SPORTLOGIA

2/2025
Vol. 21, Issue 2

Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

E-ISSN 1986-6119
Print ISSN 1986-6089
CD-ROM ISSN 1986-6097
UDK/UDC 796/799

9 771986 608009

PUBLISHER / IZDAVAČ
University of Banja Luka, Faculty of Physical Education and Sport, Banja Luka, Bosnia and Herzegovina
Univerzitet u Banjoj Luci, Fakultet fizičkog vaspitanja i sporta, Banja Luka, Bosna i Hertegovina

EDITORIAL OFFICE / URED IZDAVAŠTVA
University of Banja Luka, Faculty of Physical Education and Sport
Univeritet u Banjoj Luci, Fakultet fizičkog vaspitanja i sporta
Bulevar Vojvode Petra Bojovića 1A
78000 Banja Luka, Bosna i Hercegovina
Phone/Fax: 00387 051 31 22 80
E-mail: info@ffvs.unibl.org
Web site: https://sportlogia.ffvs.unibl.org/ (full text available free of any charge)

ABSTRACT OR INDEXED-IN:
Cross Ref, Academic Search Premier, CAB Abstracts, EBSCO SPORTDiskus with Fulltext, EBSCOhost Research Databases, Fulltext Sources Online, Cab Global Health, Google Scholar, INASP - International Network for the Availability of Scientific Publication, Open Access Map, Science Gate, WorldCat.

Sportlogia journal (print ISSN 1986-6089, e-ISSN 1986-6119, CD-ROM ISSN 1986-6097) is published two times in one volumen per year till 2014 (every June and December), in one issue per year from 2015 until 2024, and since 2024 again in two issues per year by Faculty of Physical Education end Sports, University of Banja Luka, Bulevar Vojvode Petra Bojovića 1A, 78000 Banja Luka, Bosnia and Herzegovina.

SPORTLOGIA
Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 17/05/2025
Accepted: 20/06/2025
Online first: 01/09/2025
UDK: 796.323.012.1:159.9.072

Original article

Test-Retest Reliability of the 20-Metre Shuttle Run Test in a Cohort of Police Trainees

[1,2]*Elisa F.D. Canetti, [1,2] Ben Schram, [2,3] Rodney Pope, [4] Robert G. Lockie, [5]J.Jay Dawes, [2,7,8] Patrick Campbell, & [1,2]Robin Orr

[1]Faculty of Health Sciences and Medicine, Bond University, QLD Australia;
[2]Tactical Research Unit, Bond University, QLD, Australia;
[3]School of Allied Health, Exercise & Sports Sciences, Charles Sturt University, NSW, Australia;
[4]Center for Sport Performance, Department of Kinesiology, California State University - Fullerton, CA, USA;
[5]School of Kinesiology, Applied Health and Recreation, Oklahoma State University, OK, USA;
[6]Tactical Fitness and Nutrition Lab, Oklahoma State University, OK, USA;
[7]School of Behavioural and Health Sciences, Australian Catholic University, Brisbane, QLD 4014, Australia
[8]Sports Performance, Recovery, Injury and New Technologies (SPRINT) Research Centre, Australian Catholic University, Brisbane, QLD, Australia.

*Correspondence: rorr@bond.edu.au

Abstract

The 20-metre Multistage Fitness Test (20-m MSFT) is commonly used to measure aerobic capacity in police trainees as an entry gateway or exit requirement. However, its test-retest reliability, or consistency of scores for individual candidates across successive days, has not been determined. The aim of this study was to evaluate the test-retest reliability of the 20-m MSFT in police trainees. Retrospective data for 13 police trainees who completed the 20-m MSFT on two occasions 48 hours apart (Trial 1 and Trial 2) were provided. Paired sample t-tests were used to detect differences between individual performances with intraclass correlation coefficients (ICC) investigating the test-retest reliability. A Bland Altman plot was created to inspect the limits of agreement between the two measures. Alpha levels were set at 0.05 whereby a p value of >0.05 indicated no significant difference in mean scores between the two trials. No significant differences (p=0.821) between the mean total numbers of shuttles completed in Trials 1 (mean = 70.4±19.7 shuttles; Level 8-9) and 2 (mean = 69.8 ±21.3 shuttles; Level 8-9) were found. Six trainees achieved higher total shuttle scores for Trial 1 (+8.0 ±3.2 shuttles) while seven trainees achieved higher total shuttle scores for Trial 2 (+5.9± 5.1 shuttles). Test-retest reliability across trials was 'excellent' (ICC(3,1)=0.922 [95%CI 0.766-0.976], mean difference between scores = 0.55±8.37 shuttles). While the 20-m MSFT has excellent test-retest reliability the small amount of variability in results suggests that retesting of candidates who fail to meet any discriminatory standard by a small margin should be considered.

Keywords: Law Enforcement, test consistency, police academy, shuttle run, PSRT, beep test

Citation: F.D. Canetti, E et al. (2025). Test-Retest Reliability of the 20-Metre Shuttle Run Test in a Cohort of Police Trainees. *Sportlogia. 21(*2), 1-7. https://doi.org/10.63356/spl.2025.005

Introduction

Aerobic fitness is an important physical attribute for police trainees when completing their training academy (Lockie et al., 2019; Orr et al., 2020; Shusko et al., 2017; Tomes et al., 2020). Consequently, aerobic fitness assessments are often conducted during initial applicant screening and selection or during ab initio training (Lockie et al., 2019; Orr et al., 2020; Orr et al., 2022; Shusko et al., 2017; Tomes et al., 2020; Zulfiqar et al., 2021). During these assessments minimum requirements or 'cut scores' may be used to determine an applicant's suitability to train and potential to successfully perform the job tasks required of police officers. If aerobic fitness measures are to be used for hiring and retention purposes, it is important to understand whether scores on those measures are consistent between trials. In other words, how likely is it that the candidate would obtain a similar score if testing was conducted on another day relatively close to the time at which the initial test was performed (i.e., test-retest reliability). Test-retest reliability is of great importance given that potential variation could lead to candidate inclusion or exclusion or trainee pass or fail when on any other occasion the outcome may have been reversed.

Numerous factors may impact the results of physical fitness testing. Hopkins (2000) identified biological variability, and mental (e.g., effort, motivation) and physical (e.g., fatigue) state, as the main sources of within-subject deviations influencing test reliability. Furthermore, weather conditions (e.g., temperature, wind, humidity) (Sproule et al., 1993), environment (e.g., testing surface, noise, other people, etc.,) (Cooper et al., 2005), and hydration status (Lamb & Rogers, 2007) are all factors generally acknowledged to impact aerobic fitness testing results. Most notably for trainees undergoing ab initio training, insufficient recovery from daily stressors and training activities must be considered (Orr et al., 2016), as this has been shown to have a deleterious impact on performance (French & Ronda, 2021).

The 20-metre Multistage Fitness Test (20-m MSFT) is commonly used to assess aerobic fitness within law enforcement populations (Dawes et al., 2019; Lockie et al., 2021; Lockie et al., 2020; Orr et al., 2022; Zulfiqar et al., 2021). Considering this, broader research suggests that the test-retest reliability of the 20-m MSFT is generally 'high' to 'excellent', with reliability coefficients of between 0.87 and 0.98 (Aandstad et al., 2011; Knapik et al., 2004). However, these studies were conducted in military personnel, with no known similar study conducted among law enforcement personnel. This contextualization to the specific population is of importance given the myriad of factors discussed above that could influence assessment results. As such, the aim of this study was to investigate test-retest reliability of the 20-m MSFT in a population of police trainees.

Methods

Experimental approach to the problem

Retrospective data for 13 police trainees were provided by a state police agency. This data included trainee age and 20-m MSFT scores completed on two occasions, 48 hours apart (Trial 1 and Trial 2). All data were collected between 15:00-16:30 on the 28th and 30th November 2022, with each assessment preceded by a self-paced warm up. By ensuring the assessments were conducted at the same time of day, diurnal variations were mitigated. The processes put into place addressed concerns raised by French and Ronda (2021), with limited recovery between tests, lack of warm up protocols and diurnal variations being factors known to impact performance.

Participants

Data for 17 trainees who completed Trial 1 were provided. However, four (n=4) trainees were unable to attend Trial 2, due to either being on duty (n=3) or ill (n=1), leaving data for 13 trainees (males n=7; mean age = 29.6 ± 5.1 years; females n = 6, mean age = 27.7 ± 5.0 years) available for evaluation. The Bond University Human Research Ethics Committee granted ethics approval for this study (BUHREC, Research Protocol BS02086), with approval for public release of this report provided by the state police agency within which this research took place.

Measurements and Procedures

Prior to attempting the 20-m MSFT, all trainees were briefed by a police Physical Education Officer (PEO) to identify any trainees who might be at risk of suffering from adverse events during exercise. Trainees were given 48 hours between assessments. While police academy training could not be ceased in order to mitigate fatigue, the trainees did not conduct another physical assessment in the intervening period. All activities were conducted outdoors (mean temperature = 19.2 ± 1.8 (range = 18.0-21.2) °C, mean humidity = 67.2 ± 8.1 (range = 57.8-78.4)%, mean wind speed= 7.5 ± 1.1 (range = 6.0-8.6) km/h. Following the safety briefing, trainees were given approximately 10 minutes to warm up and were instructed to include at least five 20-m shuttle runs.

The 20-m MSFT protocols are described in the literature (Dawes et al., 2019; Lockie et al., 2021; Lockie et al., 2020), but for ease will be briefly described here. The 20-m MSFT, also termed the 'progressive shuttle run test', 'beep test', or 'bleep test', has participants run back and forth between two lines spaced 20 meters apart. The speed of running starts at 8.5km/h and increases by 0.5km/h every level, with each level lasting approximately one minute (Dawes et al., 2017). The running speed was standardized by pre-recorded auditory cues (i.e., beeps), played on a handheld iPhone device (Apple Inc., Cupertino, California) connected via Bluetooth to a portable speaker (Ultimate Ears, UE Boom 3, California, US), with participants required to reach the opposing line by the next beep (Dawes et al., 2017). The test was terminated when: a) the PEO informed the trainee that they had failed to reach the lines three times in a row in accordance with the auditory cues, or b) the trainee voluntarily withdrew. The assessment was scored by the police PEO as is common in this police agency. Final scores were presented as level and shuttle (e.g., Level 7 – Shuttle 5, or 'Level 7-5') before being converted to total number of shuttles for the analysis.

