head Free)_Head Contribution Mean	163	92.89	10.69	37.5	89.01	93.44	99.55	113.27
Smooth Pursuit (Head Free)_Head Contribution Vertical	163	91.47	14.9	21.6	83.8	92.85	99.72	123.21
Smooth Pursuit (Head Free)_Mean Error	163	2.55	0.59	1.71	2.22	2.45	2.74	7.04
Smooth Pursuit (Head Free)_Mean Vergence	163	0.19	1.35	-3.47	-0.45	0.09	0.56	10.47
Smooth Pursuit (Head Free)_Number of Saccades	163	11.86	8.32	1	5	9	18	37
Smooth Pursuit (Head Free)_Vergence Standard Deviation	163	1.24	1.53	0.33	0.62	0.88	1.32	12.13
Spontaneous Nystagmus_Average Drift Horizontal	163	0	0.01	-0.04	0	0	0	0.05
Spontaneous Nystagmus_Average Drift Mean	163	0.01	0.01	0	0	0.01	0.01	0.05
Spontaneous Nystagmus_Average Drift Vertical	163	0	0	-0.01	0	0	0	0.02
Spontaneous Nystagmus_Mean Tremor Frequency	163	0.26	0.19	0	0.12	0.21	0.37	1.01
Spontaneous Nystagmus_Mean Tremor Velocity Horizontal	163	44.11	27.9	0	26.3	40.11	57.23	133.63
Spontaneous Nystagmus_Mean Tremor Velocity Mean	163	51.8	31.9	0	31.13	46.26	64.92	176.6
Spontaneous Nystagmus_Mean Tremor Velocity Vertical	163	22.58	21.66	0	8.06	16.13	30.57	175.13

For the calculation of normative ranges, some students did not perform on all subtests; consequently, each metric was based on data from a minimum of 124 students. Relevant descriptive statistics for all participants (irrespective of gender or team) and the cohort normative ranges are reported in Table 2. These resulting

normative ranges can serve as more precise comparison values for athletes, offering an alternative to the normative dashboard values from the NeuroFlex system. Unlike the dashboard data, which is based on measurements from the general population, these ranges resulting from athletes' performance, are better suited to account for the unique performance characteristics and physiological adaptations of this group.

DISCUSSION

Oculomotor measures offer robust, objective indicators of brain function and are particularly valuable for assessing concussions in athletes. Because visual information integrates with attention, cognitive control, memory, and motor systems to generate the three-dimensional representations that guide perception and action, disruptions in oculomotor performance can provide insight into neurological function (Laby & Appelbaum, 2021). Eye-tracking thus reflects more than how an athlete "sees and reacts"; it reveals how the brain is processing and coordinating information. Prior research demonstrates the clinical utility of these measures: oculomotor assessments have shown up to 90% accuracy in differentiating healthy from concussed athletes and in monitoring recovery (Cheever et al., 2018). Virtual reality–based oculomotor tools further enhance this approach by capturing head and eye responses to generate brain health metrics, facilitating concussion identification even when subjective reports are limited. Importantly, establishing individual athlete baselines allows clinicians to detect changes quickly, guide rehabilitation, and support safe return-to-play decisions (Capó-Aponte et al., 2012). Collectively, these findings reinforce the convenience and clinical importance of oculomotor assessments in concussion management for athletes.

Athletes, however, represent a unique population whose sport-specific training can shape oculomotor performance. This raises an important clinical question: can their performance be fairly compared to that of non-athletes when determining thresholds for potential brain injury? In our study, we found significant differences in the average scores of 28 oculomotor measures between athletes and the normative non-athlete sample. These results suggest that athletes do not perform at the same level as the general adult population in head and eye movement tasks, reinforcing the need for athlete-specific normative standards. Notably, significant group differences emerged in AVVOR, smooth pursuit, optokinetic nystagmus, and antisaccades—findings consistent with prior research. For example, Vera et al. (2020) reported superior visual discrimination in basketball players compared to non-athletes, while Kubitz et al. (2020) showed that baseball players demonstrated significantly faster gaze speed than amateurs or non-athletes. Similarly, Zwierko et al. (2019) found enhanced visuomotor performance, as measured by saccadic tasks, in soccer players. Together, these findings suggest that athletes outperform non-athletes in visuomotor skills such as saccades, accommodation speed, and convergence—likely as a result of repetitive practice and sport-specific demands. Consequently, comparing athletes against non-athlete norms risks misinterpretation of clinical outcomes. Ideally, normative data should not only be athlete-specific but further refined by sport type, given the distinct training regimens, visual demands, and cognitive processing requirements inherent in different sports (Laby & Appelbaum, 2021).

While previous research has demonstrated that oculomotor functions can be enhanced through training (Millard et al., 2021), the specific source of athletes' superior performance remains uncertain. It is not yet clear whether their advantages over non-athletes arise from enhanced visual search strategies, faster information processing, improved acuity, or advanced oculomotor control. Further work is needed to disentangle whether these benefits reflect general physiological factors, such as increased blood flow and central nervous system efficiency (Lange et al., 2018), or whether they are primarily shaped by the unique demands of different sports. Importantly, athletes do not always outperform non-athletes. For instance, oculomotor performance has been shown to decline immediately after strenuous activity, with decrements in

saccade duration and velocity (Zwierko et al., 2019), demonstrating that not all metrics measured between athletes and non-athletes were found to be significantly different. This suggests that fatigue can negatively impact visuomotor skills, potentially increasing injury risk and complicating assessment outcomes. Because these conditions are specific to the athletic population, normative data tailored to athletes are essential for more accurate clinical interpretation. Our study contributes to this effort by generating preliminary athlete-specific normative ranges, which—although not yet sport-specific—provide a more reliable framework for clinical use than comparisons to non-athlete norms, which risk introducing bias.

Concussions are complex injuries that affect individuals differently, requiring careful monitoring from the acute phase through recovery (Cavanaugh et al., 2005). Virtual reality–based oculomotor protocols provide an effective means of capturing objective measures, tracking change over time, and offering meaningful baseline comparisons for athletes. However, to fully support clinical decision-making, it is essential to move beyond the assumption that athletes will outperform non-athletes. Therefore, having non-athletes as the minimum and only threshold to surpass places athletes at a disadvantage because their performance and results should be evaluated based on their sport, experience level, and their own personal baseline. This further substantiates the need for visual and oculomotor abilities to be assessed with population-specific benchmarks (Soave et al., 2023). Athletes' training often enhances sport-specific visuospatial and oculomotor skills, yet their performance can also be influenced by fatigue, concussion history, and the unique demands of their sport. Understanding these distinctions is critical for interpreting test results accurately. By identifying systematic differences between athletes and non-athletes and generating normative ranges specific to the athletic population, our study provides an important step toward improving concussion assessment. Athlete-specific norms offer clinicians a more reliable and contextually appropriate framework for evaluating neurological function, guiding return-to-play decisions, and ultimately enhancing both the care and recovery of athletes.

CONCLUSIONS

This study suggests that athletes differ significantly from non-athletes in multiple oculomotor metrics, underscoring the need for population-specific normative ranges in clinical concussion assessment. While oculomotor measures offer robust and objective indicators of brain function, interpreting them against general population data risks overlooking enhancements linked to sport-specific training as well as potential deficits arising from injury, fatigue, or other sport-related conditions. Our findings contribute preliminary normative values for athletes, providing clinicians with a more reliable framework for evaluating concussion-related changes in this population. Establishing athlete-specific benchmarks is essential for accurate diagnosis, monitoring recovery, and guiding return-to-play decisions, as well as for advancing our understanding of how sport training and neurological health interact. Further research, including larger and sport-specific samples, is warranted to refine these norms and to explore the mechanisms underlying both enhanced and diminished oculomotor performance in athletes.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. Contributions are specified according to the CRediT (Contributor Roles Taxonomy) as follows:

Conceptualization: [Clea Tucker] — Developed the overall concept and design of the study. Methodology: [Clea Tucker] — Designed the methodology and conducted the experiments. Software: [Arianna Soave, Nafiseh Ghoroghchian, William Pei] — Developed or used software tools for data analysis. Validation: [Clea

Tucker] — Verified the results and the overall accuracy of the study. Formal Analysis: [Clea Tucker, Beatriz Barragan] — Analysed the data and interpreted results. Investigation: [Clea Tucker] — Conducted the research and collected data. Resources: [Clea Tucker] — Provided the resources or facilities needed for the study. Data Curation: [Clea Tucker] — Managed and curated the data for analysis. Writing – Original Draft: [Clea Tucker, Beatriz Barragan] — Wrote the initial draft of the manuscript. Writing – Review & Editing: [Clea Tucker, Beatriz Barragan] — Critically reviewed and edited the manuscript. Visualization: [Arianna Soave, Nafiseh Ghoroghchian, William Pei] — Created visual representations of the data. Supervision: [Clea Tucker] — Oversaw the research and provided guidance. Project Administration: [Clea Tucker] — Managed the project and organized the workflow. Funding Acquisition: [Clea Tucker] — Secured funding and financial support for the project. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

We would like to thank the Cal State LA Office of Research, Scholarship, and Creative Activities for the mini-grant approved in support of this project.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

ACKNOWLEDGMENTS

We would like to thank Saccade Analytics, Inc. for providing us with VR equipment, software (NeuroFlex®), and normative data for oculomotor function assessment. We would like to thank all Cal State LA athletes who volunteered to be participants for this study and the Sports Medicine Team for assisting in this project.

REFERENCES

- Broglio, S. P., Sosnoff, J. J., & Ferrara, M. S. (2009). The relationship of athlete-reported concussion symptoms and objective measures of neurocognitive function and postural control. *Clinical Journal of Sport Medicine*, 19(5), 377. <https://doi.org/10.1097/JSM.0b013e3181b625fe>
- Capó-Aponte, J. E., Tarbett, A. K., Urosevich, T. G., Temme, L. A., Sanghera, N. K., & Kalich, M. E. (2012). Effectiveness of computerized oculomotor vision screening in a military population: Pilot study. *Military Medicine*, 177(Suppl. 9), 27-34. <https://doi.org/10.1682/JRRD.2011.07.0128>
- Cavanaugh, J. T., Guskiewicz, K. M., Giuliani, C., Marshall, S., Mercer, V., & Stergiou, N. (2005). Detecting altered postural control after cerebral concussion in athletes with normal postural stability. *British Journal of Sports Medicine*, 39(11), 805-811. <https://doi.org/10.1136/bjsm.2004.015909>

- Centers for Disease Control and Prevention. (2019). Surveillance report of traumatic brain injury-related emergency department visits, hospitalizations, and deaths-United States, 2014. U.S. Department of Health and Human Services. Retrieved from [Accessed 2026, 23 April]: <https://stacks.cdc.gov/view/cdc/78062>
- Cheever, K., McDevitt, J., Tierney, R., & Wright, W. (2018). Concussion recovery phase affects vestibular and oculomotor symptom provocation. *International Journal of Sports Medicine*, 39(2), 141-147. <https://doi.org/10.1055/s-0043-118339>
- Chen, Y., Herrold, A. A., Gallagher, V. T., Reilly, J. L., Parrish, T., & Breiter, H. C. (2019). Cutting to the pathophysiology chase: Translating cutting-edge neuroscience to rehabilitation practice in sports-related concussion management. *Journal of Orthopaedic & Sports Physical Therapy*, 49(11), 811-818. <https://doi.org/10.2519/jospt.2019.8884>
- Choe, M. C. (2016). The pathophysiology of concussion. *Current Pain and Headache Reports*, 20(6), Article 42. <https://doi.org/10.1007/s11916-016-0573-9>
- Cochrane, G. D., Christy, J. B., Almutairi, A., Busettini, C., Swanson, M. W., & Weise, K. K. (2019). Visuo-oculomotor function and reaction times in athletes with and without concussion. *Optometry and Vision Science*, 96(4), 256-265. <https://doi.org/10.1097/OPX.0000000000001364>
- Dams-O'Connor, K., Cantor, J. B., Brown, M., Dijkers, M. P., Spielman, L. A., & Gordon, W. A. (2014). Screening for traumatic brain injury: Findings and public health implications. *The Journal of Head Trauma Rehabilitation*, 29(6), 479-489. <https://doi.org/10.1097/HTR.0000000000000099>
- Dorman, J. C., Valentine, V. D., Munce, T. A., Tjarks, B. J., Thompson, P. A., & Bergeron, M. F. (2015). Tracking postural stability of young concussion patients using dual-task interference. *Journal of Science and Medicine in Sport*, 18(1), 2-7. <https://doi.org/10.1016/j.jsams.2013.11.010>
- Finnoff, J. T., Peterson, V. J., Hollman, J. H., & Smith, J. (2009). Intrarater and interrater reliability of the Balance Error Scoring System (BESS). *PM&R*, 1(1), 50-54. <https://doi.org/10.1016/j.pmri.2008.06.002>
- Garner, D. P., Sparks, P. D., Chizuk, H. M., & Haider, M. N. (2023). Effect of exertion on blink reflex parameters in Division I football athletes. *Cogent Engineering*, 10(1), Article 2232166. <https://doi.org/10.1080/23311916.2023.2232166>
- Hunfalvay, M., Roberts, C.-M., Murray, N., Tyagi, A., Kelly, H., & Bolte, T. (2019). Horizontal and vertical self-paced saccades as a diagnostic marker of traumatic brain injury. *Concussion*, 4(1), Article CNC60. <https://doi.org/10.2217/cnc-2019-0001>
- Johnson, B., Zhang, K., Hallett, M., & Slobounov, S. (2015). Functional neuroimaging of acute oculomotor deficits in concussed athletes. *Brain Imaging and Behavior*, 9(3), 564-573. <https://doi.org/10.1007/s11682-014-9316-x>
- Kelly, K. M., Kiderman, A., Akhavan, S., Quigley, M. R., Snell, E. D., Happ, E., Synowiec, A. S., Miller, E. R., Bauer, M. A., Oakes, L. P., Eydelman, Y., Gallagher, C. W., Dinehart, T., Schroeder, J. H., & Ashmore, R. C. (2019). Oculomotor, vestibular, and reaction time effects of sports-related concussion: Video-oculography in assessing sports-related concussion. *Journal of Head Trauma Rehabilitation*, 34(3), 176-188. <https://doi.org/10.1097/HTR.0000000000000437>
- Kontos, A. P., Huppert, T. J., Beluk, N. H., Elbin, R. J., Henry, L. C., French, J., Dakan, S. M., & Collins, M. W. (2014). Brain activation during neurocognitive testing using functional near-infrared spectroscopy in patients following concussion compared to healthy controls. *Brain Imaging and Behavior*, 8(4), 621-634. <https://doi.org/10.1007/s11682-014-9289-9>
- Kontos, A. P., Deitrick, J. M., Collins, M. W., & Mucha, A. (2017). Review of vestibular and oculomotor screening and concussion rehabilitation. *Journal of Athletic Training*, 52(3), 256-261. <https://doi.org/10.4085/1062-6050-51.11.05>

- Kontos, A. P., Sufrinko, A., Sandel, N., Emami, K., & Collins, M. W. (2019). Sport-related concussion clinical profiles: Clinical characteristics, targeted treatments, and preliminary evidence. *Current Sports Medicine Reports*, 18(3), 82-92. <https://doi.org/10.1249/JSR.0000000000000573>
- Kroshus, E., Garnett, B., Hawrilenko, M., Baugh, C. M., & Calzo, J. P. (2015). Concussion under-reporting and pressure from coaches, teammates, fans, and parents. *Social Science & Medicine*, 134, 66-75. <https://doi.org/10.1016/j.socscimed.2015.04.011>
- Kubitz, K., Roberts, C.-M., Hunfalvay, M., & Murray, N. (2020). A comparison of cardinal gaze speed between Major League Baseball players, amateur prospects, and non-athletes. *Journal of Sports and Performance Vision*, 2(1), e17-e28. <https://doi.org/10.22374/jspv.v2i1.5>
- Laby, D. M., & Appelbaum, L. G. (2021). Review: Vision and on-field performance: A critical review of visual assessment and training studies with athletes. *Optometry and Vision Science*, 98(7), 723-731. <https://doi.org/10.1097/OPX.0000000000001729>
- Lange, B., Hunfalvay, M., Murray, N., & Roberts, C.-M. (2018). Reliability of computerized eye-tracking reaction time tests in non-athletes, athletes, and individuals with traumatic brain injury. *Optometry & Visual Performance*, 6(3), 108-120.
- Lunkova, E., Chen, J.-K., Saluja, R. S., & Ptito, A. (2024). Assessment of oculomotor functions as a biomarker in mild traumatic brain injury. *Neurotrauma Reports*, 5(1), 628-639. <https://doi.org/10.1089/neur.2024.0018>
- Lunkova, E., McCabe, S., Chen, J.-K., Saluja, R. S., & Ptito, A. (2024). Exploring oculomotor functions in a pilot study with healthy controls: Insights from eye-tracking and fMRI. *PLOS ONE*, 19(6), Article e0303596. <https://doi.org/10.1371/journal.pone.0303596>
- Lumba-Brown, A., Teramoto, M., Bloom, O. J., Brody, D., Chesnutt, J., Clugston, J. R., Collins, M., Gioia, G., Kontos, A., Lal, A., Sills, A., & Ghajar, J. (2020). Concussion guidelines step 2: Evidence for subtype classification. *Neurosurgery*, 86(1), 2-13. <https://doi.org/10.1093/neuros/nyz332>
- McLeod, T. C. V., & Leach, C. (2012). Psychometric properties of self-report concussion scales and checklists. *Journal of Athletic Training*, 47(2), 221-223. <https://doi.org/10.4085/1062-6050-47.2.221>
- Millard, L., Shaw, I., Breukelman, G. J., & Shaw, B. S. (2021). Differences in visio-spatial expertise between 1st division rugby players and non-athletes. *Heliyon*, 7(2), Article e06290. <https://doi.org/10.1016/j.heliyon.2021.e06290>
- Nauman, E. A., Talavage, T. M., & Auerbach, P. S. (2020). Mitigating the consequences of subconcussive head injuries. *Annual Review of Biomedical Engineering*, 22, 387-407. <https://doi.org/10.1146/annurev-bioeng-091219-053447>
- Pearce, A. J., Tommerdahl, M., & King, D. A. (2019). Neurophysiological abnormalities in individuals with persistent post-concussion symptoms. *Neuroscience*, 408, 272-281. <https://doi.org/10.1016/j.neuroscience.2019.04.019>
- Quintana, C., Heebner, N. R., Olson, A. D., Abt, J. P., & Hoch, M. C. (2020). Sport-specific differences in dynamic visual acuity and gaze stabilization in Division I collegiate athletes. *Journal of Vestibular Research*, 30(4), 249-257. <https://doi.org/10.3233/VES-200710>
- Slobounov, S. M., Gay, M., Zhang, K., Johnson, B., Pennell, D., Sebastianelli, W., Horovitz, S., & Hallett, M. (2011). Alteration of brain functional network at rest and in response to YMCA physical stress test in concussed athletes: RsfMRI study. *NeuroImage*, 55(4), 1716-1727. <https://doi.org/10.1016/j.neuroimage.2011.01.024>
- Soave, A., Tantin, A., Lachmann, F., & Galiana, M. (2023, July). Establishing normative ranges for students using NeuroFlex®-A virtual reality based software [Poster presentation]. *International Society for Virtual Rehabilitation Conference*, Montreal, QC, Canada.

- Tucker, C., & Rivera, Y. (2021). Subtleties in pathophysiology and function post-concussion that may be captured in a multimodal assessment [Conference session]. American Association of Sport Physical Therapy Annual Meeting & Scientific Conference, Philadelphia, PA, United States (Virtual).
- Vera, J., Jiménez, R., Cárdenas, D., Redondo, B., & García, J. A. (2020). Visual function, performance, and processing of basketball players vs. sedentary individuals. *Journal of Sport and Health Science*, 9(6), 587-594. <https://doi.org/10.1016/j.jshs.2017.05.001>
- Webb, E. K., Etter, J. A., & Kwasa, J. A. (2022). Addressing racial and phenotypic bias in human neuroscience methods. *Nature Neuroscience*, 25(4), 410-414. <https://doi.org/10.1038/s41593-022-01046-0>
- Zwierko, T., Jedziniak, W., Florkiewicz, B., Stępiński, M., Buryta, R., Kostrzewa-Nowak, D., Nowak, R., Popowczak, M., & Woźniak, J. (2019). Oculomotor dynamics in skilled soccer players: The effects of sport expertise and strenuous physical effort. *European Journal of Sport Science*, 19(5), 612-620. <https://doi.org/10.1080/17461391.2018.1538391>



Analysis of research trends in para-athletics from 2015 to 2025: Bibliometric review

 **Eduardo Antonio Reinoso González**  . Department of Kinesiology. Faculty of Medicine. University of Concepcion. Concepción, Chile.

 **Mauricio Godoy Rencoret**. Department of Clinical and Preclinical Sciences. Faculty of Medicine. Catholic University of the Holy Conception. Concepción, Chile.

ABSTRACT

Introduction: para-athletics is an adaptation of athletics for athletes with disabilities and has experienced notable interest in recent years. This expansion has generated increasing research output across disciplines related to sport, rehabilitation, and performance. **Objective:** To analyse global research trends in para-athletics between 2015 and 2025 by examining publication output, keywords, journals, countries, institutions, and emerging research topics. **Methods:** A bibliometric review was conducted using publications indexed in the Web of Science Core Collection (WoSCC). Only articles published between 2015 and 2025 were included. Bibliometric indicators were analysed using Microsoft Excel (v.2504) and VOSviewer (v.1.6.11) to examine publication patterns, collaboration networks, keyword co-occurrence, and thematic development. **Results:** A total of 1,237 articles were identified. Scientific production increased steadily between 2015 and 2022, followed by a decline in 2023. USA, England, and China showed the highest research output. The most productive journals were Journal of Athletic Training, American Journal of Sports Medicine, and Human Sport Medicine. The most frequent keywords were performance, exercise, and sport, indicating a strong focus on performance. **Conclusion:** Research in para-athletics has expanded significantly over the past decade, revealing global publication patterns and dominant research themes that may guide future studies in the field.

Keywords: Para-athletics, Adapted sport, Para-sport, Inclusive sport, Disability sport.

Cite this article as:

Reinoso González, E. A., & Godoy Rencoret, M. (2026). Analysis of research trends in para-athletics from 2015 to 2025: Bibliometric review. *Scientific Journal of Sport and Performance*, 5(3), 530-541. <https://doi.org/10.55860/GASR9396>

 **Corresponding author.** Department of Kinesiology. Faculty of Medicine. University of Concepción. Concepción, Chile.

E-mail: eduareinoso@udec.cl

Submitted for publication March 09, 2026.

Accepted for publication April 23, 2026.

Published May 07, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/GASR9396>

INTRODUCTION

Para-athletics is a sport for people with disabilities who must be eligible according to a classification process based on their disability (Van Biesen et al., 2021; van der Linden et al., 2021). Participation focuses on athletes with a physical, visual, or cognitive impairment. It is based on a variety of competitions that present physical challenges, where athletes participate in running, jumping, and throwing events. para-athletics is currently practiced by athletes with disabilities in various settings, including recreational, developmental, and competitive levels. The competitive aspect is most prominent at the high-performance level, and the most prestigious event is considered to be the Paralympic Games, with approximately 179 countries participating.

The educational and recreational development of para-athletics offers numerous advantages for people with disabilities, generating personal, social, health, and rehabilitation benefits, among other aspects encompassed by adapted sports (Sakalidis et al., 2023). However, high performance does not guarantee participation for all people with disabilities. This is because, to participate in the competitive arena, they must undergo a classification process regulated by World Para-athletics (IPC), which aims to ensure competitive equity across the spectrum of disability. Therefore, athletes in the competitive arena, and especially at the high-performance level, prioritize athletic performance over any other condition of the eligible athlete (Marques & Alves, 2021).

In recent years, para-athletics has experienced significant growth since its beginnings in the 1960s in Rome, Italy, culminating in the last major event held in Paris, France in 2024. Similarly, research in para-athletics, due to its divergence between recreational, educational, and competitive aspects, has generated new lines of development and a diversity of topics to be addressed (Heffernan, 2022). In this study, the objective was to investigate the literature on para-athletics, evaluating the impact of countries, institutions, authors, journals, and topics. The findings of this study can provide comprehensive information and suggestions on global trends in para-athletics research and future research directions.

MATERIALS AND METHODS

Bibliometric indicators included annual publication output, country productivity, journal distribution, author productivity, research categories, and keyword co-occurrence networks.

Database source

The articles published were recovered from a topic search in the Web of Science Core Collection with the search topic (TS = (Para athletic* OR Athletic* in persons with disability OR athletic* + disability OR Athletic* + Paraspport* OR special + athletic* OR unified + athletic* OR inclusive + athletic* OR Deaflympic* + athletic* OR athletic* + for + disabled OR Athletic* + for + handicap OR Athletic* + for + the + impaired OR athletics + for + physical + disability OR athletics + for + sensory + disability OR Athletics + for + special + needs OR blinds + athletics Or deaf + Athletics OR wheelchair + athletic*)) AND (PY = ("2015" OR "2016" OR "2017" OR "2018" OR "2019" OR "2020" OR "2021" OR "2022" OR "2023" OR "2024" OR "2025")).

Inclusion criteria

Studies designated as "articles" were considered for inclusion. Full records and cited references from publications were imported from Web of Science. Science Core Collection (WoSCC). Only literature published in English was included. If a publication was attributed to more than one country/region or institution, the information from the corresponding authors was considered.

Exclusion criteria

Other types of documents were excluded, such as retractions of publication (3) and Proceeding Papers (9), and book chapters (17). 1271 results were found, of which 34 articles were excluded because, after review, they were not directly related to the search topics, leaving 1237 articles for analysis.

Statistical analysis

VOSviewer 1.6.20 [6] and Excel version 2504 (compilatio 18730.20186) were used to identify and visualize critical information from the publications. VOSviewer was used for network visualisation maps were generated to analyse collaboration patterns between countries and authors, as well as keyword co-occurrence clusters representing major research themes.

RESULTS

WoSCC database, according to the search strategy mentioned above. Figure 1 shows the publication output on para-athletics research over the last 10 years (2015–2025). The number of annual publications grew steadily from 2015 to 2022, with a decrease in 2023. The majority of the research was published in 2024 ($n = 176$ publications, 14.20%). It is worth noting that the challenge for this year is to increase para-athletics research, as only 60 articles have been reported to date through the end of May 2025. Table 1.

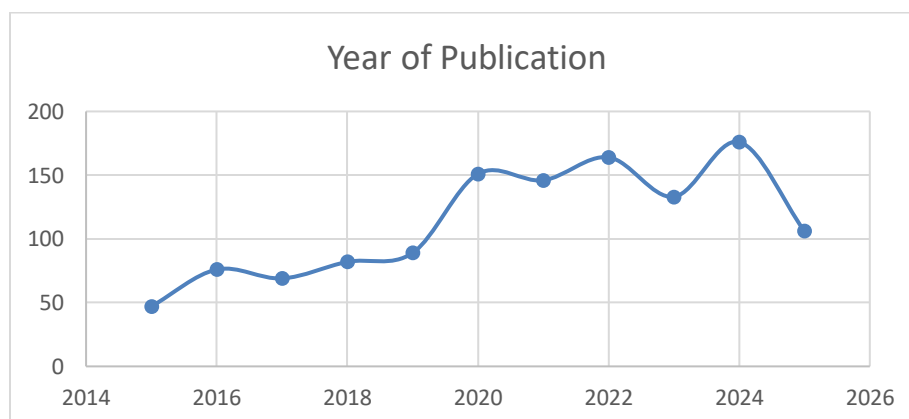


Figure 1. Annual number of publications per year.

Table 1. Annual number of publications and percentages

Year	No.	%
2025	106	8.555
2024	176	14.205
2023	133	10.734
2022	164	13.236
2021	146	11.784
2020	151	12.187
2019	89	7.183
2018	82	6.618
2017	69	5.569
2016	76	6.134
2015	47	3.793

Note. $n = 1239$ items.

According to the corresponding author's country, a total of 84 countries/regions have contributed to Para-athletics research over the past decade. Between 2015 and 2025, 34 countries contributed more than 10 articles to para-athletics research. The top ten countries/regions (by corresponding author's country) are led by North America, comprised of the United States and Canada, with the United States leading the way with 440 articles. This is followed by Europe, comprised of England, Italy, Germany, Spain, and the Netherlands, with five countries contributing a total of 288 articles. All other regions have only one representative in the top ten. South America is represented by Brazil with 89 articles, Asia by China with 87, and Oceania by Australia with 85.

Figure 2 shows 34 leading countries/regions across different time periods from 2015 to 2025 that have achieved 10 publications in the last 10 years. Table 2 presents the 83 countries/regions that have published on para-athletics.

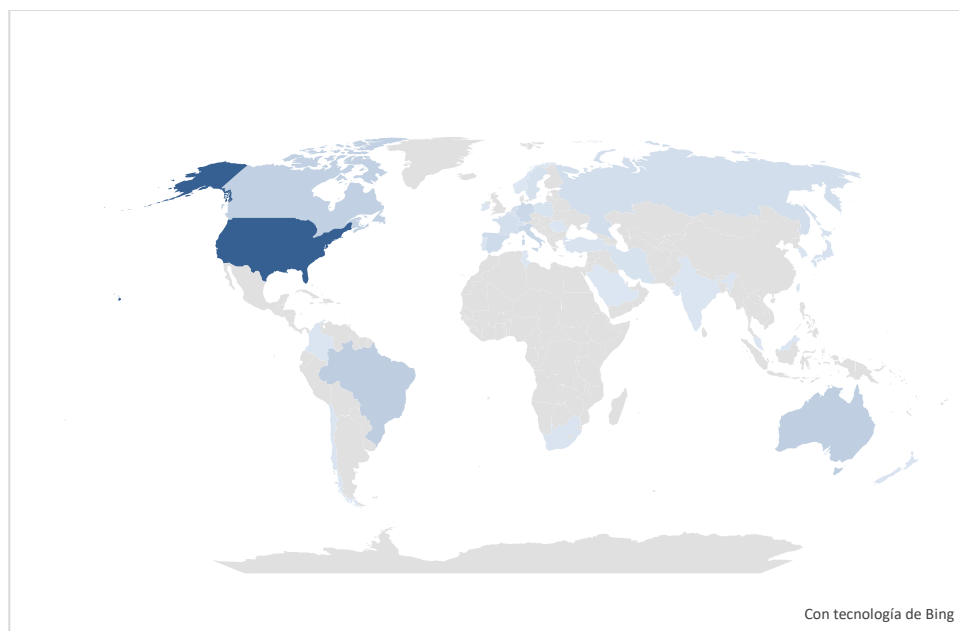


Figure 2. Map of publications by country/region that published over 10 articles in the last decade.

Table 2. Number of publications per country/region on 10 articles in the last decade.

Country	No.	%
United States	452	36.481
England	95	7.667
China	93	7.506
Brazil	91	7.345
Australia	87	7.022
Canada	82	6.618
Italy	59	4.762
Germany	57	4.600
Spain	47	3.793
Netherlands	37	2.986
Russia	37	2.986
France	34	2.744

Japan	28	2.260
Sweden	27	2.179
Switzerland	25	2018
Iran	24	1937
Belgium	22	1.776
Turkey	21	1.695
Poland	20	1.614
Türkiye	19	1.533
Portugal	18	1.453
South Africa	17	1.372
New Zealand	16	1.291
South Korea	16	1.291
Taiwan	16	1.291
Chili	14	1.130
Scotland	14	1.130
India	13	1.049
Ireland	13	1.049
Colombia	12	0.969
Romania	12	0.969
Norway	11	0.888
Saudi Arabia	11	0.888
Malaysia	10	0.807
Tunisia	10	0.807

Note. n = 84 countries.

Para-athletics was published in 556 academic journals. The top ten journals for the last 10 years and each five-year period are presented in Table 3. Over the last 10 years, the Journal of Athletic Training (n = 42 articles) published the most research about para-athletics, the following These journals were American Journal of Sports Medicine (n = 22), Human Sport Medicine (n = 22), PLOS ONE (n = 20), Sport in Society (n = 18), Journal of Sports Medicine and Physical Fitness (n = 17), Orthopaedic Journal of Sports Medicine (n = 17), International Journal of Environmental Research and Public Health (n = 16), Revista Brasileira de Medicina do Esporte (n = 16), and Journal of Sport Rehabilitation (n = 15). Most of these journals are linked to sports science and medicine.

Table 3. The top 10 journals that published articles on research in para-athletics

Publication titles	Number of records	%
Journal Athletic Training	42	3.395
American Journal of Sports Medicine	22	1.778
Human Sport Medicine	22	1.778
Plos One	20	1.617
Sport in Society	18	1.455
Journal of Sports Medicine and Physical Fitness	17	1.374
Orthopedic Journal of Sports Medicine	17	1.374
International Journal of Environmental Research and Public Health	16	1.293
Sports Medicine Magazine	16	1.293
Journal of Sport Rehabilitation	15	1.213

Note. n = 556.