Statistical analyses

Data were provided digitally in a Microsoft Excel spreadsheet and prior to analysis were examined for accuracy and cleaned for any errors, with improbable (e.g., Level 1-13 as opposed to Level 11-3) or hard to read results clarified and corrected. Data were then imported into JASP (JASP Team 2023; Version 0.16.4) for analysis. Given that Hopkins (2000) suggests a change in mean scores between trials can be used as a measure of reliability, paired sample t-tests were performed, comparing scores from the two trials. Alpha level was set at 0.05, whereby a p value of >0.05 indicated no significant difference in mean scores between the two trials. Cohen's d was used to assess the effect size indicated by the difference between the two related cohort means (Cohen, 2013), i.e. the standardized mean difference. Interpretation of the effect size (ES) followed the guidelines proposed by Hopkins (2009), whereby ES was considered very small (0.00-0.19), small (0.20-0.59), moderate (0.60-1.19), large (1.20-1.99), very large (2.00-3.99), or extremely large (≥4.00). Raincloud plots were created for visualization of trainee and mean data as well as data distribution using JASP (JASP Team 2023; Version 0.16.4) statistical software. An intraclass correlation coefficient (ICC) was calculated to investigate the test-retest reliability, as previously used in the literature (Aandstad et al., 2011; Cuenca-Garcia et al., 2022). Using JASP (JASP Team 2023; Version 0.16.4), a two-way mixed model was selected. The type of ICC was set to absolute agreement between results from the repeated testing, using single test measurements in each testing episode. The 95% confidence interval (95% CI) was also reported for the ICC. Interpretation of the level of agreement indicated by the ICC was as follows: <0.50, poor; between 0.50 and 0.75, fair, between 0.75 and 0.90 good; above 0.90, excellent (Koo & Li, 2016). A Bland Altman plot was created to inspect the limits of agreement between the two measures (Bland & Altman, 1999). Limits of agreement were calculated as mean difference ± 1.96 standard deviation.

Results

There was no significant difference (t(1,12)=0.232, p=0.821, ES=0.064) between total numbers of shuttles completed in Trial 1 (mean = 70.4 ± 19.7 shuttles; Level 8-9) and Trial 2 (mean = 69.8 ± 21.3 shuttles; Level 8-9). Six of the 13 trainees achieved a higher total shuttle score for Trial 1, completing on average 8.0 (+3.2) shuttles less in Trial 2. The other seven trainees achieved a higher total shuttle score for Trial 2, completing on average 5.9 (+ 5.1) shuttles more than in Trial 1. Trainee and mean data are shown in Figure 1.

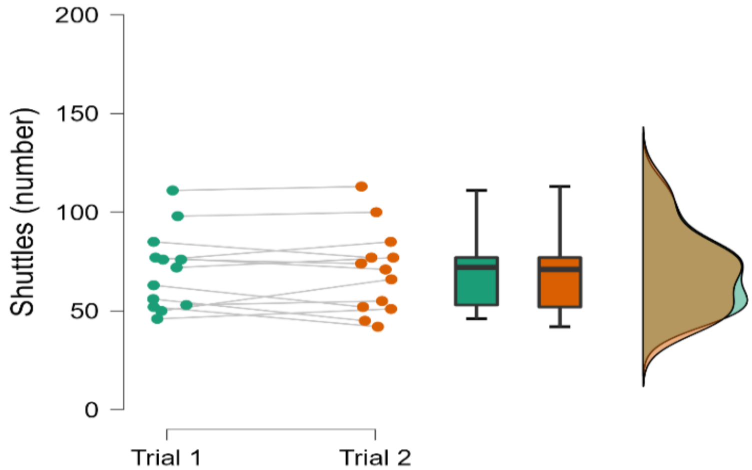


Figure 1. Depictions of individual trainee performance levels (scatter plot) and mean performance levels (box plots) as well as data distribution via a raincloud plot. Green markers denote Trial 1 results and orange markers, Trial 2. These plots show that the total shuttles performed in Trial 1 were similar to Trial 2.

Test-retest reliability of the 20-m MSFT for trial 1 and 2 was 'excellent', with an ICC(3,1)=0.922 [95%CI 0.766-0.976]. The mean difference in scores between trials was 0.55 [95%CI -4.52 to 5.60; SD 8.37] shuttles (Figure 2). The upper limit of agreement (LoA) was 16.95 [95%CI 8.19 to 25.71] shuttles and the lower LoA was -15.87 [95%CI -24.64 to -7.11] shuttles, with these LoA indicative of the 'typical error' in scores.

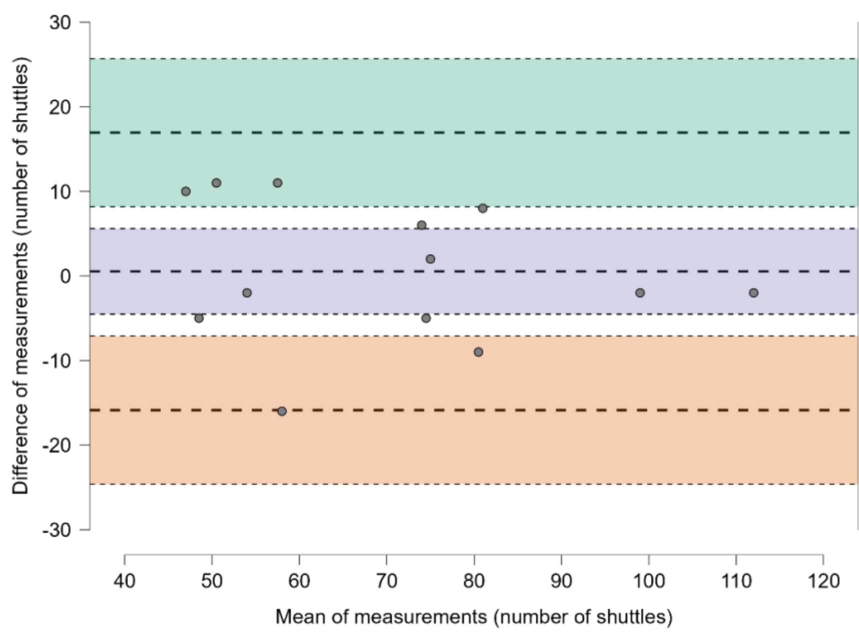


Figure 2. Bland-Altman plot indicating limits of agreement between the two trials. Figure shows 95% confidence intervals for mean (purple shaded area) and limits of agreement (upper in green and lower in orange shaded area).

Discussion

The aim of this study was to investigate the test-retest reliability of the 20-m MSFT in a population of police trainees. The results of the study showed no statistically significant difference between mean scores from two trials, indicating strong test-retest reliability. The ICC (reliability coefficient) of 0.922, which is considered to indicate an 'excellent' level of agreement between paired scores at the two testing timepoints (Koo & Li, 2016), supports such finding.

The test-retest reliability for the 20-m MSFT identified in the current study was consistent to the reliability reported in the literature. Recent systematic review by Cuenca-Garcia et al. (2022), concluded that the 20-m MSFT showed strong test re-test reliability, with seven out of nine high-quality studies reporting ICCs between 0.93–0.96 and correlation coefficients between 0.85–0.96. Similar strong reliability of the 20-MSFT was demonstrated in studies involving tactical populations. A US military report by Knapik et al (2004) documented test-retest reliability coefficients ranging from 0.87 to 0.98, which were higher than those for the 1-mile (1.6 km) and 2-mile (3.2 km) run tests (0.82–0.92). Similarly, Aandstad et al.,(2011) described a reliability coefficient of 0.95 for the 20-m MSFT in a study involving US soldiers.

While the mean difference in scores between trials was small (0.55 shuttles), the standard deviation of 8.4 shuttles indicate variability across trainees. This variability may be partly due to the small sample size or the short interval between trials (48 hours). Similar studies assessing test–retest reliability of the 20-m MSFT have reported intervals ranging from 2-4 days (Aandstad et al., 2011; Metsios et al., 2008) to 1-4 weeks (Cooper et al., 2005; Kim et al., 2011; Lamb & Rogers, 2007; Sproule et al., 1993), yet consistently reporting high reliability. In the present study, the LoAs, based on the SD, were approximately 16 shuttles. Despite the small sample size, these results are comparable to studies with larger sample sizes: Lamb and Rogers (2007) reported LoAs of 18 shuttles when assessing 35 university students, while Aandstad et al (2011) reported LoAs of ~10 shuttles for 41 Home Guard soldiers.

Although the 20-m MSFT demonstrated high group-level reliability, individual differences were evident, with six trainees performing better in Trial 1 and seven in Trial 2. Several factors may explain this variability. Familiarization with the pacing and auditory cues of the MSFT can improve performance on a second attempt, and prior studies have shown that one practice session is often sufficient to alter performance (Ramsbottom et al., 1988; Stickland et al., 2003). Conversely, poorer performance in Trial 2 may have been due to residual fatigue or incomplete recovery, as other studies have reported longer rest times between trials (Cooper et al., 2005; Kim et al., 2011; Lamb & Rogers, 2007; Sproule et al., 1993). Motivation and psychological factors may also have contributed to the variability observed as some trainees may have been more motivated during the first attempt, whereas others may have been more driven in the second trial by the opportunity to surpass their prior score (Neto et al., 2015) . Finally, mental stress has been shown to impair endurance performance by increasing perceived exertion and reducing pacing efficiency (Van Cutsem et al., 2017) , which could have contributed to lower shuttle scores in some individuals. Collectively, these findings highlight that while the 20-m MSFT is reliable at the group level, substantial within-subject variability may occur, and may be more prominent in studies with small sample size.

Several limitations of the present study should be acknowledged. First, the small sample size may have contributed to the wide 95% confidence interval (0.766–0.975) and amplified the influence of individual performance variability, as reflected by the standard deviation of 8.4 shuttles. Second, police academy training could not be suspended during the study period, meaning trainees were exposed to occupational stressors and cognitively demanding training activities (i.e., not physical), which are known to negatively affect 20-m MSFT performance and recovery (Macmahon et al., 2019; Slimani et al., 2018). Finally, although factors such as diurnal variation, physical fatigue, and rater consistency were controlled, these external stressors may have introduced additional variability in participants' physical and mental states between trials.

Conclusion

The findings of this study suggest that while there will be some variability in performance given a variety of individual factors (e.g., fatigue, etc.), the 20-m MSFT has excellent test-retest reliability. However, there were some small variations in scores and the associated between-trial variation could have implications for gateway testing. Thus, if the 20-m MSFT was to be used as a decision-making tool for training entry or completion, retesting of candidates or trainees who fail to meet any discriminatory standard by a small number of shuttles could be considered, with repeat testing within a few days feasible as the candidate or trainee could have the capacity to pass the standard on a different day.

Practical Implications

The test-retest reliability of the 20-m MSFT across two trials spaced 48 hours apart was 'excellent', with LoAs of approximately ±16 shuttles. The findings of this study demonstrate that the 20-m MSFT is generally test-retest reliable in this population. However, given there was some variability in individual results, re-testing of candidates who fail to meet any discriminatory standard by a small margin can be considered within a short timeframe (e.g., 48 hours), if allowed full rest, to ensure application of test results in decision-making is fair.

Acknowledgements: The authors acknowledge no conflicts of interest. This study was funded by the New Zealand Police Force and the authors would like to thank and acknowledge those personnel who volunteered to take part in this study.