An analysis by categories was carried out in which 101 were found in total, to simplify the visualization in Table 4 the 10 main ones are shown: Sport Sciences (n = 515), Social Sciences Other Topics (n = 155), Orthopedics (n = 145), Rehabilitation (n = 118), Physiology (n = 106), General Internal Medicine (n = 72), Education Educational Research (n = 65), Physiology (n = 64), Neurosciences Neurology (n = 57), Science Technology Other Topics (n = 49).

Table 4. The top 10 categories in which articles on research in para-athletics were published.

Web of Science categories	Number of records	%
Sport sciences	515	41.633
Social sciences other topics	155	12.530
Orthopedics	145	11.722
Rehabilitation	118	9.538
Physiology	106	8.569
General internal medicine	72	5.821
Education educational research	65	5.255
Physiology	64	5.174
Neurosciences neurology	57	4.608
Science technology other topics	49	3.961

Note. n = 101.

An analysis was performed by the most relevant author in which 5322 were found in total, to simplify the visualization in Table 5 the 10 main authors are shown Winckler C (n = 10), De Mello MT (9), Silva A (n = 9), Neto FR (n = 8), Connick MJ (n = 7), Costa RRG (7), Valier ARS (n = 7), Burke LM (n = 6), Derman W (n = 6), Tweedy SM (n = 6).

Table 5. The top 10 authors who published articles on research in para-athletics.

Author	Number of records	%
Winckler Ciro	10	0.815
De Mello MT	9	0.733
Silva A	9	0.733
Net FR	8	0.652
Connick MJ	7	0.570
Costa RRG	7	0.570
Valier ARS	7	0.570
Burke LM	6	0.489
Derman W	6	0.489
Tweedy SM	6	0.489

Note. n = 5322.

The ten keywords with the highest occurrence frequencies were: performance (n = 156), exercise (n = 135), sport (n = 109), Athletic performance (n = 106), disability (n = 104), Sports (n = 103), children (n = 79), strength (n = 68), health (n = 67) and participation (n = 58).

A keyword clustering analysis was performed using VOSviewer (Figure 3). With a threshold of ≥ 10 occurrences, 158 keywords were selected, and four clusters emerged through co-occurrence clustering analysis.

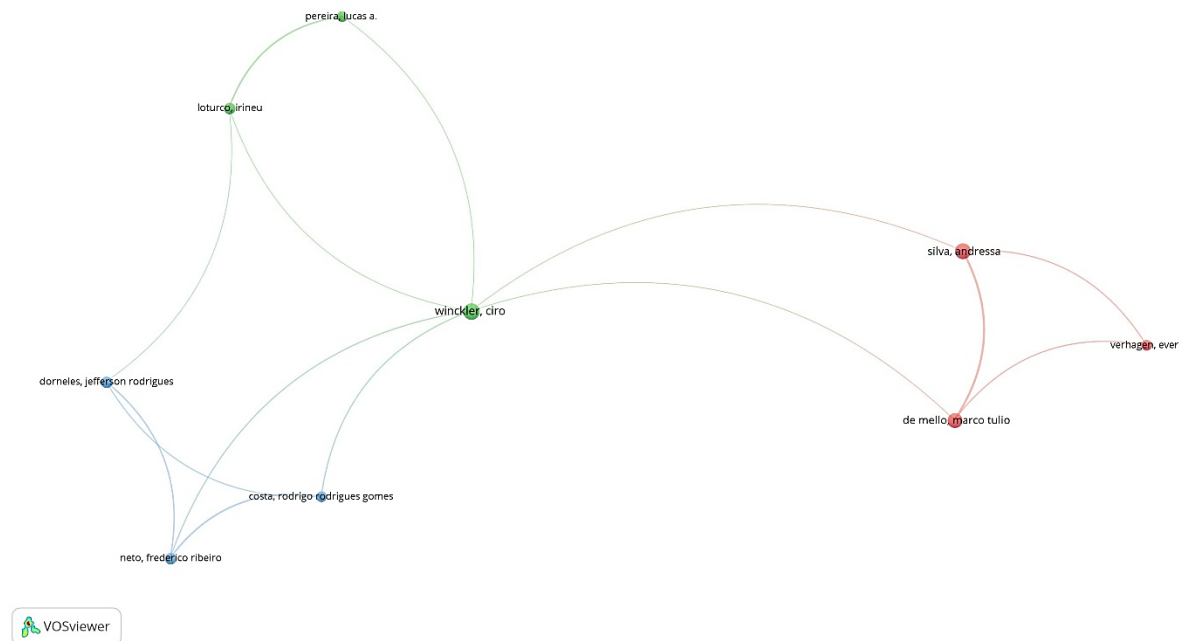


Figure 3. Top 10 authors who have published articles on Para Athletics research, co-occurrence analysis.

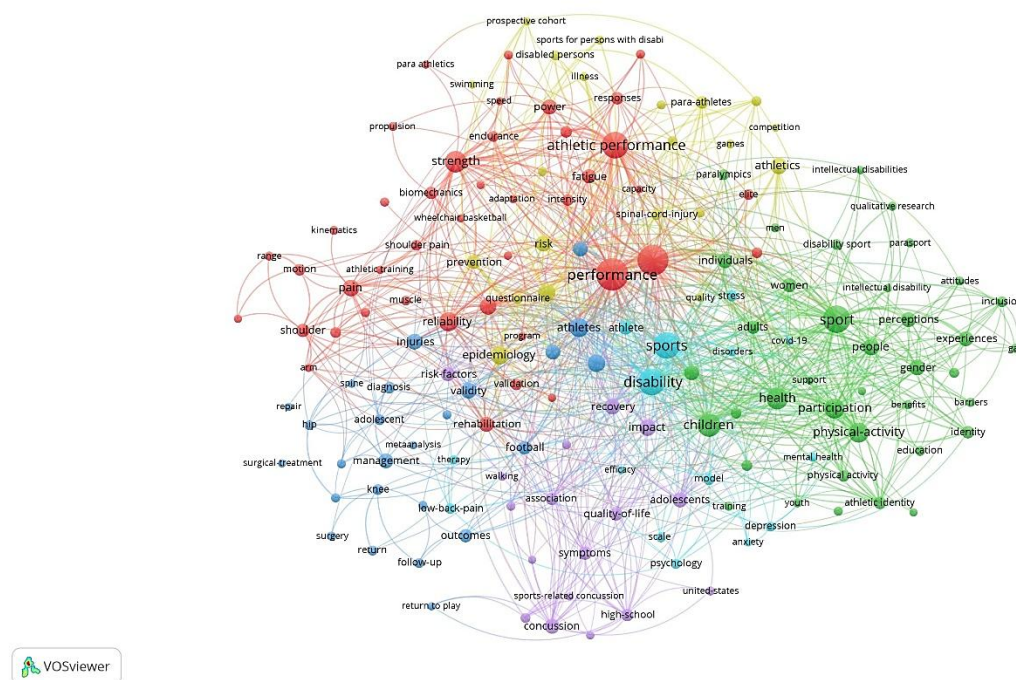


Figure 4. Keywords in publications on adapted sport over the last ten years, co-occurrence clustering.

Figure 5 shows the analysis of keywords in publications on adapted sports over the last ten years. The year of publication is shown as a gradient colour change.

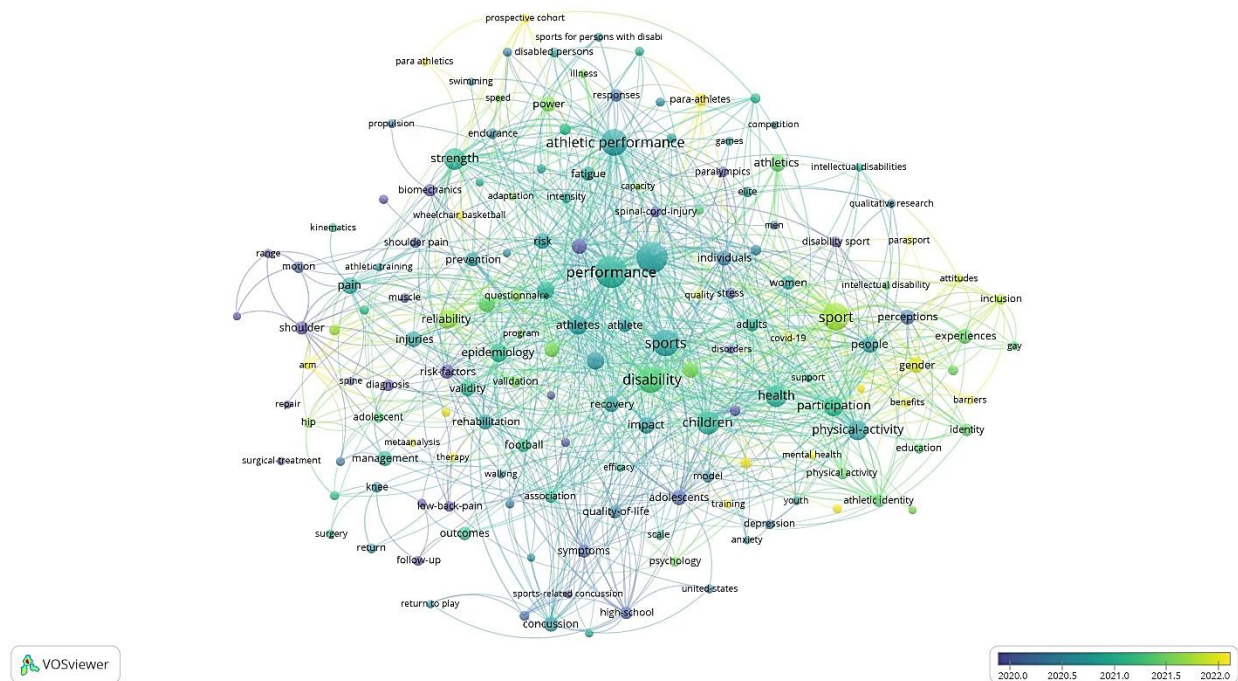


Figure 5. Keywords in publications on adapted sport, with colour variation by chronology.

DISCUSSION

The growth in publication output reflects the increasing scientific interest in disability sport and the professionalisation of para-sport research.

Bibliometrics provides a comprehensive and statistical view of the impact of research on specific topics. The inquiry in this article focused on a bibliometric analysis of studies related to para-athletics (IPC, 2024) over the last decade, identifying 1237 articles that were subjected to an analysis that provided comprehensive and objective information on the impact of para-athletics in scientific publications. This information is valuable for current and future research, as it allows for the identification of publication characteristics, research trends related to para-athletics, and critical points that can be addressed.

Observing research over the last decade allows us to reflect on the scientific contributions and development that para-athletics has undergone, both in year-on-year monitoring with a comprehensive view, and how these have increased alongside their respective exceptions. The annual results for para-athletics show a significant decline in 2023. The continuous growth in publications stagnated, which can be attributed to the socio-political or scientific context of the time, problems in scientific production, decisions by authorities or organizations that regulate para-athletics research, global economic investment in the field, or interruptions caused by sporting events held around that period. This phenomenon, reflected in this article, can open avenues for future research to investigate which factors affect scientific production in para-athletics and anticipate the research barriers that produce it (Marques & Alves, 2021).

The highest scientific output by country shows a trend toward more developed nations, which have produced a greater number of publications on para-athletics. High-income countries in North America and Western

Europe, along with Asia, lead this production of articles (Hurt et al., 2022). These countries have achieved outstanding research development in para-athletics over the last decade, consolidating their scientific leadership worldwide with a strong association to economic standards (Kudinska, Solovjova, & Korde, 2025). This analysis also suggests a strong correlation between the sporting results obtained in the Paris 2024 medal count and the last decade of scientific output in the countries that have most developed research in para-athletics. The concentration of research output in high-income countries may reflect disparities in research funding, infrastructure, and access to elite para sport programmes.

The trends by country can be attributed to the organization of sporting events where para-athletics competitions have been held (Baade & Matheson, 2016; Tien, Lo, & Lin, 2011; Kovpak, 2021; Pan, 2023), as is the case of Brazil with the Rio 2016 Games, Japan with Tokyo 2020, and France with Paris 2024. In fact, Brazil ranks fourth in publications, Japan thirteenth, and France twelfth; all three countries are also among the world's top three in publications with more than 10 articles in the last decade, reflecting this trend. This association is also found in the winter games, with a relationship with South Korea with 16 publications in the last decade with the games held in PyeongChang in 2018 and in Beijing China, held in 2022, finding China in the third place of countries with the most publication in para-athletics.

In addition, eighty-four countries have published articles related to para-athletics. The distribution of countries that managed to maintain an average of one article per year over the last decade shows a heterogeneous regional distribution, which could attribute a strong collaborative work in the northern hemisphere regions, unlike the publications in the southern hemisphere regarding para-athletics. In general, collaboration between institutions should continue to be strengthened with a view to the development of this topic.

Regarding journals, research related to it for athletics has been published in 556 journals, the main one being the Journal of Athletic Training, with 42 publications in the last decade regarding para-athletics, representing 3.40 % of the total, is the journal of Athletic Training. Of Athletic Training is by far the most influential magazine on the subject, almost doubling the number of publications of its successor, the American Journal. The Journal of Sports Medicine, with 22 publications, represents 1.78% of the total publications by journal, reflecting and confirming the research impact of the Journal. Of Athletic Training and the trend of para-athletics to publish scientific articles linked to training and the medical sciences that are linked to para-athletics over other impacts such as at the community, social, economic and humanities level (Alhumaid, Adnan, Said, Alobaid, & Khoo, 2024; Fagher, DeLuca, Derman, & Blauwet, 2022; Yamashita, 2021).

Regarding the keywords, co-occurrence was identified, reflecting the preferences of the studies, interests, trends, and research subjects. It is important to emphasize that the keywords show a frequency of non-specific information on research guidelines related to para-athletics. This study highlights the strength of keywords related to performance (Buhmann, Sayers, O'Brien, & Borg, 2024) of athletes in para-athletics, with 156 entries, followed by exercise (Fagher et al., 2022) with 135 and sport with 109. This reflects that trends in para-athletics are strongly related to athletic performance, rather than sociological and psychological aspects or athletic effects outside of performance, which accounted for a smaller proportion of publications related to the keywords.

Key words that appear isolated and infrequently in publications on the map were explored in depth to determine new guidelines or links to it for athletics. Such as the case of sports-related concussion (Lexell, Lovén, & Fagher, 2021). sports inclusion for people with disabilities (Rakaa, Bassiri, & Lotfi, 2025) and pain (Zwierzchowska, Gawel, & Rosolek, 2022).

Regarding the most prominent authors in the field of para-athletics, Winckler, C; is the only one to have published 10 articles in the last decade, averaging 1 article per year and representing 0.815% of the articles related to para-athletics. His research focuses on taxonomies of sports disciplines related to para-athletics and physical exertion tests (Winckler et al., 2024). Furthermore, the map shows a connection with other prominent authors in the area of sports, such as De Mello, MT and Silva A, who both have 9 publications each, including the characterization of para-athletes. (Pinheiro et al., 2024) and the prevalence of injuries respectively (Resende et al., 2025).

This study had some potential limitations. Notably, it focused solely on literature indexed in WoSCC because it provided the most authoritative and credible publication information. Inevitably, some useful information from other databases of similar relevance may have been omitted. Also, due to the large number of included articles, only the titles and abstracts of each publication were identified. Therefore, other sports were not classified, which could lead to potential inaccuracies in the included keywords compared to manual annotation. Finally, there could be some biases caused by linguistic restrictions in the search strategy. To achieve a more accurate range of evaluation, linguistic restrictions were imposed on the study. While only English-language publications were considered, some mixed-language publications might have a greater impact than English-only publications, potentially resulting in linguistic bias. All of the above factors could affect the accuracy of the bibliometric analysis.

CONCLUSION

This research evaluates the development of the scientific literature on para-athletics using a quantitative approach. The analysis provides a comprehensive overview of para-athletics research over the last decade, from 2015 to 2025. Annual publication output increased steadily from 2015 to 2022, before declining in 2023. The United States, England, and China stood out as the countries with the highest scientific output in the field of para-athletics. The most frequently published scientific journals are: Journal Of Athletic Training, American Journal “*The Journal of Sports Medicine and Human Sport Medicine*” were ranked as the three most productive journals in the last decade. The main categories were sports science, orthopedics, rehabilitation, and physiology. The main keywords associated with para-athletics are; performance, exercise, and sport.