References

Aandstad, A., Holme, I., Berntsen, S., & Anderssen, S. (2011). Validity and reliability of the 20 meter shuttle run test in military personnel. Mil Med, 176(5), 513-518. https://doi.org/10.7205/MILMED-D-10-00373

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. Stat Methods Med Res, 8(2), 135-160.

Cohen, J. (2013). Statistical power analysis for the behavioral sciences. Routledge.

Cooper, S. M., Baker, J. S., Tong, R. J., Roberts, E., & Hanford, M. (2005). The repeatability and criterion related validity of the 20 m multistage fitness test as a predictor of maximal oxygen uptake in active young men. Br J Soc Med, 39(4), e19-e19. https://bjsm.bmj.com/content/bjsports/39/4/e19.full.pdf

Cuenca-Garcia, M., Marin-Jimenez, N., Perez-Bey, A., Sanchez-Oliva, D., Camiletti-Moiron, D., Alvarez-Gallardo, I. C., Ortega, F. B., & Castro-Pinero, J. (2022). Reliability of field-based fitness tests in adults: a systematic review. Sports Med, 52(8), 1961-1979. https://link.springer.com/article/10.1007/s40279-021-01635-2

Dawes, J. J., Lockie, R. G., Orr, R. M., Kornhauser, C., & Holmes, R. J. (2019). Initial fitness testing scores as a predictor of police academy graduation. J Aust Strength Cond, 27(4), 30-37.

Dawes, J. J., Orr, R. M., Flores, R. R., Lockie, R. G., Kornhauser, C., & Holmes, R. (2017). A physical fitness profile of state highway patrol officers by gender and age. Annals of Occupational and Environmental Medicine, 29(16), 16. https://doi.org/10.1186/s40557-017-0173-0

French, D., & Ronda, L. T. (2021). NSCA's Essentials of Sport Science. Human Kinetics.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. Sports Med, 30, 1-15.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc, 41(1), 3-13. https://doi.org/10.1249/MSS.0b013e31818cb278

Kim, J., Jung, S., & Cho, H.-C. (2011). Validity and reliability of shuttle-run test in Korean adults. Int J Sports Med, 32(08), 580-585.

Knapik, J. J., Jones, B. H., Sharp, M. A., Darakjy, S., & Jones, S. (2004). The case for pre-enlistment physical fitness testing: Research and recommendations.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med, 15(2), 155-163.

Lamb, K. L., & Rogers, L. (2007). A re-appraisal of the reliability of the 20 m multi-stage shuttle run test. Eur J Appl Physiol, 100(3), 287-292. https://link.springer.com/article/10.1007/s00421-007-0432-9

Lockie, R., Balfany, K., Bloodgood, A. M., Moreno, M. R., Cesario, K. A., Dulla, J. M., Dawes, J. J., & Orr, R. M. (2019). The influence of physical fitness on reasons for academy separation in law enforcement recruits. Int J Env Res Pub He, 16(3), 372. https://mdpi-res.com/d_attachment/ijerph/ijerph-16-00372/article_deploy/ijerph-16-00372.pdf?version=1548735319

Lockie, R., Dawes, J. J., Moreno, M. R., Cesario, K. A., Balfany, K., Stierli, M., Dulla, J. M., & Orr, R. M. (2021). Relationship Between the 20-m Multistage Fitness Test and 2.4-km Run in Law Enforcement Recruits. J Strength Cond, 35(10), 2756-2761. https://doi.org/10.1519/JSC.0000000000003217

Lockie, R., Ruvalcaba, T. R., Stierli, M., Dulla, J. M., Dawes, J. J., & Orr, R. M. (2020). Waist circumference and waist-to-hip ratio in law enforcement agency recruits: Relationship to performance in physical fitness tests. J Strength Cond, 34(6), 1666-1675. https://doi.org/10.1519/jsc.0000000000002825

Macmahon, C., Hawkins, Z., & Schuecker, L. (2019). Beep test performance is influenced by 30 minutes of cognitive work. Med Sci Sports Exerc, 51(9), 1928.

Metsios, G. S., Flouris, A. D., Koutedakis, Y., & Nevill, A. (2008). Criterion-related validity and test–retest reliability of the 20 m square shuttle test. J Sci Med Sport, 11(2), 214-217.

Neto, J. M. D., Silva, F. B., De Oliveira, A. L. B., Couto, N. L., Dantas, E. H. M., & de Luca Nascimento, M. A. (2015). Effects of verbal encouragement on performance of the multistage 20 m shuttle run. Acta Scientiarum. Health Sciences, 37(1), 25.

Orr, R., Ferguson, D., Schram, B., Dawes, J. J., Lockie, R., & Pope, R. (2020). The relationship between aerobic test performance and injuries in police recruits. Int J Exerc Sci, 13(4), 1052. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7449329/pdf/ijes-13-4-1052.pdf

Orr, R., Knapik, J. J., & Pope, R. (2016). Avoiding program-induced cumulative overload (PICO). J Spec Oper Med : Peer Rev J SOF Med Prof, 16(2), 91-95.

Orr, R. M., Lockie, R., Milligan, G., Lim, C., & Dawes, J. (2022). Use of physical fitness assessments in tactical populations. Strength Cond J, 44(2), 106-113.

Ramsbottom, R., Brewer, J., & Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. Br J Sports Med, 22(4), 141-144.

Shusko, M., Benedetti, L., Korre, M., Eshleman, E. J., Farioli, A., Christophi, C. A., & Kales, S. N. (2017). Recruit Fitness as a Predictor of Police Academy Graduation. Occup Med (Lond), 67(7), 555-561. https://doi.org/10.1093/occmed/kqx127

Slimani, M., Znazen, H., Bragazzi, N. L., Zguira, M. S., & Tod, D. (2018). The effect of mental fatigue on cognitive and aerobic performance in adolescent active endurance athletes: insights from a randomized counterbalanced, cross-over trial. Journal of clinical medicine, 7(12), 510.

Sproule, J., Kunalan, C., McNeill, M., & Wright, H. (1993). Validity of 20-MST for predicting VO2max of adult Singaporean athletes. Br J Soc Med, 27(3), 202-204. https://bjsm.bmj.com/content/bjsports/27/3/202.full.pdf

Stickland, M. K., Petersen, S. R., & Bouffard, M. (2003). Prediction of maximal aerobic power from the 20-m multi-stage shuttle run test. Can J Appl Physiol, 28(2), 272-282.

Tomes, C. D., Sawyer, S., Orr, R., & Schram, B. (2020). Ability of fitness testing to predict injury risk during initial tactical training: a systematic review and meta-analysis. Injury Prev, 26(1), 67-81. https://injuryprevention.bmj.com/content/26/1/67.long

Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., & Roelands, B. (2017). The effects of mental fatigue on physical performance: a systematic review. Sports Med, 47(8), 1569-1588.

Zulfiqar, M. M., Wooland, J., Schram, B., Dawes, J. J., Lockie, R., & Orr, R. (2021). Battery fitness testing in law enforcement: A critical review of the literature. Int J Exerc Sci, 14(4), 613.

SPORTLOGIA
Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 09/09/2025
Accepted: 10/10/2025
Online first: 24/10/2025
UDK: 159.942:796.4.071

Original article

Differences in Personality Traits and Emotional Intelligence of Athletes and Non-Athletes: The Role of Gender and Performance Level

[1]*Duško Lepir, [2]Ivana Pedović

[1]Faculty of Physical Education and Sport, University of Banja Luka, Bosnia and Herzegovina
[2]Psychology department, Faculty of Philosophy, University of Niš, Serbia

*Correspondence: dusko.lepir@ffvs.unibl.org

Abstract

The increasing interest in the psychological impact of sports engagement has led to growing scientific efforts to understand whether athletes exhibit distinct psychological profiles compared to non-athletes. This study investigates differences in personality traits and emotional intelligence (examined both as a trait and as an ability) between athletes and non-athletes, with a particular focus on the role of gender and level of sports achievement. A total of 481 participants from Bosnia and Herzegovina took part in the study, including 277 athletes (58%) from 16 different sports and 204 non-athletes (42%). The sample comprised 260 males (163 athletes, 97 non-athletes) and 221 females (114 female athletes, 107 female non-athletes). Four validated instruments were employed to assess psychological traits: The Big Five Inventory (BFI), the Situational Test of Emotional Understanding (STEU), the Situational Test of Emotional Management (STEM), and the Emotional Skills and Competence Questionnaire (ESCQ). The results revealed that athletes score significantly higher in conscientiousness and extraversion and demonstrate better emotional management skills compared to non-athletes. However, they also exhibit lower openness to new experiences. These differences are more evident among male participants. No statistically significant differences were found between athletes of different performance levels (e.g., elite vs. recreational), suggesting that the mere involvement in sport, regardless of competitive level, may be linked to specific psychological traits. These findings underscore the importance of considering sport participation as a factor influencing personal development, particularly in relation to personality structure and emotional competence. However, the cross-sectional nature of the study limits causal interpretations. Therefore, future research should employ longitudinal designs and integrate sport-specific psychological assessments, while also considering socio-cultural contexts that may shape both sport experience and psychological functioning. Although additional investigation is necessary, the results are encouraging, indicating that sport practice may have positive effects on the development of enduring psychological characteristics, such as personality traits and emotional competence.

Keywords: emotional competence, Big Five, sports performance, diverse sports

Citation: Lepir, D. and Pedović, I. (2025). Differences in Personality Traits and Emotional Intelligence of Athletes and Non-Athletes: The Role of Gender and Performance Level. *Sportlogia. 21*(2), 8-23
https://doi.org/10.63356/spl.2025.006

Introduction

Research on the relationship between sports activities and psychological phenomena has a long tradition, with various psychological constructs being linked to (in)activity in sport and athletic success. This paper focuses on two constructs that are particularly relevant in the sports context: personality traits and emotional intelligence. Over the past 80 years, interest in personality in sport has gone through phases of intense research, neglect, and renewed attention (Allen et al., 2013). The most prevalent psychological framework in this field is trait theory, which defines traits as relatively stable characteristics of an individual that shape patterns of behavior, thinking, and emotions (Kassin, 2003). Factor analysis enabled the simplification of numerous identified traits, leading to the development of the dominant Big Five model (Goldberg, 1990), which identifies five major factors: extraversion, agreeableness, conscientiousness, neuroticism, and openness to new experiences. The application of the Big Five model in sports has taken various directions, linking personality traits with goal orientation (Wang & Erdheim, 2007), coping strategies (Allen et al., 2011; Kaiseler et al., 2012), burnout syndrome (Li et al., 2018), mental skills (Fabbricatore et al., 2023), emotional stability, and reasoning (Klatt et al., 2021). Additionally, some attention has been given to understanding the relationship between personality traits and athletic success. Research shows that higher-level athletes tend to exhibit greater conscientiousness and lower neuroticism (Piedmont et al., 1999), as well as varying levels of agreeableness (Allen et al., 2011; Steca et al., 2018), extraversion, agreeableness, and conscientiousness (Siemon & Wessels, 2022). Review studies confirm the predictive value of the Big Five, emphasizing that, aside from neuroticism, all other traits positively correlate with athletic success (Allen et al., 2013; Piepiora et al., 2024; Shuai et al., 2023).