This bibliometric study provides a comprehensive overview of global research trends in para-athletics from 2015 to 2025. The results highlight a sustained increase in scientific production, with research concentrated in economically developed countries and primarily focused on sport performance and exercise science. Future research should expand towards social inclusion, policy development, and long-term health outcomes in para-athletes to support a more holistic understanding of disability sport.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. Eduardo Reinoso González (Main Author); participated in the conceptualization (direction), research, methodological design, data curation and manuscript writing. Mauricio Godoy Rencoret (Co-Author); participated in the final editing and revision of the manuscript. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

- Alhumaid, M., Adnan, Y., Said, M., Alobaid, M., & Khoo, S. (2024). Empowerment and social inclusion through Para sports: a qualitative study on women with physical impairments in Saudi Arabia. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1366694>
- Ben Rakaa, O., Bassiri, M., & Lotfi, S. (2025). Impact of adapted physical education and para-athletics on the mental skills and inclusion of adolescents with disabilities. *Retos*, 68, 82-94. <https://doi.org/10.47197/retos.v68.111520>
- Baade, R., & Matheson, V. (2016). Going for the Gold: The Economics of the Olympics. *Journal of Economic Perspectives*, 30, 201-218. <https://doi.org/10.1257/jep.30.2.201>
- Buhmann, R., Sayers, M., O'Brien, J., & Borg, D. (2024). Important features of bench press performance in non-disabled and Para athletes: A scoping Review. *PLoS one*, 19 (11), e0310127. <https://doi.org/10.1371/journal.pone.0310127>
- Fagher, K., DeLuca, S., Derman, W., & Blauwet, C. (2022). Optimizing health equity through for sport. *British Journal of Sports Medicine*, 57, 131 - 132. <https://doi.org/10.1136/bjsports-2022-106229>
- Heffernan, A. (2022). Paralympic assemblages globalizing international relations: an autoethnographic account of global politics at the Paralympic Games. *Sport in Society*, 25(1), 56-69. <https://doi.org/10.1080/17430437.2020.1772234>
- Hurt, J., Yang, L., Sorger, J., Lampoltshammer, T., Pulda, N., Rosenbichler, U., Thurner, S., & Klimek, P. (2022). Evidence-based policy-making in sports funding using a data-driven optimization approach. *PLOS ONE*, 17. <https://doi.org/10.1371/journal.pone.0312179>
- International Paralympic Committee. (2024). Paris 2024 Paralympic Games qualification regulations. Retrieved from [Accessed 2026, 24 April]: <https://www.paralympic.org/>
- Kovpak, V. (2021). Influence of the organization of the Olympic Games as a mega-event on urbanization processes in the city and the region. *Urban Development and Spatial Planning*, 77, 241-252. <https://doi.org/10.32347/2076-815x.2021.77.241-252>
- Kudinska, M., Solovjova, I., & Korde, Ž. (2025). Research trends in the field of sport impact on the economy: a bibliometric analysis. *Frontiers in Sports and Active Living*, 7. <https://doi.org/10.3389/fspor.2025.1545264>

- Lexell, J., Lovén, G., & Fagher, K. (2021). Incidence of sports-related concussion in elite athletes - a 52-week prospective study. *Brain Injury*, 35, 971 - 977. <https://doi.org/10.1080/02699052.2021.1942551>
- Marques, M., & Alves, A. (2021). Investigating environmental factors and paralympic sports: an analytical study. *Disability and Rehabilitation: Assistive Technology*, 16, 414 - 419. <https://doi.org/10.1080/17483107.2020.1780483>
- Pan, Q. (2023). Leveraging Benefits from Winter Olympic Games, Beijing 2022 through Leveraging Strategies. *Open Journal of Social Sciences*, 11(1), 125-131. <https://doi.org/10.4236/jss.2023.111011>
- Pinheiro, LSP, Silva, A., Madaleno, FO, Verhagen, E., de Mello, MT, Ocarino, JM, & Resende, RA (2024). Prevalence and incidence of health problems and their characteristics in Brazilian para athletes: A one-season single- center prospective pilot study. *Disability and Health Journal*, 17(1), 101511. <https://doi.org/10.1016/j.dhjo.2023.101511>
- Resende, R., Madaleno, F., Verhagen, E., Wezenbeek, E., de Mello, MT, Chagas, MH, Gonçalves, D., Silva, A., Pinheiro, L., Ocarino, J., & Witvrouw, E. (2025). Comprehensive 1-year multilevel study of sports injuries in for athletes: impact of season timing, years of sports experience, impairment and sports type. *BMJ Open Sport & Exercise Medicine*, 11(2), e002474. <https://doi.org/10.1136/bmjsem-2025-002474>
- Sakalidis, K.E., Fadeeva, A., Hettinga, F.J., & Ling, F.C.M. (2023). The role of the social environment in including sports participation - Identifying similarities and challenges in athletes with and without Intellectual Disabilities through coaches' eyes: A qualitative inquiry. *PLOS ONE*, 18(1), e0280379. <https://doi.org/10.1371/journal.pone.0280379>
- Tien, C., Lo, H., & Lin, H. (2011). The Economic Benefits of Mega Events: A Myth or a Reality? A Longitudinal Study on the Olympic Games. *Journal of Sport Management*, 25, 11-23. <https://doi.org/10.1123/jsm.25.1.11>
- Van Biesen, D., Burns, J., Mactavish, J., Van de Vliet, P., & Vanlandewijck, Y. (2021). Conceptual model of sport- specific classification for para-athletes with intellectual impairment. *Journal of Sports Sciences*, 39(sup1), 19-29. <https://doi.org/10.1080/02640414.2021.1881280>
- van der Linden, M.L., Corrigan, O., Tennant, N., & Verheul, M.H.G. (2021). Cluster analysis of impairment measures to report an evidence-based classification structure in RaceRunning, a new World Para-athletics event for athletes with hypertonia, ataxia or Athetosis. *Journal of Sports Sciences*, 39(sup1), 159-166. <https://doi.org/10.1080/02640414.2020.1860360>
- Winckler, C., Costa, SG da, Campos, LFCC de, & Lourenço, TF (2025). Differences and relationships between velocity attained at VO 2max and time trial performances in Para athletes with vision impairment and their guides. *British Journal of Visual Impairment*, 43(2), 565-572. <https://doi.org/10.1177/02646196241252611>
- Yamashita, R. (2021). Mega-Para-Sporting Event Social Impacts Perceived by Tokyo Residents: Comparison of Residents' Vitality. *Sustainability*, 13(16), 9311. <https://doi.org/10.3390/su13169311>
- Zwierzchowska, A., Gawel, E., & Rosolek, B. (2022). Determinants of the prevalence and location of musculoskeletal pain in elite for athletes. *Medicine*, 101(42), e31268. <https://doi.org/10.1097/MD.00000000000031268>



External oblique activation is augmented with a water-filled resistance training implement compared to a static load

- Dylan Cho** . Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Joshua A. Kidwell**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Trent Yamamoto**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Sean Taylor**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Joshua Thomas**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Nathan Truneh**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Jacob J. Bright**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Joseph Hoang**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Vishruth Shatagopam**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.
- Phillip Goldman**. Baylor College of Medicine. Houston, TX, United States of America.
- Eric V. Neufeld**. Northwell Orthopedics. New Hyde Park, NY, United States of America.
- Brett A. Dolezal**. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.

ABSTRACT

This study investigated the effects of a water-filled resistance training device (WTD) compared to a static training load on core muscle activation in the rectus abdominis (RA) and external obliques (EO) during a rotational core exercise. Twelve college-aged male participants with moderate resistance training experience performed 4 sets of 12 repetitions on each side of the Russian twist exercise using both the WTD and a static-weight control in a randomized crossover design on two separate days. Surface electromyography (sEMG) was used to measure muscle activation in both conditions. Results indicated that the WTD condition elicited greater mean and peak muscle activation in both muscle groups compared to the control. EO activation was 85% higher in mean ($p = .002$, $d = 0.95$) and 77% greater in peak ($p = .002$, $d = 1.59$) values, while no significant difference in RA activation was identified between groups. These findings suggest that the WTD enhances activation of the external obliques relative to a static load, supporting its potential use in instability-based training for moderately trained males.

Keywords: Performance analysis, Core muscle, Instability, sEMG.

Cite this article as:

Cho, D., Kidwell, J. A., Yamamoto, T., Taylor, S., Thomas, J., Truneh, N., Bright, J. J., Hoang, J., Shatagopam, V., Goldman, P., Neufeld, E. V., & Dolezal, B. A. (2026). External oblique activation is augmented with a water-filled resistance training implement compared to a static load. *Scientific Journal of Sport and Performance*, 5(3), 542-551. <https://doi.org/10.55860/CEMI8585>

Corresponding author. Airway and UC Fit Digital Health-Exercise Physiology Laboratory. David Geffen School of Medicine. University of California Los Angeles. Los Angeles, CA, United States of America.

E-mail: dylan.j.cho@outlook.com

Submitted for publication March 27, 2026.

Accepted for publication April 30, 2026.

Published May 07, 2026.

Scientific Journal of Sport and Performance. ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/CEMI8585>

INTRODUCTION

Core strength plays a pivotal role in human movement, providing stability in the context of both daily activities and athletic performance (Escamilla et al., 2010). Strengthening the core musculature, particularly the rectus abdominis (RA) and external obliques (EO), has been shown to improve spine stability while simultaneously reducing the risk of injury and other musculoskeletal disorders (Kim and Oh, 2015; Oliva-Lozano and Muyor, 2020). Traditionally, core training has relied on static resistance exercises involving fixed weights, such as dumbbells, barbells, and various resistance machines. However, recent research in regards to instability resistance training, specifically within the core musculature, suggests that introducing unstable conditions may elicit greater neuromuscular engagement and subsequent training adaptations (Batistia et al., 2024).

Instability resistance training primarily involves unstable surfaces or shifting loads designed to challenge an individual's ability to stabilize themselves while experiencing dynamic movement. This style of training has been shown to increase core activation relative to stable conditions, likely due to increased proprioceptive demands and muscle recruitment (Kim and Oh, 2015; Batista et al., 2024; Saeterbakken et al., 2022). Commonly utilized instability tools, such as yoga balls, Swiss balls, and foam rollers, require greater muscle coordination in order to control their movement. More recently, water-based instability implements have become more popular as an alternative instability training modality. Unlike traditional methods, water-based implements create shifting loads through the unpredictable movement of water, potentially leading to increased core activation by necessitating compensatory adjustments for stabilization (Glass et al., 2016).

Despite the growing interest in water-based resistance tools, research largely remains limited in regards to their effectiveness relative to static weights. Many prior studies have focused on unstable surfaces rather than unstable loads, such as the one produced by the shifting of the water (Saein et al., 2024; Behm et al., 2010). Furthermore, research has shown conflicting results, as although some studies have shown increased core muscle activation in exercises that have greater stability requirements (Oliva-Lozano and Muyor, 2020; Batista et al., 2024; Bao et al., 2025), others have not (Saeterbakken et al., 2022). In regard to the core musculature, some evidence suggests that unstable loads can enhance neuromuscular engagement (Batista et al., 2024). However, it remains to be determined if water-filled resistance training provides a unique stimulus beyond that of traditional weights. Previous studies using sEMG have shown increased core muscle activation when core exercises were performed with other instability modalities, such as Swiss and BOSU balls (Escamilla et al., 2010; Czaprowski et al., 2014). Furthermore, although prior studies have included a water-based implement, they have primarily focused on the activation of the core through secondary core exercises, such as bicep curls, squats, and the clean and jerk (Ditroilo et al., 2018; Glass et al., 2016; Calatayud et al., 2015). As such, no studies have directly compared core muscle activation between water-filled training tubes and a static weight using surface electromyography (sEMG) during core-focused exercises.

The primary aim of this study is to assess the impact of a water-filled training tube on core muscle activation in the rectus abdominis and external obliques relative to a static weight during the execution of a core-focused exercise. We hypothesize that the water-filled training tube will elicit greater core muscle activation, measured via sEMG, compared to that of the static weight.

METHODS

Participants

Twelve apparently-healthy male participants ($n = 12$) from Southern California volunteered for this study. Inclusion criteria included individuals aged 18-35 years old with moderate experience in resistance training,

defined as engaging in structured, resistance-based training (e.g., weightlifting, bodyweight training, or similar exercises) at least 2 to 3 times per week for a minimum of 3 to 6 months. This was done to ensure participants possess the strength and familiarity necessary to effectively lift and stabilize the water-filled resistance training device (WTD). Exclusion criteria included the presence of recent upper-body or core-related injuries that may inhibit performance, any significant musculoskeletal, cardiovascular, or pulmonary disorders that may limit exercise ability, and the use of any drugs and/or supplements that enhance anabolic performance. A sample size of $n = 10$ was calculated based on α priori power analysis using sEMG data from an unpublished exploratory study of a similar study $\alpha = 0.05$ and $B = 0.20$, using five resistance-trained, collegiate-aged males. All exploratory participants provided written informed consent while ethical approval was obtained from UCLA (IRB: 11-003190). Off-site participants provided written informed consent and single IRB approval (sIRB: BRANY, NY, USA). Research practices were conducted in full accordance with the ethical standards established by the Declaration of Helsinki and the International Journal of Exercise Science (Navalta et al., 2019).

Study design

This study was a one-week, single-blind, randomized crossover research design. Each participant completed two sessions, no less than 48 hours apart: one in the intervention condition, using the WTD, and the other in the control condition, using a static weight of equivalent mass. The order of conditions was randomized using an online random number generator. During each session, participants performed Russian Twists, a seated rotational core exercise in which the torso is twisted from side to side while holding a weight. The exercise was performed for 4 sets of 12 repetitions per side under each condition.

Participants were instructed to follow the tempo of an online metronome with a cadence of 40 beats per minute, switching sides during the exercise every 1.5 seconds. Prior to their participation in the study, participants were asked to avoid any strenuous or additional resistance training for the duration of their participation in the study. A two-minute break was provided between sets. All sessions were monitored by researchers under the guidance of the lab director at the UC Fit Digital Health - Exercise Physiology Research Laboratory at UCLA.

Anthropometric measures

Body mass and height

Body mass and Height: Body mass was measured on a calibrated medical scale (accuracy ± 0.1 kg), and height was determined using a precision stadiometer (Seca, Hanover, MD, United States; accuracy ± 0.01 m). In a fasted state and after voiding their bladder, participants were instructed to remove unnecessary clothing and accessories prior to being weighed, as well as remove their shoes prior to taking height measurements.

Body composition

Body fat percentage was measured using a validated octipolar, multi-frequency, multi-segmental bioelectrical impedance analyser (BIA) (InBody Co., Seoul, Korea Republic) (Dolezal et al., 2013). To ensure accuracy, participants adhered to standard pre-measurement BIA guidelines recommended by the American Society of Exercise Physiologists (Heyward, 2001). Briefly, the test was performed after at least 3 hours of fasting and voiding, with participants instructed to remain hydrated and not exercise 2 hours before testing. After investigators explained the procedure, the participant stood upright with their feet on two metallic footpads while holding a handgrip with both hands. The instrument measured resistance and reactance using proprietary algorithms.

Protocol

Participants reported to the UCFIT Digital Health - Exercise Physiology Research Laboratory on two non-consecutive days (within no less than 48 hours apart) at the same time of day under controlled temperature conditions. Body mass was measured on a calibrated medical scale (accuracy ± 0.1 kg), and height was determined using a precision stadiometer (Seca, Hanover, MD, United States; accuracy ± 0.01 m). In a fasted state and after voiding their bladder, participants were instructed to remove unnecessary clothing and accessories prior to being weighed, as well as remove their shoes prior to taking height measurements.

Following body mass measurement, body fat percentage was assessed using a validated octipolar, multi-frequency, multi-segmental bioelectrical impedance analyser (BIA) (InBody Co., Seoul, Korea Republic) (Dolezal et al., 2013). To ensure accuracy, participants adhered to standard pre-measurement BIA guidelines recommended by the American Society of Exercise Physiologists (Heyward, 2001). Briefly, the test was performed after at least 3 hours of fasting and voiding, with participants instructed to remain hydrated and not exercise 2 hours before testing. After investigators explained the procedure, the participant stood upright with their feet on two metallic footpads while holding a handgrip with both hands. The instrument measured resistance and reactance using proprietary algorithms.

Participants were then fitted with two MR EMG sEMG sensor recording devices (MR EMG, Dunedin, New Zealand). In accordance with previously established sEMG practices, one MR EMG sensor was placed 4 cm to the right of the participant's navel to record the sEMG activity of their RA.16 The second MR EMG sensor was placed halfway between the anterior superior iliac spine (ASIS) and the most inferior point of the costal margin (Huebner et al., 2015). The sEMG sensors were placed on the right side muscles only, as prior research has demonstrated sEMG reading symmetry between the left and right core muscles (Escamilla et al., 2010; Escamilla et al., 2006; Ng et al., 1998). Prior to placement, the area was shaved and rubbed with an alcohol pad by the researcher to ensure proper connectivity for each sensor.

Before beginning the exercise, participants received standardized verbal instructions and a live demonstration of the Russian Twist exercise by a trained research assistant. To ensure proper form, participants were instructed to begin in a seated position with their legs extended in front of them and their heels lightly contacting the floor. Participants maintained a straight back while leaning slightly backward at an approximately 45° angle from vertical.

The WTD was held with both hands using an underhand grip on its horizontal handles, positioned at chest level. During the exercise, participants were instructed to rotate their torsos alternatively to the right and left while keeping their lower bodies stable and elbows tucked close to their torso. Rotation was paced using an online metronome set at 1.5-second intervals. At each beep of the metronome, participants rotated the WTD to the opposite side, completing one full repetition after returning the WTD to its original position.

Four sets, each consisting of 12 repetitions per side, were completed with a 90-second rest period in between each set. A trained research assistant was present throughout all sessions to monitor safety, provide real-time feedback strictly related to form, and ensure adherence to the standardized protocol.

Resistance tube specifications

The experimental condition in this experiment used a WTD known as the Water Pill (Finesse Performance, Los Angeles, CA). The Water Pill is a cylindrical, transparent plastic tube with handles attached to the sides. By filling it with water and air, the Water Pill creates an unstable internal load through the displacement of

water. For this study, the WTD was filled with water to 50% of its total capacity, corresponding to approximately 8.5 kilograms to balance instability while also ensuring adequate weight stimulus.

The control condition utilized a custom-made resistance tube. To ensure that the control condition adequately approximated the experimental condition, a single Water Pill unit was disassembled by removing the end caps. A mailing tube was inserted centrally into the Styrofoam discs. The internal cavity was then filled with several stacked Styrofoam discs to replicate the typical volume of the device, as well as to hold the mailing tube in place. Sand and metal weights were added to the mailing tube until the total weight of the control device matched that of the WTD (8.5 kilograms). Once the desired weight was achieved, the ends were resealed to ensure structural integrity and to prevent the internal movement of the weight. These steps were taken to ensure that both conditions were identical in size, shape, and handling, differing only in the presence (experimental) or absence (control) of fluid movement.



Figure 1. Water Pill (above) and Static Control (below) used in the experiment.

Statistical analysis

Descriptive statistics are presented as median and interquartile range (IQR). Data deviated significantly from normality per Shapiro-Wilk tests. Wilcoxon signed rank-tests were utilized to compare control and intervention data within each group. All tests were two-tailed. Statistical significance was determined if $p < .05$ after employing a Holm-Bonferroni correction to control the familywise error rate. Effect size was determined as Cohen's d , along with 95% confidence intervals (CI). All analyses were performed using R version 4.5.0 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

All twelve participants successfully completed both training sessions, including one session with the water-filled resistance training device (WTD) and another session with the control device (CON), with ≥ 48 -hours in between sessions. No significant differences were observed in mean ($p = .182$) or maximum rectus abdominis activation ($p = .328$) between sessions. However, the WTD condition demonstrated a significantly greater mean ($p = .002$, $d = 0.95$) and maximum ($p = .002$, $d = 1.59$) activation of the external oblique compared to CON (Table 1).

Both increases were associated with large effect sizes as determined by Cohen's d . Mean external oblique activation was approximately 85% greater in WTD relative to control ($71.8 \pm 36.0 \mu\text{V}$ versus $38.8 \pm 21.2 \mu\text{V}$, respectively). Maximum external oblique activation was approximately 77% higher in WTD compared to control ($258.2 \pm 102.5 \mu\text{V}$ versus $145.8 \pm 77.5 \mu\text{V}$, respectively).

Table 1. External oblique and rectus abdominis mean and max activation values.

Group	External oblique activation				Rectus abdominis activation (μV)			
	Mean (μV)	p-value	Max (μV)	p-value	Mean (μV)	p-value	Max (μV)	p-value
CON	38.8 (21.2)	.002	145.8 (77.5)	.002	11.5 (24.5)	.182	40.0 (70.3)	.328
WTD	71.8 (36.0)		258.2 (102.5)		13.7 (20.4)		44.5 (82.9)	

DISCUSSION

The results of the study demonstrated that using a water-filled resistance training device (WTD) during a rotational core exercise significantly increased muscle activation in the external obliques (EO) compared to a static weight control but did not show any significant increases in activation of the rectus abdominis (RA). The mean and maximum EO activation was 85% ($p = .002$, $d = 0.95$) and 77% ($p = .002$, $d = 1.59$) greater, respectively, in WTD compared to CON. However, differences in mean and maximum rectus abdominis (RA) activation between WTD and CON were nonsignificant. These findings partially support our hypothesis, which proposed that the WTD would significantly enhance core muscle activation in the rectus abdominis and external oblique relative to a static weight device. The significant increase in activation for the EO only and not the RA suggests that while the WTD's unstable load demands led to heightened core muscular engagement relative to CON's stable load, its efficacy may be restricted to certain muscle groups.

Although the Russian twist movement investigated in this study is designed to directly target the core, our findings align with broader research demonstrating that instability—particularly when introduced through water-filled implements—enhances core muscle activation even during exercises not primarily focused on the core. For instance, one paper reported a significant increase in rectus abdominis activation when participants performed bicep curls using a dynamically flowing water-filled tube (Glass et al., 2016). Despite not being the directly targeted muscle of the exercise, this finding suggests that the shifting water necessitates at least a partial compensatory response from trunk musculature in order to maintain balance. Similarly, another paper found that bench pressing with a water-filled tube elicited a near five-fold enhancement in core muscle activation (as measured by %MVC) compared to using a traditional barbell (Nairn et al., 2015). These findings, along with our own, support the premise that water-based instability elicits a compensatory response from both deep (e.g. transversus abdominis) and superficial (e.g. external oblique) core stabilizers, regardless of whether the movement is inherently core-focused.

However, the efficacy of instability training in enhancing core activation has long been debated. For instance, Anderson & Behm (2004) found similarly to our results that while performing upper-body exercises on unstable surfaces (e.g. Swiss balls), rectus abdominis muscle activation was not significantly different relative to a firm surface (e.g. flat bench). Additionally, Cressey et al. (2007) noted that instability can reduce the amount of external load that can be used during resistance exercises, potentially limiting hypertrophic gains in the primary target muscles. In these cases, the increased neuromuscular demand placed on the core—likely driven by anticipatory postural adjustments—may come at the expense of training other muscles involved in completing the exercises. As such, the mechanistic demands of using the WTD during the Russian Twist exercise, which emphasize rotational and lateral stabilization rather than trunk flexion, provide a plausible explanation for the non-significant difference observed in the rectus abdominis, a primarily flexion-oriented muscle.

In core-focused rotational movement exercises, both the rectus abdominis and external obliques are recruited to support trunk stability and proper spinal alignment during dynamic movements. Among other activated muscles such as the transverse abdominis, the RA maintains trunk flexion, while the EO contributes to lateral

flexion and rotation (Oliva-Lozano and Muyor, 2020). The activation of these stabilizing core muscles can be categorized into the central nervous system's anticipatory (APA) and compensatory (CPA) postural adjustments. APAs refer to the activation or inhibition of trunk musculature prior to anticipated balance perturbations. In contrast, CPAs are reactive responses that activate or inhibit trunk musculature initiated by post-perturbation sensory feedback (Santos et al., 2009).

Traditional static resistance exercises demonstrate greater APAs, as the predictability of these exercises allows the body to prepare. However, the unpredictable, dynamic fluid movement of instability devices, such as the WTD, results in greater CPAs (Santos et al., 2009). The device's unpredictability elicits greater proprioceptive and neuromuscular demands, recruiting both deep and superficial core stabilizers to rapidly restore trunk stability during rotation. The reactive mechanism elicited by instability devices underscores our study's demonstration of heightened core muscular engagement when utilizing water-based instability training.

Water-based instability training may be leveraged for rehabilitation, specifically in programs designed to improve postural control and core muscle activation. The present study's use of the simple Russian Twist exercise allows it to be incorporated into nearly any array of existing programs. The instability engages more motor units than static weights, offering significant use for patients unable to walk or readily activate their core without risk of injury (Glass et al., 2016). Additionally, the unpredictability introduced by the device better mirrors the functional demands of daily life, which are key to practice for recovering patients, compared to isolated static movements (Rozevink et al., 2021). The increased substitution of weight for instability offers promising results for more holistic muscle activation and key motor activation in clinical and rehabilitation settings (Rozevink et al., 2021).

Additionally, water-based instability training presents a significant opportunity for athletes to improve existing techniques for both routine and rehabilitation training. Our results, along with those of similar studies have shown increased activation of the rectus abdominis and other core muscles when using water-based instability modalities (Glass et al., 2016; Ditroilo et al., 2018). The relevance of heightened activation of core musculature to athletes has previously been demonstrated and has been shown to improve general athletic performance due to improved core endurance and balance (Yu et al., 2025). Further research also suggests that core training can lead to improved performance in select sports, including football, basketball, and golf (Luo et al., 2022), with general increases in physical strength and smaller increases in speed, agility, and power (Dong et al., 2023). Given the impact of core muscle training on athletic performance, further experimentation could evaluate the relevance of water-based instability training for enhancing control and stability-maintenance abilities in athletes of different respective sports.

Limitations

The present study is limited by the use of a small, homogenous sample size. Data was collected exclusively from male, college-aged, apparently healthy individuals. As such, interpretations of results should be limited to this demographic, as it is unclear if the effects shown in the study are generalizable to all populations. Given the short duration of the study, additional longitudinal research is necessary to understand the long-term impacts of using the water-filled resistance training device (WTD) on core muscle activation during training. Further limitations include the use of a single weight (18.8 pounds), making it unclear if the observed effects persist at different intensities. Lastly, while sEMG is validated for surface-level muscle activity, it may be affected by outside factors, including body fat and body water content.

CONCLUSION

The present study investigated the effects of a water-filled resistance training device (WTD) on core muscle activation during a rotational core exercise, specifically targeting the rectus abdominis and external obliques. Results demonstrated that both mean and peak muscle activation were greater when participants used the WTD compared to a static-weight control, suggesting that instability-based resistance can elicit higher levels of neuromuscular engagement. These findings support the potential use of water-filled resistance devices in enhancing the efficacy of core training. For individuals aiming to improve trunk stability and muscular activation, incorporating dynamic, fluid-based resistance may provide a valuable supplement to traditional resistance training protocols.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. The study was conceived and designed by D. C. and B. A. D., whereas S. T., J. T., N. T., J. J. B., and J. H. performed data collection. P. G., T. Y., E. V. N., J. A. K., and B. A. D. completed data analysis. P. G., T. Y., J. A. K., D. C., and B. A. D. interpreted data and composed the manuscript, while S. T., J. T., N. T., J. J. B., J. H., and V. S. made crucial edits. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

- Anderson, K. G., & Behm, D. G. (2004). Maintenance of EMG Activity and Loss of Force Output With Instability. *J Strength Cond Res*, 18(3), 637-640. <https://doi.org/10.1519/00124278-200408000-00043>
- Bao, Z., Wang, S., & Li, Z. (2025). Effects of unstable training on muscle activation: a systematic review and meta-analysis of electromyographic studies. *PeerJ*, 13, e19751. <https://doi.org/10.7717/peerj.19751>
- Batista G. A., Beltrán S. P., Passos, M. H. P. d., Calixtre L. B., Santos, L. R. d. H., & de Araújo, R. C. (2024). Comparison of the Electromyography Activity during Exercises with Stable and Unstable Surfaces:

- A Systematic Review and Meta-Analysis. *Sports* (Basel), 12(4), 111. <https://doi.org/10.3390/sports12040111>
- Behm, D. G., Drinkwater, E. J., Willardson, J. M., & Cowley, P. M. (2010). The use of instability to train the core musculature. *Appl Physiol Nutr Metab*, 35(1), 91-108. <https://doi.org/10.1139/H09-127>
- Calatayud, J., Borreani, S., Colado, J. C., Martin, F., Tella, V., & Andersen, L. L. (2015). Bench press and push-up at comparable levels of muscle activity results in similar strength gains. *J Strength Cond Res*, 29(1), 246-253. <https://doi.org/10.1519/JSC.0000000000000589>
- Cressey, E. M., West, C. A., Tiberio, D. P., Kraemer, W. J., & Maresch, C. M. (2007). The Effects of Ten Weeks of Lower-Body Unstable Surface Training on Markers of Athletic Performance. *J Strength Cond Res*, 21(2), 561-567. <https://doi.org/10.1519/00124278-200705000-00047>
- Czaprowski, D., Afeltowicz, A., Gębicka, A., Pawłowska, P., Kędra, A., Barrios, C., & Hadała, M. (2014). Abdominal muscle EMG-activity during bridge exercises on stable and unstable surfaces. *Phys Ther Sport*, 15(3), 162-168. <https://doi.org/10.1016/j.ptsp.2013.09.003>
- Ditroilo, M., O'Sullivan, R., Harnan, B., Crossey, A., Gillmor, B., Dardis, W., & Grainger, A. (2018). Water-filled training tubes increase core muscle activation and somatosensory control of balance during squat. *J Sports Sci*, 36(17), 2002-2008. <https://doi.org/10.1080/02640414.2018.1431868>
- Dolezal, B. A., Lau, M. J., Abrazado, M., Storer, T. W., & Cooper, C. G. (2013) Validity of two commercial grade bioelectrical impedance analyzers for measurement of body fat percentage. *J Exerc Physiol Online*, 16(4), 74-83.
- Dong, K., Yu, T., & Chun, B. (2023). Effects of Core Training on Sport-Specific Performance of Athletes: A Meta-Analysis of Randomized Controlled Trials. *Behav Sci* (Basel), 13(2), 148. <https://doi.org/10.3390/bs13020148>
- Escamilla, R. F., Babb, E., DeWitt, R., Jew, P., Kelleher, P., Burnham, T., Busch, J., D'Anna, K., Mowbray, R., & Imamura, R. T. (2006). Electromyographic analysis of traditional and nontraditional abdominal exercises: implications for rehabilitation and training. *Phys Ther*, 86(5), 656-671. <https://doi.org/10.1093/ptj/86.5.656>
- Escamilla, R. F., Lewis, C., Bell, D., Bramblett, G., Daffron, J., Lambert, S., Pecson, A., Imamura, R., Paulos, L., & Andrews, J. R. (2010). Core Muscle Activation During Swiss Ball and Traditional Abdominal Exercises. *J Orthop Sports Phys Ther*, 40(5), 265-276. <https://doi.org/10.2519/jospt.2010.3073>
- Glass, S. C., Blanchette, T. W., Karwan, L. A., Pearson, S. S., O'Neil, A. P., & Karlik, D. A. (2016). Core Muscle Activation During Unstable Bicep Curl Using a Water-Filled Instability Training Tube. *J Strength Cond Res*, 30(11), 3212-3219. <https://doi.org/10.1519/JSC.0000000000001418>
- Heyward, V. (2001). ASEP methods recommendation: body composition assessment. *J Exerc Physiol Online*, 4(4), 1-12.
- Huebner, A., Faenger, B., Schenk, P., Scholle, H., & Anders, C. (2015) Alteration of Surface EMG amplitude levels of five major trunk muscles by defined electrode location displacement. *J Electromyogr Kinesiol*, 25(2), 214-223. <https://doi.org/10.1016/j.jelekin.2014.11.008>
- Kim, M., & Oh, J. (2015). Effects of performing an abdominal hollowing exercise on trunk muscle activity during curl-up exercise on an unstable surface. *J Phys Ther Sci*, 27(2), 501-503. <https://doi.org/10.1589/jpts.27.501>
- Luo, S., Soh, K. G., Soh, K. L., Sun, H., Nasiruddin, N. J. M., Du, C., & Zhai, X. (2022). Effect of Core Training on Skill Performance Among Athletes: A Systematic Review. *Front Physiol*, 13, Article 915259. <https://doi.org/10.3389/fphys.2022.915259>
- Nairn, B. C., Sutherland, C. A., & Drake, J. (2015). Location of Instability During a Bench Press Alters Movement Patterns and Electromyographical Activity. *J Strength Cond Res*, 29(11), 3162-3170. <https://doi.org/10.1519/JSC.0000000000000973>