Personality traits have also been linked to sports (in)activity. A review study indicates that athletes are generally more extraverted, emotionally stable, and open to new experiences compared to non-athletes (Allen et al., 2013). However, other research has found that athletes score higher on all Big Five dimensions except for openness to new experiences (Steca et al., 2018). The examination of gender effects on the relationship between personality traits and sport (in)activity has mostly been of an intra-group nature, with general conclusions indicating that results found in the non-athlete population are mirrored in the athlete population, given that differences between men and women were established in both cases (Allen et al., 2011, 2013). However, a question that has remained in the background of previous research, but which deserves deeper analysis, involves understanding the differences between male non-athletes and athletes, and analogously among females. A few studies indicate the existence of such differences, emphasizing that male athletes are more conscientious than male non-athletes (Malinauskas et al., 2014), and that elite athletes display higher levels across all personality traits except openness (Steca et al., 2018). Even more scarce are findings about differences between female athletes and non-athletes. One study in the domain of martial arts found that female athletes exhibited lower levels of emotional reactivity, though it should be noted that this study was not based on the Big Five model (Burdzicka-Wołowik & Góral, 2014). Given the plurality of theoretical approaches, the variety of instruments applied, and the wide research potential of the personality trait construct, it is not surprising that experts today describe it as an "ever-present research direction" in sport (Piepiora et al., 2024). In addition to arguments supporting the need for further investigation of the link between personality traits and sport, researchers also point out several shortcomings of previous studies. These include small sample sizes per study (a minimum of 40 participants, with most studies not exceeding 200), excessive heterogeneity, limitations to one sport without consideration of categorizing athletes by skill level (Steca et al., 2018), and the lack of studies addressing the socio-cultural context (Shuai et al., 2023).

Sports activities are rich in complex socio-psychological dynamics that generate intricate affective states, which is why experts emphasize that emotions are essential aspects of sporting activities (Vallerand & Blanchard, 2000) and deserve special attention. Accordingly, the construct of emotional intelligence (EI), which represents an individual's ability to recognize and effectively manage emotions to achieve personal and social well-being (Mayer et al., 2008), has attracted researchers' attention in the sports domain. Naturally, the interest was focused on understanding the relationship between EI and sports success, with analyses confirming statistically significant correlations indicating that higher-level athletes exhibit greater emotional intelligence (Kopp & Jekauc, 2018; Laborde et al., 2016; Merino Fernández et al., 2022; Montenegro-Bonilla et al., 2024), and also that athletes with higher emotional intelligence achieve better results on ability tests or in actual competitions (Gatsis et al., 2021; Haryanto et al., 2024). Studies also focused on the link between EI and participation in sports or physical exercise, with results generally confirming that individuals who engage in sports or are more physically active possess higher levels of EI (Bostani & Saiiari, 2011; Lepir et al., 2018; Ubago-Jiménez et al., 2019). One explanation is that EI is a good precondition for developing self-regulated learning (planning, self-monitoring, effort, and self-confidence), which ultimately serves as a significant predictor of greater involvement in activities that are generally beneficial for the individual, such as physical exercise (Cecchini et al., 2019). When it comes to gender effects, studies of the general population usually indicate higher levels of EI in women (Extremera et al., 2006; Kafetsios, 2004), while results in sports samples are more contradictory, ranging from higher EI in men or in women, to no observed differences at all (Merino Fernández et al., 2019; Mon-López et al., 2023; Popovych et al., 2024; Rodriguez-Romo et al., 2021). The inconsistency of these findings, along with the pronounced neglect in understanding differences between physically active and inactive men and women, calls researchers to devote more thorough attention to these relationships. The examination of EI, both within and outside the context of sports, also raises the issue of its conceptualization and operationalization. One group of researchers views EI as a personality trait and accordingly uses self-report questionnaires for its assessment, while another group considers it an ability that should be measured using performance-based tests (Laborde et al., 2016). The same authors also point to the dominance of the former approach in the field of sports, indicating the need for integrating both approaches in future research. In addition to the unresolved effects of sociodemographic characteristics and methodological limitations related to operationalization, experts also cite other arguments supporting the continued investigation of EI in the sports context – such as the overrepresentation of student populations in existing studies (Ubago-Jiménez et al., 2019).

Based on recommendations from previous studies regarding the need for further exploration of these constructs in sports, the aim of this paper is to examine differences between non-athletes and athletes across the entire sample, as well as within male and female subgroups, and to assess differences among athletes of varying performance levels in terms of personality traits (using the Big Five model) and emotional intelligence (using a dual conceptualization approach: EI as a trait and as an ability). It is expected that athletes will show higher conscientiousness, emotional stability, extraversion, and emotional intelligence than non-athletes. Regarding gender, it is anticipated that both male and female athletes will demonstrate stronger adaptive and weaker maladaptive traits, as well as higher emotional intelligence compared to their non-athlete counterparts, and similar pattern is anticipated between higher and lower performing athletes. Acknowledging the methodological challenges of previous studies, and in addition to emphasizing the specificities in the application of conceptualizations and operationalizations of these constructs, this study, by including a culturally specific, relatively larger, and sport-type diverse sample of participants from Bosnia and Herzegovina, aims to contribute to clarifying insufficiently explored aspects of this issue.

Methods

Participants and Procedure

The study included 481 participants, comprising 277 athletes (58%) from 16 different sports and 204 non-athletes (42%). The sample included 260 male participants (54%), of whom 163 were athletes (92 from team sports and 71 from individual sports) with an average age of 22.2 ± 5.9 years and 11.5 ± 5.7 years of sporting experience, and 97 were non-athletes with an average age of 20.4 ± 3.4 years. The female sample consisted of 114 athletes (52%) with an average age of 20.61 ± 5.5 years and 9.5 ± 5 years of sporting experience (78 from team sports and 36 from individual sports), and 107 non-athletes with an average age of 19 ± 3.6 years. To be classified as an athlete, participants were required to have at least two years of active sports participation and to regularly compete at a minimum of the regional level. In addition to the non-athlete category, three categories of athletic achievement were created based on the participants' current level of competition: low (n = 84), medium (n = 97), and high achievement (n = 96). This sample size was found to be adequate for conducting ANOVA with post hoc Bonferroni corrections for four participant groups, assuming a medium effect size (Cohen's f = 0.25). A power analysis conducted using G*Power software (Faul et al., 2007) indicated a statistical power of 0.98, exceeding the recommended threshold of 0.80 for detecting significant effects (α = .005 after Bonferroni correction).

Creating unified categories of sport expertise based on objective competitive achievement is a highly challenging task, especially when working with a diverse sample of athletes from various sports, genders, and age groups. To address this challenge, team sport athletes were first categorized based on the following criteria: those competing at the regional level (entity levels of Republic of Srpska and the Federation of BiH) were classified as low-level athletes; those competing in the first league of BiH were placed in the medium category; and athletes from clubs that were current national champions and competing at European international competitions were placed in the high category. For individual sports, a three-level categorization of performance was conducted using an expert method. The experts were individuals holding a doctoral degree in sport sciences and with at least ten years of coaching experience at the national or international level in the respective sport. They developed precise classification criteria for each individual sport. For example, in karate, individuals whose highest achievement was participation in national championships were categorized at the lowest level; national champions and participants in Balkan Championships, Mediterranean Games, Universiade, and the World Premier League were classified as medium level; while those ranked among the top 8 at European Championships, top 16 at World Championships, and top 3 at Balkan or Mediterranean Championships were placed in the high-level category.

The data collection process was carried out across various locations in Bosnia and Herzegovina using the selected instruments. Athletes were invited to participate through sports organizations such as clubs, federations, and sporting events, as well as through educational institutions, including secondary schools and universities. Before taking part in the study, all participants were thoroughly briefed on its purpose and signed a consent form, with parental or guardian consent obtained for underage athletes. The data collection was conducted anonymously, and the presence of the researcher ensured equal conditions for all participants, facilitated the completion of the instruments, and minimized any potential negative influence from coaches or other officials within the clubs. Ethical approval was obtained from the Faculty Ethics Committee (approval no. 11/1.624-2/25). The study was conducted in accordance with the Declaration of Helsinki.

Instruments

Big Five Inventory – BFI (John et al., 1991) is a self-report instrument for assessing personality traits based on the Big Five model. The instrument includes 44 items measuring five personality dimensions: conscientiousness, extraversion, neuroticism, agreeableness, and openness to experience. For each item, participants evaluate how characteristic the given statement or specific behavior is for them using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) (example: "I see myself as someone who does a thorough job"). The BFI has shown good internal consistency (α = .70–.85), acceptable construct validity, with confirmatory factor analyses supporting the five-factor structure (CFI ≈ .90), and convergent validity demonstrated through strong correlations with the NEO-FFI and other measures of personality (John & Srivastava, 1999). For this sample, the following alpha coefficients were obtained: conscientiousness .81, extraversion .80, neuroticism .71, agreeableness .72, and openness to experience .79.

The Emotional Skills and Competence Questionnaire – ESCQ45 (Takšić, 2002) is designed for the self-assessment of emotional competence. It is based on the Mayer and Salovey (1997) model and measures three key aspects of emotional intelligence: perceiving and understanding emotions (PUE), expressing and labelling emotions (ELE), and managing and regulating emotions (MRE). It consists of 45 items, with participants rating their level of agreement with each statement on a five-point Likert scale, where 1 indicates "never" and 5 "always" (example: "I can almost always describe my feelings and emotions in words"). The ESCQ45 has demonstrated good internal consistency (α = .78–.89) and construct validity, with confirmatory factor analysis supporting the three-factor model (CFI = .91; RMSEA = .06), and convergent validity confirmed through positive correlations with the EQ-i and other measures of emotional competence (Takšić et al., 2009). For this sample, the following alpha coefficients were obtained: PUE (α = .83), ELE (α = .79), and MRE (α = .78).

Situational Test of Emotional Understanding – STEU (MacCann & Roberts, 2008). This test consists of 42 scenarios, and participants are asked to determine which of the five offered emotions is most likely to be experienced by a person in a given situation (example: "Something unpleasant is happening. Neither the person involved, nor anyone else can make it stop. The person involved is most likely to feel: a) Guilty, b) Distressed, c) Sad, d) Scared, e) Angry"). Each scenario has only one correct answer, determined through the target method, i.e., the authors defined the correct emotion (MacCann et al., 2004). Scoring is based on a mode system, meaning that the correct answer receives one point, while all others receive zero. STEU has been shown to have acceptable reliability (Cronbach's α ≈ .71 in the original version), good convergent validity with MSCEIT - Understanding branch, vocabulary, personality traits, and emotional-/stress-related outcomes, and independence from trait EI, and discriminant validity indicated by low correlations with other EI components and intelligence (Austin, 2010; MacCann & Roberts, 2008). Following the authors' guidelines (see MacCann et al., 2011), five items with zero or negative correlations with the corrected total score were removed, reducing the test to 37 items. For this version, the obtained alpha coefficient was .71.