- Navalta, J. W., Stone, W. J., & Lyons, T. S. (2019). Ethical issues relating to scientific discovery in exercise science. *Int J Exerc Sci*, 12(1), 1-8. <https://doi.org/10.70252/EYCD6235>
- Ng, J. K., Kippers, V., & Richardson, C. A. (1998). Muscle fibre orientation of abdominal muscles and suggested surface EMG electrode positions. *Electromyogr Clin Neurophysiol*, 38(1), 51-58.
- Oliva-Lozano J. M., & Muyor J. M. (2020). Core Muscle Activity during Physical Fitness Exercises: A Systematic Review. *Int J Environ Res Public Health*, 17(12), 4306. <https://doi.org/10.3390/ijerph17124306>
- Rozevink, S. G., Hijmans, J. M., Horstink, K.A., & van der Sluis, C. K. (2021). Effectiveness of task-specific training using assistive devices and task-specific usual care on upper limb performance after stroke: a systematic review and meta-analysis. *Disabil Rehabil Assist Technol*, 18(7), 1245-1258. <https://doi.org/10.1080/17483107.2021.2001061>
- Saein, A. M., Kahrizi, S., & Boozari, S. (2024) Effects of unstable load and unstable surface on trunk muscles activation and postural control in healthy subjects. *J Biomech*, 173, Article 112257. <https://doi.org/10.1016/j.jbiomech.2024.112257>
- Saeterbakken, A. H., Stien, N., Pedersen, H., & Andersen, V. (2022). Core Muscle Activation in Three Lower Extremity Exercises With Different Stability Requirements. *J Strength Cond Res*, 36(2), 304-309. <https://doi.org/10.1519/JSC.0000000000003465>
- Santos, M. J., Kanekar, N., & Aruin, A. S. (2009). The role of anticipatory postural adjustments in compensatory control of posture: 1. Electromyographic analysis. *J Electromyogr Kinesiol*, 20(3), 388-397. <https://doi.org/10.1016/j.jelekin.2009.06.006>
- Yu, T., Xu, Y., Zhang, Z., Sun, Y., Zhong, J., & Ding, C. (2025). The impact of core training on overall athletic performance in different sports: a comprehensive meta-analysis. *BMC Sports Sci Med Rehabil*, 17(1), 112. <https://doi.org/10.1186/s13102-025-01159-6>



Changes in cross-punch speed and torso motion associated with round progression in amateur boxers

 **Shinya Kimoto.** Graduate School of Physical Education. National Institute of Fitness and Sports in Kanoya. Kanoya-city, Japan.

 **Kanzi Ohyama.** Graduate School of Physical Education. National Institute of Fitness and Sports in Kanoya. Kanoya-city, Japan.

 **Akira Maeda**  . National Institute of Fitness and Sports in Kanoya. Kanoya-city, Japan.

ABSTRACT

The aim of this study was to examine changes in cross-punch speed and trunk motion associated with the progression of rounds in amateur boxing athletes. We included 11 male amateur boxers (height: 167.3 ± 3.8 cm, weight: 65.3 ± 10.1 kg). Cross-punch performance was assessed before and after a 3-minute maximal-effort punching test, based on a previous study by Dunn et al. (2019). Performance was evaluated using an optical three-dimensional motion-capture system, with measurements conducted four times: once under a baseline condition (Pre) and three times under post-test conditions (Post1–Post3). No significant changes in punch speed were observed from Pre to Post1–Post3. The negative values of lower and upper torso axial-rotation angular velocities (rotation opposite to the punching direction) were significantly higher in Post1–Post3 than in Pre. The negative values of torso separation angular velocity and angle (upper torso twisting relative to the lower torso) were significantly higher than Pre in Post1–Post3 for angular velocity and in Post3 for angle. These findings suggest that rapid countermovement of the lower and upper torso and torso separation play an important role in maintaining high punch speed even with increased progression of rounds.

Keywords: Performance analysis, Combat sports, Kinematics, Countermovement motion, Torso separation.

Cite this article as:

Kimoto, S., Ohyama, K., & Maeda, A. (2026). Changes in cross-punch speed and torso motion associated with round progression in amateur boxers. *Scientific Journal of Sport and Performance*, 5(3), 552-565. <https://doi.org/10.55860/LAQA2181>

 **Corresponding author.** National Institute of Fitness and Sports in Kanoya, 1 Shiromizu, Kanoya-city, Kagoshima 891-2393, Japan.

E-mail: amaeda@nifs-k.ac.jp

Submitted for publication March 09, 2026.

Accepted for publication April 23, 2026.

Published May 12, 2026.

[Scientific Journal of Sport and Performance](#). ISSN 2794-0586.

©Asociación Española de Análisis del Rendimiento Deportivo. Alicante. Spain.

doi: <https://doi.org/10.55860/LAQA2181>

INTRODUCTION

In amateur boxing, boxers typically compete in three rounds of three minutes each. Victory is determined either by points based on a scoring system that evaluates the number of high-quality strikes to target areas, or by knocking out the opponent. Amateur boxing as an activity is characterized by intermittent high-intensity exercise, and athletes are required to maintain high performance even as rounds progress (Davis et al., 2013; Davis et al., 2015).

Punches in amateur boxing mainly consist of four types: jabs, crosses, hooks, and uppercuts. Davis et al. (2018) analyzed high-level male amateur boxing matches and reported that winning requires the ability to continuously deliver a large number of cross punches and maintain punch accuracy even in the later rounds. A previous three-dimensional motion analysis study of cross punches by Cheraghi et al. (2014) reported that amateur boxers prioritize fast punches over heavy punches. Therefore, maintaining a fast and effective cross punch that is difficult for opponents to evade even in the later stages of a bout is considered crucial for amateur boxing performance.

Dunn et al. (2021) reported that repeated maximal rowing ergometer efforts significantly reduced cross-punch force during combination punching, suggesting that fatigue of the trunk muscles impairs trunk rotation and hinders the execution of powerful punches. In contrast, Hukkanen and Häkkinen (2017) reported that the force of a single cross punch after three rounds of a simulated bout was higher than before the bout, suggesting that fatigue-induced changes in body movement increase punch force; however, the detailed kinematic changes underlying this phenomenon remain to be elucidated. Previous studies have reported that torso rotation is important even in single cross punches (Lenetsky et al., 2013; Satkunskiene et al., 2024; Tong-lam et al., 2017). The torso has the greatest mass and moment of inertia among body segments (Ae, 1996) and contains a high proportion of muscle mass (Abe et al., 2003), indicating that it plays a major role in generating mechanical energy during movement. Nevertheless, to our knowledge, no study has specifically focused on the torso to examine changes in single cross-punch performance across rounds. El-Ashker (2011) reported that both single punches and combinations must be effectively utilized to win, indicating that investigating performance changes in single cross punches with a focus on torso movement across rounds will offer insights for enhancing amateur boxing performance.

This study employed an amateur boxing-specific punching performance test designed based on the movement patterns and exercise intensity of amateur boxing matches (Dunn et al., 2019). The reason for employing this test was to examine changes in single cross-punch speed and torso movement under conditions that replicate movement patterns and high-intensity progression similar to real matches. By replicating movement patterns and high-intensity progression similar to real matches, this approach can provide practical insights into key torso movements that enable the production of fast punches even in later stages of a bout.

The aim of this study was to investigate changes in cross-punch performance associated with round progression under match-like intensity conditions in amateur boxers.

METHODS

Participants

Eleven healthy male amateur boxers (age: 17.0 ± 1.0 years, height: 167.3 ± 3.8 cm, body mass: 65.3 ± 10.1 kg) participated in this study. Seven participants had an orthodox stance (left foot and left hand forward,

performing cross punches with the right hand), and four participants had a southpaw stance (right foot and right hand forward, performing cross punches with the left hand).

Prior to the experiment, the aim of the study, measurement procedures, and potential risks associated with the measurements were explained to the participants and their parents/guardians. Written informed consent for participation and publication was obtained. This study was conducted with the approval of the Ethics Committee of the National Institute of Fitness and Sports in Kanoya (approval number: 24-1-51).

Experimental protocol

After performing sufficient warm-up exercises, such as dynamic stretching, the participants put on boxing gloves and performed a 3-minute maximal-effort punch test (hereafter abbreviated as “3MPT”) against a suspended punching bag set at a height of 150 cm. The 3MPT (Dunn et al., 2019) is an amateur boxing-specific punching performance test developed based on the exercise intensity and punch frequency observed in amateur boxing matches. The test consists of a total of 3 minutes, divided into six cycles of 30 seconds each (Figure 1). Within each cycle, participants performed combination punches consisting of jabs, crosses, and hooks with maximal effort in a predetermined sequence in response to external cues given every 5 seconds, resulting in a total of 126 punches over the 3-minute period.

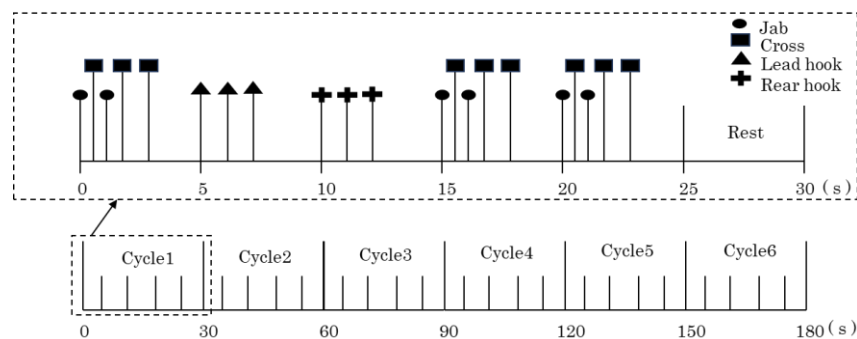


Figure 1. Schematic of the 3MPT (Dunn et al., 2019).

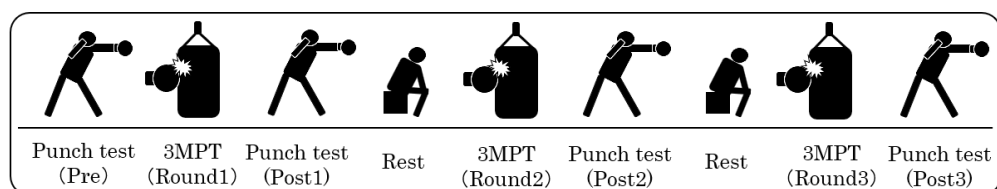


Figure 2. Flow of performance assessment.

We recruited participants who had no prior experience with the 3MPT. They practiced the task sufficiently beforehand and began the actual experiment after taking adequate rest. In the experimental session, participants first performed one trial consisting of a cross punch in a resting state (hereafter “Pre”). They then completed the first 3MPT (hereafter “Round1”), immediately followed by one trial consisting of a cross punch (hereafter “Post1”). Subsequently, after a 1-minute rest period, which aimed to replicate the conditions of an actual amateur boxing match, participants performed the second 3MPT (hereafter “Round2”), immediately followed by one trial consisting of a cross punch (hereafter “Post2”). After another 1-minute rest, they completed the third 3MPT (hereafter “Round3”), immediately followed by a trial consisting of one cross punch (hereafter “Post3”) (Figure 2). For each cross-punch trial, the participants’ standing position during Pre was

standardized to a fixed location that facilitated punching, and they were instructed to perform the cross punches at the same position for Post1 to Post3. In all trials, participants were instructed to perform with maximal effort. The cross-punch condition was limited to only one trial to minimize the influence of fatigue recovery over time and to maintain consistent experimental conditions.

Data collection

During each 3MPT, heart rate was measured using a chest-worn heart rate monitor (POLAR), and ratings of perceived exertion (RPE) were assessed using the Borg scale (Onodera & Miyashita, 1976). For RPE measurement, the Borg scale was displayed on a television monitor immediately after the completion of each 3MPT, and participants reported their perceived exertion based on the displayed scale. For the measurement of cross-punch performance in each trial, a 16-camera optical three-dimensional motion analysis system (Mac3D, Motion Analysis Corporation) equipped with Kestrel cameras was used. For motion capture, 47 reflective markers were attached to specific anatomical landmarks, and three-dimensional coordinates of body segments were recorded at a sampling frequency of 250 Hz with a shutter speed of 1/1500 s (Figure 3). The 47 reflective marker locations were as follows: three points on the head (vertex and bilateral tragus); 16 points on the upper limbs (bilateral anterior shoulder, bilateral posterior shoulder, bilateral acromion, bilateral medial epicondyle of the humerus, bilateral lateral epicondyle of the humerus, bilateral medial wrist, bilateral lateral wrist, and bilateral third metacarpal); 10 points on the trunk (anterior and posterior aspects of the suprasternal notch, anterior and posterior aspects of the xiphoid process, bilateral inferior borders of the ribs, bilateral anterior superior iliac spines, and bilateral posterior superior iliac spines); and 18 points on the lower limbs (bilateral higher trochanter, bilateral medial femoral condyle, bilateral lateral femoral condyle, bilateral medial malleolus, bilateral lateral malleolus, bilateral medial metatarsophalangeal joint, bilateral lateral metatarsophalangeal joint, bilateral calcaneal tuberosity, and bilateral toes). The global coordinate system was defined with the punching direction, representing the horizontal anterior–posterior axis, assigned to the X-axis; the mediolateral direction perpendicular to the X-axis assigned to the Y-axis; and the vertical direction assigned to the Z-axis.

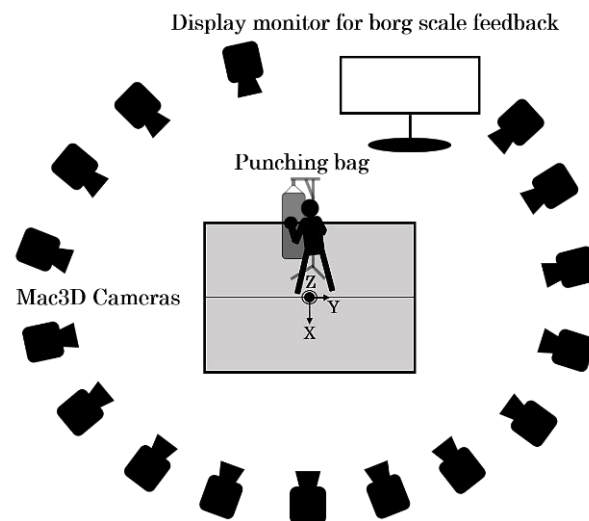


Figure 3. Overview of the performance assessment.

Data processing

The three-dimensional coordinates of the reflective marker attachment points were calculated using the Direct Linear Transformation method in the motion capture system control software (Cortex 6.0.0.1645,

Motion Analysis Corporation). The obtained three-dimensional coordinate data were smoothed using a fourth-order Butterworth low-pass filter. The cut-off frequency was set to 12 Hz, based on previous studies (Piorkowski et al., 2011; Stanley et al., 2018).

Motion analysis phase

The motion analysis phase was defined from the peak of the Z-coordinate of the centre of mass of the lead foot (0%) to one frame before the glove contacted the punching bag (100%), in consideration of potential coordinate distortion caused by impact with the punching bag, and this time point was normalized as the moment of bag contact (Figure 4). The moment at which the glove contacted the punching bag was defined as one frame before the point at which the three-dimensional resultant velocity calculated from the unsmoothed reflective marker on the third metacarpal rapidly decreased due to contact with the punching bag, using the impact detection method used in baseball batting (Ae et al., 2014). The centre of mass of the lead foot was calculated based on the body segment inertia coefficients reported by Ae (1996).

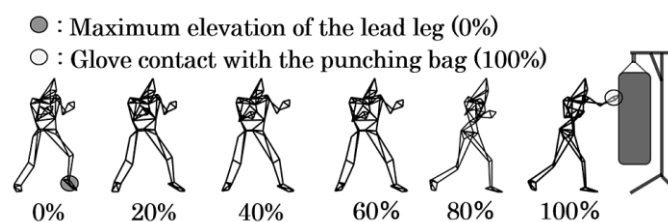


Figure 4. Definition of analysis phase.

Calculated variables

Motion time

The time from the peak of the Z-coordinate of the centre of mass of the lead foot (0%) to one frame before the glove contacted the punching bag (100%) within the motion analysis phase was defined as motion time.