Situational Test of Emotional Management – STEM (MacCann & Roberts, 2008). The original test comprises 44 items and assesses the most complex aspect of emotional intelligence: the ability to modify one's own or others' emotional states. For each of the 44 situations, participants select one of four possible actions they believe would be most effective for the person in the given situation (example: "Andre moves away from the city his friends and family are in. He finds his friends make less effort to keep in contact than he thought they would. What action would be the most effective for Andre? a) Try to adjust to life in the new city by joining clubs and activities there, b) He should make the effort to contact them, but also try to meet people in his new city, c) Let go of his old friends, who have shown themselves to be unreliable, d) Tell his friends he is disappointed in them for not contacting him"). The expert method was used to determine the correct answers, and scoring was conducted using a proportional scoring system (MacCann et al., 2004; Takšić et al., 2006). STEM demonstrated reliability of about α = .68 in the original study, meaningful correlations with emotional understanding, agreeableness, and life satisfaction, supporting its construct and criterion validity. (STEM; MacCann & Roberts, 2008). The STEM also shows good convergent validity through moderate correlations with the MSCEIT "Managing Emotions" branch, indicating it captures emotional management ability, and demonstrates discriminant validity by not correlating with trait emotional intelligence measures, confirming it assesses ability rather than self-perceived emotional traits (Austin, 2010). In this study, due to zero or negative correlations, three items were excluded from further operationalization, resulting in a reduced version of the test with 41 items. For this sample, the obtained alpha coefficient was .80.

Statistical analyses

In the first step, descriptive statistical parameters were calculated (M – mean, SD – standard deviation, Mdn – median, Min – minimum value, Max – maximum value), as well as coefficients of skewness and kurtosis. Values of skewness and kurtosis between -1 and 1 were considered acceptable for the use of parametric tests (Kim, 2013). To examine multicollinearity between variables, Pearson correlation analysis was used. Independent samples t-tests were conducted to test differences between groups of athletes and non-athletes, as well as between male and female athletes and non-athletes. In order to eliminate potential effects of demographic factors (gender and age), residuals obtained through linear regression analysis were used in further group difference analyses. The large number of tests conducted necessitated the use of the Bonferroni correction. In the case of examining differences among four groups of participants with varying levels of sports activity, a one-way analysis of variance (ANOVA) was applied, with Bonferroni post hoc analyses. Effect sizes were calculated using Cohen's d for pairwise comparisons, with thresholds of 0.2, 0.5, and 0.8 representing small, medium, and large effects, respectively (Cohen, 2013). For ANOVA, partial eta squared (η^2) was calculated to indicate the proportion of variance explained by group differences, with 0.01, 0.06, and 0.14 considered small, medium, and large effects, respectively (Cohen, 2013). Statistical analyses were conducted using SPSS software version 21 (IBM Corp., 2012).

Results

Descriptive parameters for all examined variables are presented in Table 1. Analysis of the obtained values indicates that skewness and kurtosis are within acceptable limits (between -1 and 1) for all variables, suggesting that the distributions are not significantly distorted, and thus the use of parametric statistical methods is justified. To examine potential multicollinearity among the investigated variables, Pearson correlation coefficients were analyzed. It was determined that the coefficients did not exceed the critical threshold of r < .70 (see Pallant, 2020), indicating that the variables measure different constructs and should be analyzed separately.

Table 1. Descriptive statistics.

Measures	***M***	***SD***	***Mdn***	***Skew***	***Kurt***	***Min***	***Max***
Extraversion	3.70	0.65	3.75	-0.35	-0.27	1.63	5.00
Agreeableness	3.61	0.57	3.67	-0.49	0.30	1.67	4.89
Conscientiousness	3.72	0.60	3.78	-0.27	0.05	1.67	5.00
Neuroticism	2.78	0.59	2.75	0.22	-0.27	1.25	4.63
Openness	3.48	0.66	3.44	0.02	-0.34	1.44	5.00
ESCQ - PUE	3.70	0.51	3.67	0.08	0.16	2.13	5.00
ESCQ - ELE	3.64	0.55	3.64	0.11	-0.27	2.07	5.00
ESCQ - MRE	3.91	0.42	3.94	-0.33	0.60	2.50	5.00
STEU	21.29	5.08	22	-0.57	0.35	5.00	33.00
STEM	15.77	4.25	16.92	-0.39	-0.74	6.08	25.33

Note. *N* = 481. *M*= mean, *SD*= Standard Deviation, *Mdn*= Median**,** *Skew***=** Skewness, *Kurt***=** Kurtosis, Min= Minimum, Max= Maximum, ESCQ – PUE = emotional competence of perceiving and understanding emotions, ESCQ – ELE = emotional competence of expressing and labelling emotions, ESCQ – MRE = emotional competence of managing and regulating emotions, STEU = situational test of emotional understanding, STEM = situational test of emotion management

Table 2 presents the differences between athletes and non-athletes in terms of the five personality traits, three self-report emotional competence questionnaire scales, and two emotional intelligence test scales. In order to eliminate potential influences of demographic characteristics, residuals controlled for gender and age were used. This ensured a more objective interpretation of group differences, i.e., it ensured that the observed differences between groups were not due to varying distributions in terms of gender and age, but rather reflected actual differences in the characteristics being examined. Due to the large number of conducted t-tests, the Bonferroni correction was applied (setting the significance threshold at p < .005) in order to control for cumulative Type I errors. According to this correction, statistically significant differences in personality traits were found for extraversion, conscientiousness, and openness to experience. The first two traits were more pronounced among athletes, while openness to experience was more characteristic of non-athletes. Among the three self-report questionnaire scales, a statistically significant difference was found only for the highest branch of emotional intelligence, with analysis indicating that athletes report greater *emotional competence in managing and regulating emotions* compared to non-athletes. Regarding emotional intelligence tests, statistically significant differences in favor of athletes were also confirmed for the highest branch of emotional intelligence, labeled as *situational emotion management*.

Table 2. Comparison of athletes and non - athletes (residuals controlled for gender and age).

Measures	Athletes (n=277) (*M*±*SD*)	Non-Athletes (n=204) (*M*±*SD*)	*t*(479)	*p*	Cohen's *d*
Extraversion	0.15 ± 0.94	-0.20 ± 1.05	3.88	<.001*	0.36
Agreeableness	0.07 ± 0.95	-0.09 ± 1.06	1.70	.089	0.16
Conscientiousness	0.14 ± 0.98	-0.19 ± 0.99	3.66	<.001*	0.34
Neuroticism	-0.06 ± 1.00	0.09 ± 0.99	-1.61	.108	0.15
Openness	-0.12 ± 0.95	0.16 ± 1.04	-3.03	.003*	0.28
ESCQ - PUE	0.03 ± 0.91	-0.04 ± 1.10	0.79	.432	0.07
ESCQ - ELE	0.09 ± 0.96	-0.12 ± 1.03	2.31	.021	0.21
ESCQ - MRE	0.20 ± 0.94	-0.27 ± 1.02	5.19	<.001*	0.47
STEU	0.09 ± 1.01	-0.12 ± 0.98	2.22	.027	0.20
STEM	0.12 ± 0.95	-0.17 ± 1.04	3.14	.002*	0.29

Note. *t*= t-value, Cohen's *d*= effect size. * $p < .005$ indicates significance after Bonferroni correction

Given that the Bonferroni correction method is quite conservative, which increases the likelihood of missing genuinely significant effects, it should be noted that according to the conventional threshold ($p < .05$), athletes also achieve statistically significantly higher scores on the *expressing and labeling emotions* scale of the self-report emotional competence questionnaire and on the *situational understanding of emotions* test. An analysis of effect sizes based on Cohen's d values suggests that, although statistically significant, the differences mostly fall within the category of small to moderate effects (Cohen, 2013). This indicates that, while there are differences between athletes and non-athletes, they are not particularly large.

Differences in personality traits and emotional intelligence between male athletes and non-athletes are presented in Table 3, with age controlled as a covariate. Athletes reported statistically significantly higher ($p < .005$) self-assessed *emotional competence in managing and regulating emotions*, as well as higher scores on the situational test of *emotional management*. On the other hand, non-athletes exhibited significantly higher levels of *openness*. Cohen's d values indicate that these differences fall within the category of moderate effect size. According to the conventional significance threshold ($p < .05$), athletes also scored significantly higher on conscientiousness and situational understanding of emotions.

Table 3. Comparison of male athletes and non - athletes (residuals controlled for age).

Measures	Male Athletes (*n*= 163) (*M* ± *SD*)	Male Non-Athletes (*n*= 97) (*M* ± *SD*)	*t*(258)	*p*	Cohen's *d*
Extraversion	.05 ± .59	-.08 ± .69	1.610	.109	.21
Agreeableness	.02 ± .49	-.03 ± .61	.634	.527	.08
Conscientiousness	.07 ± .58	-.11 ± .60	2.419	.016	.31

Measures	**Male Athletes (*n*= 163) (*M* ± *SD*)**	**Male Non-Athletes (*n*= 97) (*M* ± *SD*)**	***t*(258)**	***p***	**Cohen's *d***
Neuroticism	-.03 ± .59	.06 ± .60	-1.157	.248	.15
Openness	-.09 ± .61	.15 ± .63	-3.025	.003*	.39
ESCQ - PUE	.00 ± .49	-.01 ± .56	.181	.857	.02
ESCQ - ELE	.04 ± .52	-.07 ± .54	1.748	.082	.22
ESCQ - MRE	.07 ± .41	-.12 ± .43	3.631	<.001*	.47
STEU	.56 ± 5.13	-.94 ± 4.96	2.308	.022	.30
STEM	.66 ± 3.97	-1.11 ± 4.36	3.364	<.001*	.43

Note. * *p* < .005 indicates significance after Bonferroni correction

Regarding the female participants, differences between female athletes and non-athletes are presented in Table 4. Female athletes exhibit significantly higher levels of *extraversion* and greater *emotional competence in managing and regulating emotions* compared to non-athletic women, with these differences falling into the category of medium effect size. It can also be reasonably concluded that female athletes demonstrate significantly higher levels of *conscientiousness*, as the obtained p-value of .006 is at the threshold of the highly conservative Bonferroni correction, and the analysis of Cohen's d values indicates that this difference falls within the category of moderate effect.

Table 4. Comparison of female athletes and non - athletes (residuals controlled for age).

Measures	**Female Athletes (*n*= 114) (*M* ± *SD*)**	**Female Non-Athletes (*n*= 107) (*M* ± *SD*)**	***t*(219)**	***p***	**Cohen's *d***
Extraversion	0.17 ± 0.62	-0.18 ± 0.66	3.961	< .001*	.53
Agreeableness	0.07 ± 0.59	-0.07 ± 0.59	1.769	.078	.24
Conscientiousness	0.11 ± 0.60	-0.11 ± 0.58	2.789	.006	.38
Neuroticism	-0.04 ± 0.58	0.05 ± 0.57	-1.133	.259	.15
Openness	-0.06 ± 0.64	0.06 ± 0.73	-1.343	.181	.18
ESCQ - PUE	0.03 ± 0.42	-0.03 ± 0.57	0.963	.337	.13
ESCQ - ELE	0.06 ± 0.54	-0.06 ± 0.60	1.529	.128	.21
ESCQ - MRE	0.10 ± 0.37	-0.11 ± 0.44	3.748	< .001*	.50
STEU	0.26 ± 5.06	-0.27 ± 4.86	0.794	.428	.11
STEM	0.29 ± 3.97	-0.31 ± 4.37	1.066	.288	.14

Note. * *p* < .005 indicates significance after Bonferroni correction

In order to examine differences between individuals with varying levels of sports activity, specifically between non-athletes and three categories of athletic achievement (low, medium, and high levels of success), a one-way ANOVA was conducted (see Table 5). Based on the conventional threshold of statistical significance (p < .05), significant differences were observed between the four groups for 9 out of the 10 variables examined. However, when applying the Bonferroni correction in this case as well, it was concluded that at the stricter level (p < .005), statistically significant differences between the groups appear only for the traits of *extraversion* and *conscientiousness*, as well as for *emotional competence of managing and regulating emotions*. Effect size analysis indicates that the variable emotional competence of managing and regulating emotions falls into the category of medium effect (Cohen, 2013), explaining 6% of the total variance in differences across levels of sports activity ($\eta^2 = .060$). A slightly lower percentage of variance is explained by extraversion (3.6%) and conscientiousness (2.9%). Due to these significant effects, these three variables were included in post hoc analyses to determine the specifics of these differences, while the remaining variables were excluded from further examination.

Table 5. Comparison across four levels of sports activity (residuals controlled for gender and age).

Measure	*F* (3, 478)	*p*	η^2
Extraversion	6.005	.001*	.036
Agreeableness	2.657	.048	.017
Conscientiousness	4.800	.003*	.029
Neuroticism	3.218	.023	.020
Openness	3.245	.022	.020
ESCQ - PUE	2.971	.031	.018
ESCQ - ELE	3.684	.012	.023
ESCQ - MRE	10.046	<.001*	.060
STEU	2.021	.110	.013
STEM	3.712	.012	.023

Note. *F* = test for variance between groups and within groups; η^2 = Effect size; * $p < .005$ indicates significance after Bonferroni correction.

The Bonferroni post hoc analyses, presented in Table 6, show that non-athletes display statistically significantly lower values on the *extraversion* scale compared to athletes with low and medium levels of achievement, with these differences falling into the category of moderate effect size. However, the differences between non-athletes and highly successful athletes, as well as the differences among all three groups of athletic success in terms of extraversion, were not found to be statistically significant.

For the *conscientiousness* scale, it was also found that non-athletes report significantly lower values, in this case compared to athletes with low and high levels of success, with these differences also classified as moderate in effect size. Differences between non-athletes and moderately successful athletes, as well as among the three groups of athletic achievement, were not confirmed as statistically significant.

When it comes to *emotional competence of managing and regulating emotions*, the only EI scale (across both self-report and ability-based measures) for which significant differences were confirmed between groups after applying the Bonferroni correction, the post hoc analysis indicates that non-athletes report significantly lower values via self-assessment compared to athletes of all three levels of success. These differences fall into the category of moderate to medium effect sizes. No significant differences were found among athletes of different levels of success.

Table 6. Post hoc comparisons across four levels of sports activity for extraversion, conscientiousness, and emotional competence in managing and regulating emotions (residuals controlled for gender and age).

Measures	Sports Level		*MD*	*SE*	*t*	*pbonf*	Cohen's *d*
		Low level	-.422	.127	-3.310	.006*	.43
Extraversion	Non-athletes	Medium level	-.429	.121	-3.541	.003*	.44
		High level	-.213	.122	-1.751	.483	.22

Measures	Sports Level		*MD*	*SE*	*t*	*pbonf*	Cohen's *d*
Conscientiousness	Low level	Medium level	-.008	.146	-0.052	1.000	.01
		High level	.209	.147	1.421	.935	.21
	Medium level	High level	.216	.141	1.528	.763	.22
	Non-athletes	Low level	-.380	.128	-2.973	.019*	.39
		Medium level	-.252	.122	-2.071	.234	.26
		High level	-.374	.122	-3.061	.014*	.38
	Low level	Medium level	.128	.147	.873	1.000	.13
		High level	.006	.147	.044	1.000	.01
	Medium level	High level	-.122	.142	-.858	1.000	.12
ESCQ - MRE	Non-athletes	Low level	-.619	.126	-4.920	< .001*	.64
		Medium level	-.371	.120	-3.100	.012*	.38
		High level	-.425	.120	-3.536	.003*	.44
	Low level	Medium level	.248	.145	1.714	.523	.26
		High level	.194	.145	1.340	1.000	.20
	Medium level	High level	-.054	.140	-.384	1.000	.06

Note. MD= Mean Difference, SE= Standard Error, pbonf= Bonferroni corrected p-value; * $p < .05$ indicates significance after Bonferroni correction

Discussion

The renewed focus on personality traits in sport, along with a moderate number of studies addressing EI in the sports context, points to the relevance and potential of these constructs for understanding the diverse dynamics of athletic activity. At the same time, it highlights the need for further investigation of these phenomena. With this in mind, we conducted a study aimed at examining whether athletes and non-athletes differ in terms of personality traits and EI (operationalized through self-report questionnaires and ability-based tests), and whether specific distinctions exist within gender groups (male athletes vs. male non-athletes; female athletes vs. female non-athletes), as well as across different levels of sport participation and achievement. The final conclusions suggest that personality traits and EI may serve as useful predictors for understanding engagement in sport, with *emotional competence in managing and regulating emotions* standing out as particularly important.

Comparing the personality traits of athletes and non-athletes, this study confirms the findings of previous research that athletes are primarily more *extroverted* than non-athletes (see Allen et al., 2013; Paunonen, 2003), but also more *conscientious*, which is also in line with some earlier works (Steca et al., 2018). A tendency toward intensive social interactions, organization, discipline, and goal-oriented behavior are characteristics that certainly find fertile ground in the realm of sports activities. However, we cannot prejudge the causal direction of this relationship (whether extroverted and conscientious individuals enter sports, or if sports shape them, or whether it is both), considering that our research design is cross-sectional. In addition to these two traits, a significant difference between athletes and non-athletes in our sample was confirmed only for *openness to experience*. Allen et al. (2013) cite several studies emphasizing athletes' greater inclination toward openness to new ideas and experiences, creativity, and curiosity, whereas Steca et al. (2018) report no significant differences for their sample, pointing to the consistency of these findings with the majority of existing research. What is specific to our study is that non-athletes exhibited significantly higher *openness to experience* than athletes. At first glance, sports are characterized by constant unpredictable and challenging situations, reflected in changing technical-tactical tasks, as well as variable conditions (opponents, surfaces, crowds, etc.), which we assume attracts and develops openness to experience. However, we hypothesize that the organization of sports activities, especially in certain cultures with pronounced autocratic leadership styles, can easily take the form of a highly structured, routinized, rigid system with clear rules, which as such encourages conformity more than a tendency to explore, experiment, and be open to new ideas. It is important to note that, although statistically significant, the differences obtained for the three mentioned traits are of moderate effect, which implies that they can serve as predictors of sports engagement, but they are not key factors, and there are other factors that shape sports participation.

We found it interesting to examine the differences between athletes and non-athletes separately for men and women, and it was determined that the mentioned higher degree of *openness to new experiences* in non-athletes compared to athletes for the entire sample is particularly pronounced in the male population. This raises the question for future research to focus on better understanding this phenomenon, especially from the perspective of the hypothesis that in male sports in Bosnia and Herzegovina, as well as in the Balkans in general, a highly structured and rigid work model predominates, which "stifles" curiosity, creativity, and inquisitiveness. It is also important to highlight the existence of moderate effect differences on the *conscientiousness* scale. Athletes show higher values, although this difference is significant only at the level of conventional values ($p < .05$), which is consistent with some earlier studies (Malinauskas et al., 2014; Steca et al., 2018). This is also the case when comparing female athletes and female non-athletes. Regarding other traits in the female population, a statistically significant difference with a medium effect was confirmed only for the *extraversion* scale, where female athletes show a greater tendency toward this characteristic. Although there is a serious limitation in researching differences between female athletes and non-athletes concerning the Big Five personality traits, our results indicate that there are no specificities for the female population, as was the case for the male population. In other words, the greater tendency toward *extraversion* and *conscientiousness* in female athletes is complementary to existing knowledge about the general differences between athletes and non-athletes of both sexes.

We also focused on the question of personality trait differences between athletes of higher and lower levels of success and non-athletes, considering that the analysis of existing research (see the introduction) concludes that more successful athletes are characterized by more pronounced adaptive personality traits (extraversion, conscientiousness, agreeableness, openness to new experiences) and are less prone to maladaptive traits (neuroticism). By categorizing our sample into 4 groups (non-athletes and 3 levels of sports success), significant differences were found between the groups for *extraversion* and *conscientiousness* traits (it is important to note that significant differences were also found for the other three traits, but according to conventional values of $p < .05$, which were of moderate effect). Further analysis showed that the significance of the differences on the *extraversion* and *conscientiousness* scales, which had a moderate effect, stemmed exclusively from the differences between non-athletes on one side and certain groups of athletes with varying levels of success on the other. However, in both cases, no differences were found between athletes of different levels of success. What draws attention is the fact that a higher degree of *extraversion* in athletes is characteristic only for athletes of lower and middle levels of success, while those at the highest level are somewhere between these two groups of athletes and non-athletes and do not significantly deviate from them. This nonlinear nature of the relationship between *extraversion* and *success in sports* may be explained by the different nature of the demands required at different levels of sports success. We assume that lower sports levels are characterized by a need for sociability and openness to the outside world, while at the highest levels, other factors such as discipline, focus, mental stability, and even egoism (self-centeredness) may play a role in enabling better functioning.

The different nature of the differences in *conscientiousness* partly justifies this hypothesis. On the *conscientiousness* scale, differences between the groups stem exclusively from the differences between non-athletes on one side and athletes at lower and higher levels on the other. However, within these three groups of athletes, no significant differences were found. One possible explanation is that entering sports and reaching the regional competitive level requires a certain degree of conscientiousness, which is not a priority for advancing to the higher, national level of competition, where talent may play a more significant role.