Punch speed

In impact-related movements, it has been reported that appropriate peak velocities may not be obtained when coordinate data are smoothed in the usual manner prior to numerical differentiation (Kawamura et al., 2008). Therefore, punch speed was defined as the maximum value of the three-dimensional resultant velocity obtained by numerically differentiating the unsmoothed displacement of the reflective marker attached to the third metacarpal within the motion analysis phase. For participants using an orthodox stance, velocity was calculated from the reflective marker on the right third metacarpal, whereas for participants using a southpaw stance, velocity was calculated from the marker on the left third metacarpal.

Joint angles and angular velocities

Figure 5 illustrates the definitions of the movements of the lower torso, upper torso, and torso separation. The angles of each movement were calculated based on previous studies (Fuchs et al., 2018; Kageyama et al., 2014).

The rotational angle of the lower torso was defined as the angle between the vector connecting the midpoints of the pelvis (midpoint between the anterior superior iliac spine and posterior superior iliac spine on both sides) projected onto the XY plane and the Y-axis of the global coordinate system (Figure 5(a)). The rotational angle of the upper torso was defined as the angle between the vector connecting the midpoints of the left and right shoulder joints (midpoint of the anterior shoulder, posterior shoulder, and acromion) projected onto

the XY plane and the Y-axis of the global coordinate system (Figure 5(b)). For both the upper torso and lower torso, leftward rotation was defined as a positive angle, and rightward rotation as a negative angle. Furthermore, because the direction of movement is reversed depending on stance, the sign of the angles for the southpaw stance was adjusted to correspond to the rotation direction of the orthodox stance. Torso separation was defined as the difference between the upper torso angle and the lower torso angle (upper torso angle – lower torso angle) (Figure 5(c)). Accordingly, a negative torso separation angle indicates that the upper torso is rotated to the right relative to the lower torso, whereas a positive value indicates rotation to the left.

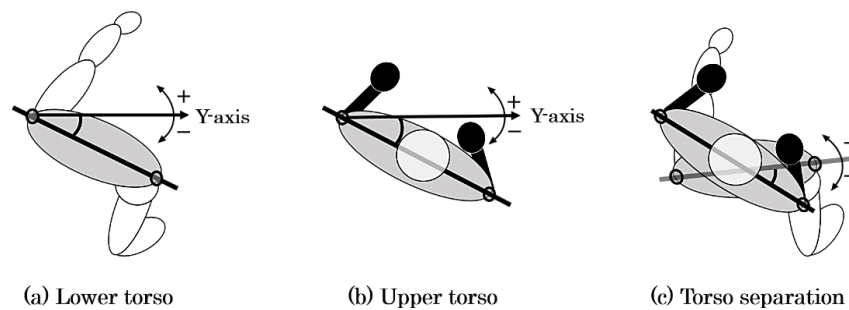


Figure 5. Definition of joint angles.

The angular velocities of the lower torso, upper torso, and torso separation were calculated by differentiating each angle with respect to time. The signs of the angular velocities were defined to correspond to the direction of rotation of each joint angle.

Statistical processing

Performance outcomes of cross-punch measurements in Pre and each Post condition (Post1, Post2, and Post3), as well as physiological indicators during the 3MPT, were compared using a linear mixed model adapted from Dunn et al. (2021) to examine changes in punch performance associated with fatigue. In the linear mixed model, the trial condition was treated as a fixed effect and participant as a random effect, and differences in performance outcomes and physiological indicators among trial conditions were compared. Bonferroni correction was applied as a post hoc adjustment for multiple pairwise comparisons. For normalized data, the same statistical procedure was conducted at each 1% time point. All statistical analyses were performed using IBM SPSS Statistics 29 (IBM Corp., Armonk, NY, USA), and the significance level was set at $p < .05$ for all tests.

RESULTS

Table 1 shows the highest heart rate and RPE during each 3MPT. The values of highest heart rate and RPE were significantly higher in Round2 and Round3 than Round1. Furthermore, Round3 had significantly higher values than Round2.

Table 1. Physiological indicators in each 3MPT round.

	Round1	Round2	Round3
Highest HR (bpm)	178.9 ± 10.0	186.6 ± 9.6 * #	190.1 ± 9.0 * #
RPE (Borg Scale)	14.0 ± 1.7	16.3 ± 1.8 * #	18.1 ± 2.0 * #

Note. Values are expressed as mean ± SD. * = significantly different to Round1 ($p < .05$). # = significantly different between Round2 and Round3 ($p < .05$).

Table 2 shows punch speed and motion time for Pre and each Post trial. No significant differences were observed between Pre and any of the Post conditions for either punch speed or motion time.

Table 2. Punch speed and motion time at Pre and each Post.

	Pre	Post1	Post2	Post3
Punch speed (m/s)	8.53 ± 0.98	8.66 ± 0.75	8.84 ± 0.87	8.93 ± 0.95
Motion time (s)	0.32 ± 0.04	0.32 ± 0.05	0.33 ± 0.08	0.32 ± 0.05

Note. Values are expressed as mean ± SD. * = significantly different to Round1 ($p < .05$). # = significantly different between Round2 and Round3 ($p < .05$).

Figure 6(a) shows the time-series changes in the rotational angle of the lower torso for Pre and each Post. In all trials, the lower torso rotated in the negative direction from 0% to approximately 40%, then changed to the positive direction after approximately 40%, reaching a peak near 100%. In Post1, positive rotation was significantly higher than Pre from 78% to 85%. In Post2, no significant changes were observed compared with Pre. In Post3, positive rotation was significantly higher than Pre from 69% to 92%. No significant changes were observed among the Post conditions.

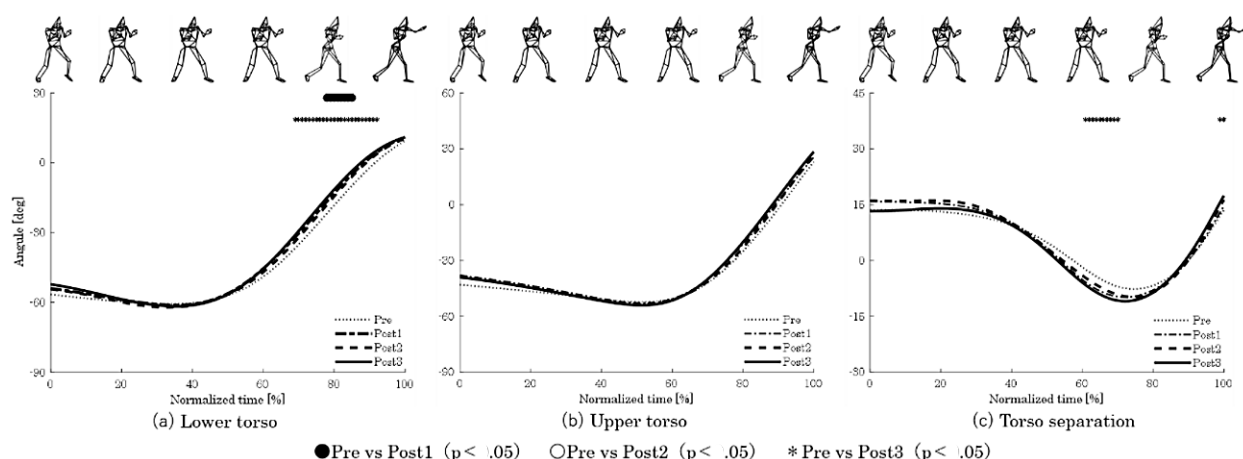


Figure 6. Time series data of axial rotation angle for each torso movement at Pre and each Post.

Figure 6(b) shows the time-series changes in the rotational angle of the upper torso for Pre and each Post. In all trials, the upper torso rotated in the negative direction from 0% to approximately 60%, then changed to the positive direction after approximately 60%, reaching a peak near 100%. No significant differences were observed between Pre and any Post condition at any time point. In addition, no significant changes were observed among the Post conditions.

Figure 6(c) shows the time-series changes in the torso separation angle for Pre and each Post. In all trials, the torso separation rotated in the negative direction from approximately 40% to 70%, then changed to the positive direction after approximately 70%, reaching a peak near 100%. In Post1 and Post2, no significant changes were observed compared with Pre. In Post3, the negative values were significantly higher than Pre from 61% to 70%, and the positive values were significantly higher than Pre from 99% to 100%. No significant changes were observed among the Post conditions.

Figure 7(a) shows the time-series changes in the rotational angular velocity of the lower torso for Pre and each Post. In all trials, the lower torso angular velocity showed negative values from 0% to approximately 40%, then reached a peak around 70%, followed by a decrease. In Post1, the positive values were

significantly higher than Pre from 50% to 58%, and in addition, the positive values were significantly lower than Pre from 90% to 100%. In Post2, the negative values were significantly higher than Pre from 13% to 22%, and the positive values were significantly lower than Pre from 90% to 100%. In Post3, the negative values were significantly higher than Pre from 7% to 26%, and the positive values were significantly higher than Pre from 51% to 71%. Furthermore, the positive values were significantly lower than Pre from 86% to 100%. No significant changes were observed among the Post conditions.

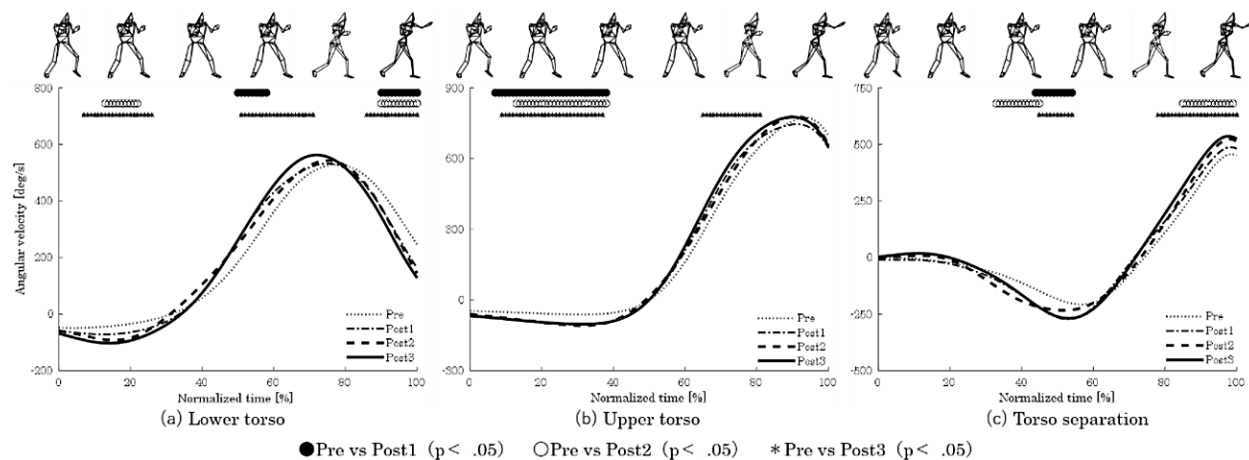


Figure 7. Time series data of axial rotation angular velocity for each torso movement at Pre and each Post.

Figure 7(b) shows the time-series changes in the rotational angular velocity of the upper torso for Pre and each Post. In all trials, the upper torso angular velocity showed negative values from 0% to approximately 50%, then reached a peak around 90%, followed by a decrease. In Post1, the negative values were significantly higher than Pre from 7% to 38%. In Post2, the negative values were significantly higher than Pre from 13% to 38%. In Post3, the negative values were significantly higher than Pre from 9% to 37%, and in addition, the positive values were significantly higher than Pre from 65% to 81%. No significant changes were observed among the Post conditions.

Figure 7(c) shows the time-series changes in the angular velocity of torso separation for Pre and each Post. In all trials, the angular velocity of torso separation showed negative values from approximately 20% to approximately 70%, then reached a peak near 100%. In Post1, the negative values were significantly higher than Pre from 44% to 54%. In Post2, the negative values were significantly higher than Pre from 33% to 45%, and the positive values were significantly higher than Pre from 85% to 99%. In Post3, the negative values were significantly higher than Pre from 44% to 54%, and the positive values were significantly higher than Pre from 78% to 100%. No significant changes were observed among the Post conditions.

DISCUSSION

This study aimed to investigate changes in cross-punch speed and torso movement associated with round progression in amateur boxers. We found that cross-punch speed was maintained despite the progression of rounds, whereas there were changes in torso movement.

Changes in physiological indicators with round progression

As shown in Table 1, the mean values of all physiological indicators in Round2 and Round3 were significantly higher than those in Round1. In addition, Round3 values were significantly higher than Round2. In this study,

Round1 to Round3 are positioned as trials corresponding to three rounds in an actual boxing match. A previous study that simulated boxing matches (El-Ashker et al., 2018) reported lower mean highest heart rate values for each corresponding round than the present study (Round1: 176 ± 4.1 bpm; Round2: 179 ± 1 bpm; Round3: 182 ± 5 bpm). Furthermore, in Round3 of the present study, the highest heart rate reached was 190.1 ± 9.0 bpm, which is higher than that reported by El-Ashker et al. (2018), indicating that an extremely high physiological load was imposed. On the other hand, the values were similar to those reported by Hukkanen and Häkkinen (2017) for Round1 (183 ± 6 bpm) and Round3 (191 ± 7 bpm). Regarding RPE, the mean values in Round2 and Round3 were significantly higher than in Round1, and a significant difference was also observed between Round2 and Round3. In particular, the RPE in Round3 reached 18.1 ± 2.0 , which indicates a very hard level of exertion. The mean RPE values for each round were higher in all corresponding rounds compared with a group that rested in a standing posture during simulated matches in a previous study (Nikolaidis et al., 2017).

Taken together, these findings suggest that each round in the present study elicited physiological responses that were higher than or comparable to those observed in previous simulated match studies in terms of both heart rate and RPE, and thus likely imposed a high-intensity load equivalent to that of actual amateur boxing matches.

Changes in punch speed and motion time with round progression

In the previous study by Dunn et al. (2021), a decline in punch performance was reported after high-intensity exercise targeting the lower limbs and trunk. However, Dunn et al. analysed cross punches during combination punching, whereas the present study focused on single punches immediately after the 3MPT. In Dunn et al. (2021), the researchers performed nine sets of 1-minute maximal rowing ergometer exercise with 1-minute rest intervals between sets, totalling 18 minutes. Therefore, compared with the present study using the 3MPT, the movement pattern differed and the physical load was higher.

Combination punching performance may decrease following high-intensity exercise. The present study used the 3MPT based on Dunn et al. (2019), which was designed based on the exercise intensity and punch frequency of amateur boxing matches. We found no changes in punch speed or motion with round progression (Table 2). These findings suggest that even under conditions similar to actual amateur boxing matches, it is possible to maintain single punch performance by modifying body movements. Hukkanen and Häkkinen (2017) reported that the force of a single cross punch after three rounds of a simulated bout was higher than before the bout, suggesting that changes in body movement with round progression are likely to influence punch performance. Furthermore, Piorkowski et al. (2011) reported that single punches exhibited significantly higher punch speed than combination punches. They attributed this to single punches allowing more time for countermovement compared with combination punches, and that the associated stretch-shortening cycle (SSC) contributes to increased punch speed. Faster punches are more likely to inflict higher damage when contacting the opponent's head or torso, potentially leading to knockouts and control of the bout. Therefore, because single punches are as important as combination punches for achieving victory, the present finding that single-punch speed did not change as the rounds progressed represents a novel contribution.