However, to reach the international level, talent alone is no longer enough, and along with it, a high level of discipline, persistence, and focus on improving one's abilities is required. In addition to these assumptions about the nature of the differences, we are certainly aware of the problem and the potential effect of applying the method of categorizing athletes for different sports. While the majority of studies, due to the complexity of operationalizing sports success for various sports, use homogeneous samples (athletes from the same sport), we opted for a heterogeneous sample (athletes from different sports) in order to obtain more generalizable conclusions, and we developed a complex classification protocol, which presents certain methodological challenges and potentially questions the criterion validity of the given method. Future research should address this issue to enable the formulation of general conclusions about the relationship between personality traits and sports success, rather than obtaining data that are only applicable to one sport.

When it comes to EI, although athletes show significantly higher scores (according to the conventional value of $p < .05$) both on the self-assessment scale of their emotional competence of expressing and labeling emotions and on the situational test of emotional understanding, after the Bonferroni correction, a significant difference of moderate effect was confirmed exclusively for the highest hierarchical branch – the *ability to manage emotions*, both in the self-assessment questionnaire and in the test. In other words, athletes not only rate their ability to consciously regulate their own and others' emotions to achieve desired behaviors higher than non-athletes, but the tests also confirm that their ability to resolve emotionally complex situations is at a higher level. We assume that hierarchically lower and less complex emotional competencies, such as understanding and labeling emotions, do not require complex life situations to develop, and are developed through everyday life activities. On the other hand, sports activities are filled with emotionally intense and complex situations, thus forcing athletes to develop a higher degree of emotional competence – *emotion management*, which is a good prerequisite for successful performance. Previous research generally confirms a higher degree of EI in athletes compared to non-athletes (see Ubago-Jiménez et al., 2019). However, these studies use different instruments and often do not specify the results of individual EI components. This obliges us to focus on the standardization of the methodological approach to researching EI, specifically in the context of sports activity, in order to enable the comparison of results and synthesis of findings, i.e., drawing conclusions about the relationship between individual branches of EI and sports activity.

By separating the sample according to gender and examining the differences between male athletes and non-athletes, similar results to the overall sample were found. Male athletes rated their competencies in *regulating and managing emotions* at a significantly higher level than male non-athletes, which was also confirmed by significant differences on the situational emotion management test, with differences approximately at the level of moderate effect. Additionally, male athletes scored higher on the situational test of *emotional understanding*, but this difference was of moderate effect and significant only at the conventional significance level. When these differences were examined in the female sample, the only significant difference, of moderate effect, was confirmed for the *emotional competence of managing and regulating emotions*, where female athletes showed higher values compared to female non-athletes. Interestingly, no significant differences were found on the other self-assessment scales, especially on the EI tests, even when the conventional significance level was applied. This somewhat aligns with the results of the study by Rodriguez-Romo and colleagues (2021), one of the few on this topic. The authors found that men at higher sports levels scored higher on the emotional repair scale (which complements the emotional management scale) compared to men at lower sports levels and men who engage in sports but do not compete, while a comparison of the results of female athletes and non-athletes did not show significant differences. The explanation for these results may lie in the findings of numerous studies indicating that women in the general population tend to exhibit higher EI than men (Extremera et al., 2006; Kafetsios, 2004). We assume that, regardless of whether the causes are biological or social, women, through growing up in "typical" life conditions, achieve a higher level of EI (Thompson & Voyer, 2014), and according to the "ceiling effect" principle, their ability to further improve their EI through sports is more limited compared to men, who start at a lower level. In other words, sport emerges as a useful tool for a more significant improvement in EI among men, particularly in its most important aspect, which is *emotion management*.

The final question regarding EI concerns understanding the differences between respondents in the four categories of sports activity (non-athletes and 3 levels of athletic success), where it was found that, according to the conventional value ($p < .05$), there are significant differences for all EI scales except for situational emotional understanding. According to the Bonferroni correction, a statistically significant difference was found only for *emotional competence in managing and regulating emotions*, and further analysis indicated that non-athletes perceive their ability to manage and regulate emotions as significantly weaker than athletes in all three success categories. In other words, even the lowest level of athletic success is associated with a higher degree of *emotional management competence* (or at least a subjective assessment of it) compared to individuals who are not involved in sports. What is interesting in our research, and inconsistent with many previous studies (Kopp & Jekauc, 2018; Laborde et al., 2016; Merino Fernández et al., 2022; Montenegro-Bonilla et al., 2024), is that athletes across different levels of athletic achievement do not show significant differences on this scale. However, it is important to note that there are also studies that align with our findings (Alhabsyi et al., 2022). We are aware that the previously mentioned issue of applying the method of categorizing athletic success levels in a heterogeneous sample of athletes may be the cause of the lack of statistically significant differences. However, it should be considered that this may not be a methodological limitation, but rather the results are due to the specificity of the sample, and that the reference studies were conducted in different socio-cultural contexts.

Conclusion

The increasing emphasis on the importance of physical activity for contemporary man leads to the re-evaluation of the relationship between this activity and the psychological characteristics of the person, highlighting the need for more comprehensive scientific approaches. Aiming for this, the results obtained point to the significance of personality traits and emotional intelligence in understanding sports activity, or inactivity. Specifically, in our case, participants in sports activities are more strongly characterized by highly desirable traits such as *conscientiousness* and *extraversion* compared to non-athletes, which is consistent with the results of previous studies. What is not consistent with previous research is that, in our case, non-athletes exhibit a greater degree of *openness to new experiences*. This is particularly pronounced in the male population, which suggests that future research should seriously consider socio-cultural parameters (e.g., leadership styles, social values, organizational systems, tradition, etc.). An additional argument for the specificity of this sample is the absence of personality trait differences between athletes of different levels of sports success.

Athletes also show a higher level of EI, primarily in the highest hierarchical branch, *emotion regulation ability*, which was confirmed by both self-assessment questionnaires and tests. Dividing the sample by gender suggests that sports activity may have more pronounced effects on the male population than on females. While women involved in sports assess their ability to regulate emotions higher than non-athletes, this is not confirmed by the tests, whereas the differences in men are confirmed by both instruments. Moreover, there are additional arguments indicating a tendency for male athletes to demonstrate greater situational *emotion understanding* than non-athletic men. Interestingly, as with personality traits, no significant differences in EI were found between athletes of different levels of success.

The multidimensional research framework, the dual method of EI operationalization, and the culturally specific, relatively large, and sport-type diverse sample of participants represent key methodological strengths of this study, justifying the serious consideration and interpretation of the findings. However, it should be noted that the study has methodological limitations. Given that this is a cross-sectional study, causal conclusions cannot be drawn. The study primarily relies on self-assessment instruments that are not specifically adapted to the sports context, which may introduce bias affecting the validity of the findings. Generating a universal method for classifying levels of sports success is methodologically challenging and may undermine the consistency and reliability of the obtained categorical distinctions. Furthermore, the sample is limited to the territory of Bosnia and Herzegovina, which restricts the generalizability of the findings. Considering these limitations, it is recommended that future research employ a longitudinal design to clarify causal mechanisms, and utilize a culturally broader sample to ensure the generalizability of the conclusions. Further refinement of findings regarding performance levels requires dedicated investigation of these relationships separately across different sports. If future research manages to integrate all these elements, it could lead to more significant and deeper insights into this complex area. This would result in high-quality practical guidelines for experts in contemporary sports and strengthen the interest of the scientific community in deepening knowledge on the essential psychological aspects of sports

Acknowledgements: The authors acknowledge no conflicts of interest.

References

Austin, E. J. (2010). Measurement of ability emotional intelligence: Results for two new tests. British Journal of Psychology, 101(3), 563–578. https://doi.org/10.1348/000712609X474370

Alhabsyi, R., Rahim, R., PA, W., Mohamed Shapie, M. N., Meera, N., & Parnabas, V. (2022). Assessing the big five personality on categories of athletes and gender among tennis players. Journal of Physical Education and Sport, 22, 2413–2419. https://doi.org/10.7752/jpes.2022.10308

Allen, M. S., Greenlees, I., & Jones, M. (2011). An investigation of the five-factor model of personality and coping behaviour in sport. Journal of Sports Sciences, 29(8), 841–850. https://doi.org/10.1080/02640414.2011.565064

Allen, M. S., Greenlees, I., & Jones, M. (2013). Personality in sport: A comprehensive review. International Review of Sport and Exercise Psychology, 6(1), 184–208. https://doi.org/10.1080/1750984X.2013.769614

Bostani, M., & Saiiari, A. (2011). Comparison Emotional Intelligence and Mental Health between Athletic and Non-Athletic Students. Procedia - Social and Behavioral Sciences, 30, 2259–2263. https://doi.org/10.1016/j.sbspro.2011.10.441

Burdzicka-Wołowik, J., & Góral, K. (2014). Selected Personality Traits of Women Training Combat Sports. Polish Journal of Sport and Tourism, 21. https://doi.org/10.2478/pjst-2014-0001

Cecchini, J., Méndez-Giménez, A., & fernández losa, J. (2019). Self-regulation of learning as a mediator between emotional intelligence and moderate- to-vigorous physical activity. International Journal of Sport Psychology, 50, 391–410. https://doi.org/10.7352/IJSP.2019.50.391

Cohen, J. (2013). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Extremera, N., Fernández-Berrocal, P., & Salovey, P. (2006). Spanish version of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). Version 2.0: Reliabilities, age and gender differences. Psicothema, 18 Suppl, 42–48.

Fabbricatore, R., Iannario, M., Romano, R., & Vistocco, D. (2023). Component-based structural equation modeling for the assessment of psycho-social aspects and performance of athletes. AStA Advances in Statistical Analysis, 107(1), 343–367. https://doi.org/10.1007/s10182-021-00417-5

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Gatsis, G., Strigas, A., & Ntasis, L. (2021). Contribution of emotional intelligence to Taekwondo athletes' performance. Journal of Physical Education and Sport, 21(3), 1976–1980. https://doi.org/10.7752/jpes.2021.s3251

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. Journal of Personality and Social Psychology, 59(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Haryanto, J., Malagoli Lanzoni, I., Nikolakakis, A., Drenowatz, C., Edmizal, E., Apriyano, B., Padli, Padli, Milovanovic, M., Lukacova, T., & Becerra-Patino, B. (2024). Exploring cognitive processing speed, emotional intelligence, and topspin shot accuracy in table tennis. Journal of Physical Education and Sport, 24(3), Article 3082. https://doi.org/10.7752/jpes.2024.03082

IBM Corp. (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

John, O. P., Donahue, E. M., & Kentle, L. R. (1991). The Big Five Inventory—Versions 4a and 54. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (2nd ed., pp. 102–138). New York: Guilford Press.

Kafetsios, K. (2004). Attachment and emotional intelligence abilities across the life course. Personality and Individual Differences, 37, 129–145.

Kaiseler, M., Polman, R. C. J., & Nicholls, A. R. (2012). Effects of the Big Five personality dimensions on appraisal coping, and coping effectiveness in sport. European Journal of Sport Science, 12(1), 62–72. https://doi.org/10.1080/17461391.2010.551410

Kassin, S. M. (2003). Psychology. Pearson/Prentice Hall.