Changes in torso movement with round progression

A maximal-effort cross punch is initiated by lifting the lead leg and subsequently stepping forward, during which the pelvis and shoulders rotate in the opposite direction of the target, stretching the trunk musculature and inducing an SSC, which enhances cross-punch performance (Turner et al., 2011). The countermovement actions of the lower torso and upper torso are considered an important factor for generating high cross-punch

speed. As shown in Figure 7(a, b), significantly higher negative angular velocities than Pre were observed for the lower torso and upper torso in Post1 (upper torso: 7%–38%), Post2 (lower torso: 13%–22%; upper torso: 13%–38%), and Post3 (lower torso: 7%–26%; upper torso: 9%–37%). In addition, although no significant difference was observed in lower torso angular velocity in Post1, higher negative angular velocity than Pre was observed in the approximately 5%–30% interval, in which significant differences were observed for the lower and upper torso. In contrast, as shown in Figure 6(a, b), the negative rotational angles of the lower torso and upper torso in each Post did not differ significantly from Pre. These findings suggest that the amateur boxers unconsciously completed the movement within a limited time and as the rounds progressed, attempted to execute punches by increasing the speed of the countermovement of the lower torso and upper torso rather than the range of motion. This is also supported by the finding that motion time did not change significantly (Table 2). Previous studies examining the effect of countermovement speed on countermovement jump performance (Hernández-Davó et al., 2024; Pérez-Castilla et al., 2021) reported that performing the countermovement phase rapidly (from the initial posture to the negative peak of vertical velocity of the body centre of mass) improved performance in the propulsion phase (from the negative peak of vertical velocity to take-off), suggesting that this effect is due to higher force production early in the propulsion phase. Takahashi et al. (2018) reported that in bar-twist trunk rotation exercises, conditions involving countermovement resulted in significantly higher angular velocities of the upper torso and lower torso during the phase in which the barbell rotated in the positive direction, indicating that utilization of countermovement contributes to higher angular velocity production in the trunk. Considering these findings, the increase in countermovement speed of the lower torso and upper torso with round progression is likely related to the significantly higher positive angular velocities in Post1 (lower torso: 50%–58%) and Post3 (lower torso: 51%–71%; upper torso: 65%–81%) compared with Pre (Figure 7(a, b)), thereby contributing to the maintenance of punch speed across rounds. Even in trials where no significant differences were detected, the increased countermovement speed of the lower and upper torso may have contributed to maintaining positive angular velocity between approximately 50% and 80% of the movement. Furthermore, the increased positive angular velocity of the lower torso likely led to the significant increases in positive rotational angles of the lower torso for Post1 (78%–85%) and Post3 (69%–92%) (Figure 6(a)). In addition, significantly higher negative angular velocities of torso separation than Pre were observed in Post1 (44%–54%), Post2 (33%–45%), and Post3 (44%–54%) (Figure 7(c)). The increase in negative angular velocity indicates that in each Post, the upper torso rotated more rapidly to the right relative to the lower torso compared with Pre. This suggests that in each Post, in addition to the faster countermovement of the lower torso and upper torso, rapid torso separation may have further emphasized the SSC of the trunk. Effective force production via the SSC is reported to occur through the reuse of elastic energy stored in the muscle–tendon complex during the stretch phase as mechanical energy in the shortening phase (Asmussen & Bonde-Petersen, 1974; Cavagna, 1977), as well as through reflexive muscle contraction induced by rapid muscle stretch via muscle spindle activation (Komi, 2000; Van Hooren & Zolotarjova, 2017). Previous studies have reported that the SSC of the trunk influences performance in throwing and striking movements involving trunk rotation (Kageyama et al., 2014; Tauchi et al., 2005). A similar mechanism may apply to the participants in the present study. In particular, it has been shown that the enhancement of muscle output via the SSC is influenced by the velocity of muscle stretching during the stretch phase (Cavagna et al., 1968; Aura & Komi, 1986). Therefore, in each Post, where the countermovement speed of the lower torso and upper torso was higher than Pre and the upper torso rotated rapidly to the right relative to the lower torso, the enhanced SSC of the trunk likely contributed to maintaining punch speed. Furthermore, the effects of the SSC are known to be most pronounced immediately after the transition from the stretch phase to the shortening phase (Takamatsu et al., 1991; Walshe et al., 1998). This aligns with the present study, as we found that the positive angular velocities of torso separation were significantly higher than Pre in Post2 (85%–99%) and Post3 (78%–100%), suggesting that the SSC of the trunk becomes increasingly important for maintaining punch speed in later

rounds (Figure 7(c)). Myers et al. (2008), in their analysis of golf swings, reported that higher torso separation angles contribute to increased upper body rotational velocity and torso separation angular velocity during the downswing, ultimately enhancing ball speed. In the present study, the significantly higher negative torso separation angles in Post3 (61%–70%) compared with Pre may similarly have contributed to maintaining punch speed in the later rounds (Figure 6(c)). Moreover, the positive angular velocity of the lower torso was significantly smaller in Post1 (90%–100%), Post2 (90%–100%), and Post3 (86%–100%) compared with Pre (Figure 7(a)). Putnam (1993) reported that, in a kinetic chain, deceleration of proximal segments facilitates the transfer of kinetic energy to distal segments, promoting distal acceleration. Therefore, the significant reduction in lower torso angular velocity in each Post compared with Pre may have increased inertia and facilitated the transfer of energy to subsequent distal segments, thereby contributing to the maintenance of punch speed with round progression.

Limitations and future directions

This study had several limitations. First, changes in torso movement, such as countermovement actions of the lower torso and upper torso and torso separation, were observed with round progression in this study. However, such preparatory movements may increase the risk of opponents anticipating the intended action in actual competitive situations. Therefore, it remains unclear to what extent the present findings are applicable to real amateur boxing matches. Further research involving actual opponents is warranted in the future. Second, the 3MPT used in this study is a whole-body exercise involving not only the trunk but also the limbs. Future studies should employ force plates and electromyography to investigate changes in whole-body movement patterns associated with round progression. This approach may help elucidate the mechanical factors underlying why punch speed did not exceed Pre values in each Post despite changes in torso movement. Third, most participants in this study were junior amateur boxers. Although previous studies involving national-level athletes have also reported no decline in punch performance with round progression (Hukkanen & Häkkinen, 2017), further research involving elite boxers is necessary to determine whether the present findings can be generalized to higher-level competitors.

CONCLUSION

This study evaluated changes in cross-punch speed and torso movement with round progression in amateur boxers using the 3MPT. The results showed that punch speed did not change with round progression. However, the cross-punch movement changed such that faster countermovement actions of the lower torso and upper torso, and more pronounced torso separation were emphasized as rounds progressed. These changes may have contributed to maintaining punch speed across rounds.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. Shinya Kimoto was responsible for the conceptualization of the study, experimental design, data collection, data analysis and interpretation, and writing of the original draft. Kanzi Ohyama provided methodological guidance, assisted in data interpretation, and critically reviewed and edited the manuscript. Akira Maeda contributed through supervision and project administration, provided methodological guidance, assisted in data interpretation, critically reviewed and edited the manuscript, and handled correspondence with the journal. All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing, with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

- Abe, T., Kearns, C. F., & Fukunaga, T. (2003). Sex differences in whole-body skeletal muscle mass measured by magnetic resonance imaging and its distribution in young Japanese adults. *Br. J. Sports Med.*, 37(5), 436-440. <https://doi.org/10.1136/bjism.37.5.436>
- Ae, K., Koike, S., & Kawamura, T. (2014). Kinetic analysis of individual upper limbs during baseball tee-batting motion at different hitting-point heights. *Jpn. J. Phys. Educ. Health Sport Sci.*, 59(2), 431-452. <https://doi.org/10.5432/jpehss.13067>
- Ae, T. (1996). Body partial inertia coefficient of Japanese athletes and infants. *Jpn. J. Sports Sci.*, 15(3), 155-162.
- Asmussen, E., & Bonde-Petersen, F. (1974). Storage of elastic energy in skeletal muscles in man. *Acta Physiol. Scand.*, 91(3), 385-392. <https://doi.org/10.1111/j.1748-1716.1974.tb05693.x>
- Aura, O., & Komi, P. V. (1986). Effects of prestretch intensity on mechanical efficiency of positive work and on elastic behavior of skeletal muscle in stretch-shortening cycle exercise. *Int. J. Sports Med.*, 7(3), 137-143. <https://doi.org/10.1055/s-2008-1025751>
- Cavagna, G. A., Dusman, B., & Margaria, R. (1968). Positive work done by a previously stretched muscle. *J. Appl. Physiol.*, 24(1), 21-32. <https://doi.org/10.1152/jappl.1968.24.1.21>
- Cavagna, G. A. (1977). Storage and utilization of elastic energy in skeletal muscle. *Exerc. Sport Sci. Rev.*, 5, 89-129. <https://doi.org/10.1249/00003677-197700050-00004>
- Cheraghi, M., Alinejad, H. A., Arshi, A. R., & Shirzad, E. (2014). Kinematics of straight right punch in boxing. *Ann. Appl. Sport Sci.*, 2(2), 39-50. <https://doi.org/10.18869/acadpub.aassjournal.2.2.39>
- Davis, P., Wittekind, A., & Beneke, R. (2013). Amateur boxing: activity profile of winners and losers. *Int. J. Sports Physiol. Perform.*, 8(1), 84-91. <https://doi.org/10.1123/ijsp.8.1.84>
- Davis, P., Benson, P. R., Pitty, J. D., Connorton, A. J., & Waldock, R. (2015). The activity profile of elite male amateur boxing. *Int. J. Sports Physiol. Perform.*, 10(1), 53-57. <https://doi.org/10.1123/ijsp.2013-0474>
- Davis, P., Connorton, A. J., Driver, S., Anderson, S., & Waldock, R. (2018). The activity profile of elite male amateur boxing after the 2013 rule changes. *J. Strength Cond. Res.*, 32(12), 3441-3446. <https://doi.org/10.1519/JSC.0000000000001864>

- Dunn, E. C., Humberstone, C. E., Iredale, K. F., & Blazeovich, A. J. (2019). A damaging punch: Assessment and application of a method to quantify punch performance. *Transl. Sports Med.*, 2(3), 146-152. <https://doi.org/10.1002/tsm2.71>
- Dunn, E. C., Humberstone, C. E., Franchini, E., Iredale, K. F., & Blazeovich, A. J. (2021). The effect of fatiguing lower-body exercise on punch forces in highly trained boxers. *Eur. J. Sport Sci.*, 22(7), 964-972. <https://doi.org/10.1080/17461391.2021.1916085>
- El Ashker, S. (2011). Technical and tactical aspects that differentiate winning and losing performances in boxing. *Int. J. Perform. Anal. Sport*, 11(2), 356-364. <https://doi.org/10.1080/24748668.2011.11868555>
- El-Ashker, S., Chaabene, H., Negra, Y., Prieske, O., & Granacher, U. (2018). Cardio-respiratory endurance responses following a simulated 3 × 3 minutes amateur boxing contest in elite level boxers. *Sports*, 6(4), 119. <https://doi.org/10.3390/sports6040119>
- Fuchs, P. X., Lindinger, S. J., & Schwameder, H. (2018). Kinematic analysis of proximal-to-distal and simultaneous motion sequencing of straight punches. *Sports Biomech.*, 17(4), 512-530. <https://doi.org/10.1080/14763141.2017.1365928>
- Hernández-Davó, J. L., Sabido, R., Omar-García, M., & Boullosa, D. (2024). Why should athletes brake fast? Influence of eccentric velocity on concentric performance during countermovement jumps at different loads. *Int. J. Sports Physiol. Perform.*, 19(4), 375-382. <https://doi.org/10.1123/ijsp.2023-0273>
- Hukkanen, E., & Häkkinen, K. (2017). Effects of sparring load on reaction speed and punch force during the precompetition and competition periods in boxing. *J. Strength Cond. Res.*, 31(6), 1563-1568. <https://doi.org/10.1519/JSC.0000000000001885>
- Kageyama, M., Iwamoto, M., Sugiyama, T., Mizutani, M., Kanehisa, H., & Maeda, A. (2014). Effect of stretch-shortening cycle and pitching movements elicited by trunk rotation on ball velocity in university baseball pitchers. *Jpn. J. Phys. Educ. Health Sport Sci.*, 59(1), 189-201. <https://doi.org/10.5432/ijpehss.13014>
- Kawamura, T., Shimada, K., Takahashi, K., Morimoto, Y., Koike, S., & Ae, M. (2008). Comparison of kinematics of upper limb motion in baseball batting between high and low bat-speed groups. *Jpn. J. Phys. Educ. Health Sport Sci.*, 53(2), 423-438. <https://doi.org/10.5432/ijpehss.a530219>
- Komi, P. V. (2000). Stretch-shortening cycle: A powerful model to study normal and fatigued muscle. *J. Biomech.*, 33(10), 1197-1206. [https://doi.org/10.1016/S0021-9290\(00\)00064-6](https://doi.org/10.1016/S0021-9290(00)00064-6)
- Lenetsky, S., Harris, N., & Brughelli, M. (2013). Assessment and contributors of punching forces in combat sports athletes: Implications for strength and conditioning. *Strength Cond. J.*, 35(2), 1-7. <https://doi.org/10.1519/SSC.0b013e31828b6c12>
- Myers, J., Lephart, S., Tsai, Y.-S., Sell, T., Smoliga, J., & Jolly, J. (2008). The role of upper torso and pelvis rotation in driving performance during the golf swing. *J. Sports Sci.*, 26(2), 181-188. <https://doi.org/10.1080/02640410701373543>
- Nikolaidis, P. T., Clemente, F. M., Busko, K., & Knechtel, B. (2017). Physiological responses to simulated boxing: The effect of sitting versus standing body position during breaks: A pilot study. *Asian J. Sports Med.*, 8(3), e55434. <https://doi.org/10.5812/asjms.55434>
- Onodera, K., & Miyashita, M. (1976). A study on Japanese scale for rating of perceived exertion in endurance exercise. *Jpn. J. Phys. Educ. Health Sport Sci.*, 21(4), 191-203. <https://doi.org/10.5432/ijpehss.KJ00003405473>
- Pérez-Castilla, A., Rojas, F. J., Gómez-Martínez, F., & García-Ramos, A. (2021). Vertical jump performance is affected by the velocity and depth of the countermovement. *Sports Biomech.*, 20(8), 1015-1030. <https://doi.org/10.1080/14763141.2019.1641545>
- Piorkowski, B. A., Lees, A., & Barton, G. J. (2011). Single maximal versus combination punch kinematics. *Sports Biomech.*, 10(1), 1-11. <https://doi.org/10.1080/14763141.2010.547590>

- Putnam, C. A. (1993). Sequential motions of body segments in striking and throwing skills: Descriptions and explanations. *J. Biomech.*, 26(Suppl. 1), 125-135. [https://doi.org/10.1016/0021-9290\(93\)90084-R](https://doi.org/10.1016/0021-9290(93)90084-R)
- Satkunskiene, D., Bruzas, V., Mickevicius, M., Snieckus, A., & Kamandulis, S. (2024). Impact of leg strength on the force produced during a cross-punch in highly trained amateur boxers. *J. Strength Cond. Res.*, 38(10), 1739-1744. <https://doi.org/10.1519/JSC.0000000000004867>
- Stanley, E., Thomson, E., Smith, G., & Lamb, K. L. (2018). An analysis of the three-dimensional kinetics and kinematics of maximal effort punches among amateur boxers. *Int. J. Perform. Anal. Sport*, 18(5), 835-854. <https://doi.org/10.1080/24748668.2018.1525651>
- Takahashi, K., Kariyama, Y., Yoshida, T., Hayashi, R., & Asai, T. (2018). Characteristics of force and power outputs during the trunk-twist exercise accompanying stretch-shortening cycle movement. *Jpn. J. Phys. Educ. Health Sport Sci.*, 63(2), 641-657. <https://doi.org/10.5432/jjpehss.17110>
- Takamatsu, K., Aida, H., & Zushi, K. (1991). Effects of isometric and eccentric preliminary muscle contractions on elbow flexion velocity: With special reference to load and range of motion during concentric contraction. *Jpn. J. Phys. Educ. Health Sport Sci.*, 36(2), 127-139. <https://doi.org/10.5432/jjpehss.KJ00003391806>
- Tauchi, K., Minagata, K., Kawamura, T., & Takamatsu, K. (2005). Influence of the trunk twist on bat speed in baseball tee batting. *Jpn. J. Coach. Stud.*, 18(1), 1-9. https://doi.org/10.24776/jcoaching.18.1_1
- Tong-lam, R., Rachanavay, P., & Lawsirirat, C. (2017). Kinematic and kinetic analysis of throwing a straight punch: The role of trunk rotation in delivering a powerful straight punch. *J. Phys. Educ. Sport*, 17(4), 2538-2543. <https://doi.org/10.7752/jpes.2017.04287>
- Turner, A., Baker, E., & Miller, S. (2011). Increasing the impact force of the rear hand punch. *Strength Cond. J.*, 33(6), 2-9. <https://doi.org/10.1519/SSC.0b013e318232fdcb>
- Van Hooren, B., & Zolotarjova, J. (2017). The difference between countermovement and squat jump performances: A review of underlying mechanisms with practical applications. *J. Strength Cond. Res.*, 31(7), 2011-2020. <https://doi.org/10.1519/JSC.0000000000001913>
- Walshe, A. D., Wilson, G. J., & Ettema, G. J. C. (1998). Stretch-shorten cycle compared with isometric preload: Contributions to enhanced muscular performance. *J. Appl. Physiol.*, 84(1), 97-106. <https://doi.org/10.1152/jappl.1998.84.1.97>