Kim, H.-Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. Restorative Dentistry & Endodontics, 38(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Klatt, S., Rückel, L.-M., Wagener, S., & Noël, B. (2021). Personality Traits and Emotion Regulation Styles of Elite Beach Volleyball Dyads: Examination of Intra-Team Differences, Performance and Satisfaction Levels. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.719572

Kopp, A., & Jekauc, D. (2018). The Influence of Emotional Intelligence on Performance in Competitive Sports: A Meta-Analytical Investigation. Sports, 6(4), 175. https://doi.org/10.3390/sports6040175

Laborde, S., Dosseville, F., & Allen, M. S. (2016). Emotional intelligence in sport and exercise: A systematic review. Scandinavian Journal of Medicine & Science in Sports, 26(8), 862–874. https://doi.org/10.1111/sms.12510

Lepir, D., Lakić, S., & Takšić, V. (2018). Relations between sport participation and emotional intelligence at high school age. Primenjena psihologija, 11(3), 285–300. https://doi.org/10.19090/pp.2018.3.285-300

Li, C., Kee, Y. H., Wang, X., & Guo, Q. (2018). The Big Five personality traits and physical and emotional exhaustion among athletes: The mediating role of autonomous and controlled motivation. International Journal of Sport Psychology, 49, 1–16. https://doi.org/10.7352/IJSP2018.49.001

MacCann, C., Pearce, N., & D. Roberts, R. (2011). Emotional Intelligence as Assessed by Situational Judgment and Emotion Recognition Tests: Building the Nomological Net. Psihologijske Teme, 20(3), 393–412.

MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. Emotion, 8(4), 540–551. https://doi.org/10.1037/a0012746

MacCann, C., Roberts, R. D., Matthews, G., & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based Emotional Intelligence (EI) tests. Personality and Individual Differences, 36(3), 645–662. https://doi.org/10.1016/S0191-8869(03)00123-5

Malinauskas, R., Dumciene, A., Mamkus, G., & Venckunas, T. (2014). Personality traits and exercise capacity in male athletes and non-athletes. Perceptual and Motor Skills, 118(1), 145–161. https://doi.org/10.2466/29.25.PMS.118k13w1

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. Annual Review of Psychology, 59, 507–536. https://doi.org/10.1146/annurev.psych.59.103006.093646

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), Emotional development and emotional intelligence: Educational implications (pp. 3–34). Basic Books.

Merino Fernández, M., Dal Bello, F., Brabec, L., Brito, C., Miarka, B., & López Díaz de Durana, A. (2019). State-trait anxiety and reduced emotional intelligence in combat sport athletes of different genders and competitive levels. Journal of Physical Education and Sport, 19, 363–368. https://doi.org/10.7752/jpes.2019.s2054

Merino Fernández, M., Vanenzuela Perez, D. I., Aedo-Muñoz, E., Mota Barreto, L. B., José Brito, C., Andrade De Brito, M., Miarka, B., & Díaz De Durana, A. L. (2022). Emotional intelligence contributes to increased ranking in combat athletes with anxiety disorders. Journal of Physical Education and Sport, 22(2), 336–344. https://doi.org/10.7752/jpes.2022.02042

Mon-López, D., Blanco-García, C., Acebes-Sánchez, J., Rodríguez-Romo, G., Marquina, M., Martín-Castellanos, A., de la Rubia, A., Cordente Martínez, C., Oliván Mallén, J., & Garrido-Muñoz, M. (2023). Emotional Intelligence in Spanish Elite Athletes: Is There a Differential Factor between Sports? Sports, 11(8), Article 8. https://doi.org/10.3390/sports11080160

Montenegro-Bonilla, A., Becerra-Patiño, B. A., Pino Ortega, J., Hernandez Beltran, V., & Gamonales, J. M. (2024). Influence of Emotional Intelligence on Sports Performance: A Systematic Review. Cuadernos de Psicología Del Deporte, 24, 34–52. https://doi.org/10.6018/cpd.617181

Pallant, J. (2020). SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS (7th ed.). Routledge. https://doi.org/10.4324/9781003117452

Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. Journal of Personality and Social Psychology, 84(2), 411–424. https://doi.org/10.1037/0022-3514.84.2.411

Piedmont, R. L., Hill, D. C., & Blanco, S. (1999). Predicting athletic performance using the five-factor model of personality. Personality and Individual Differences, 27(4), 769–777. https://doi.org/10.1016/S0191-8869(98)00280-3

Piepiora, P. A., Čaplová, P., Zimoń, P., & Gumienna, R. (2024). On research into the relationship between personality traits and the sporting level of competitive, professional and elite athletes. Frontiers in Psychology, 15. https://doi.org/10.3389/fpsyg.2024.1428107

Popovych, I., Charkina, O., Halian, A., Danko, D., Zaverukha, O., Haponenko, L., & Kryzhanovskyi, O. (2024). Relationship between emotional intelligence and coping strategies in junior athletes' behavior during extreme competitive situations. 24, 1518–1527. https://doi.org/10.7752/jpes.2024.11278

Rodriguez-Romo, G., Blanco-Garcia, C., Diez-Vega, I., & Acebes-Sánchez, J. (2021). Emotional Intelligence of Undergraduate Athletes: The Role of Sports Experience. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.609154

Shuai, Y., Wang, S., Liu, X., Kueh, Y. C., & Kuan, G. (2023). The influence of the five-factor model of personality on performance in competitive sports: A review. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1284378

Siemon, D., & Wessels, J. (2022). Performance Prediction of Basketball Players Using Automated Personality Mining with Twitter Data. Sport, Business and Management: An International Journal, 13. https://doi.org/10.1108/SBM-10-2021-0119

Steca, P., Baretta, D., Greco, A., D'Addario, M., & Monzani, D. (2018). Associations between personality, sports participation and athletic success. A comparison of Big Five in sporting and non-sporting adults. Personality and Individual Differences, 121, 176–183. https://doi.org/10.1016/j.paid.2017.09.040

Takšić, V. (2002). Upitnici emocionalne inteligencije (kompetentnosti). In K. Lacković-Grgin, A. Bautović, V. Ćubela, & Z. Penezić (Eds.), Zbirka psihologijskih skala i upitnika (pp. 27–45). Filozofski fakultet, Zada.

Takšić, V., Mohoric, T., & Duran, M. (2009). Emotional skills and competence questionnaire (ESCQ) as a self—report measure of emotional intelligence [Dataset]. https://doi.org/10.1037/t63691-000

Takšić, V., Mohorić, T., & Munjas, R. (2006). Emocionalna inteligencija: Teorija, operacionalizacija, primjena i povezanost sa pozitivnom psihologijom. Društvena istraživanja: časopis za opća društvena pitanja, 15(4-5 (84-85)), 729–752.

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. Cognition and Emotion, 28(7), 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Ubago-Jiménez, J. L., González-Valero, G., Puertas-Molero, P., & García-Martínez, I. (2019). Development of Emotional Intelligence through Physical Activity and Sport Practice. A Systematic Review. Behavioral Sciences, 9(4), 44. https://doi.org/10.3390/bs9040044

Vallerand, R. J., & Blanchard, C. M. (2000). The study of emotion in sport and exercise: Historical, definitional, and conceptual perspectives. In Y. L. Hanin, (Ed.), Emotions in sport (pp. 3–37). Human Kinetics.

Wang, M., & Erdheim, J. (2007). Does the five-factor model of personality relate to goal orientation? Personality and Individual Differences, 43(6), 1493–1505. https://doi.org/10.1016/j.paid.2007.04.024

SPORTLOGIA

Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 30/07/2025
Accepted: 20/11/2025
Online first: 26/11/2025
UDK: 159.928:37.015.3-057.875

Original article

Multiple Intelligences and Their Relationship to Physical and Skill Achievement among middle School Students

[1,*]Shukry Basem Jamal, [2]Qasim Yehya Zyad

[1]College of Physical Education and Sports Sciences, University of Mosul, Iraq
[2]Department of Sports Activities, Nineveh Governorate Education Directorate, Iraq

*Correspondence: jamal@uomosul.edu.iq

Abstract

Purpose: This study investigated the levels of multiple intelligences, physical fitness, and sport-specific technical skills among elementary school students and examined the relationships between these three constructs to identify key cognitive predictors of motor and athletic development in young learners. Methods: Ninety-two fifth-grade male Iraqi students participated. Physical fitness was assessed via medicine ball throw, vertical jump, 30-meter sprint, Barrow agility run, and sit-and-reach test. Technical skills included basketball free throws and dribbling, handball shooting, and soccer zigzag dribbling. Multiple intelligences were measured using McKenzie's inventory. Results: Results showed below-norm upper and lower limb strength (medicine ball throw: M = 4.5 m; vertical jump: M = 4.62 cm), moderate sprinting (M = 5.11 s), agility (M = 21.8 s), and flexibility (M = 17 cm). Technical skills were moderate (basketball free throws: M = 3; basketball dribbling: M = 11.32 s; handball shooting: M = 17; soccer dribbling: M = 14.2 s). Bodily-kinesthetic intelligence was dominant (M = 3.6). Correlations revealed bodily-kinesthetic intelligence predicted sprinting ($r = .209, p < .05$), basketball dribbling ($r = .221, p < .05$), and handball shooting ($r = .212, p < .05$); visual-spatial intelligence correlated with upper limb strength ($r = .247, p < .05$); logical-mathematical intelligence with handball shooting ($r = .245, p < .05$); and interpersonal intelligence with lower limb strength ($r = .299, p < .01$) and sprinting ($r = .209, p < .05$). Conclusion: Bodily-kinesthetic intelligence consistently predicted both physical fitness and technical skills, while other intelligences showed selective associations.

Keywords: psychomotor performance, physical performance testing, cognitive development, elementary students, educational measurement

Citation: Shukry, B.J. and Qasim, Y. Z. (2025). Multiple intelligences and their relationship to physical and skill achievement among middle school students. *Sportlogia. 21(*1), 24-33. https://doi.org/10.63356/spl.2025.011

Introduction

Physical education (PE) is not only a setting for developing motor skills and physical fitness, but also a learning domain where perception, decision-making, problem solving, and social interaction influence performance. Within this context, multiple intelligences (MI) theory provides a framework that helps explain individual differences in movement learning and sport performance, especially because several intelligence types align directly with perceptual–motor and tactical demands in PE (Mitchell & Kernodle, 2004; Armstrong, 2009). Among these, bodily-kinesthetic, visual-spatial, logical-mathematical, and interpersonal intelligences are particularly relevant to coordination, spatial awareness, strategic reasoning, and team interaction in games and sport tasks (Mitchell & Kernodle, 2004; Demirel, 2000).

Recent studies support MI's potential as a predictor of sport performance. Visual-spatial intelligence has been associated with youth motor coordination and physical activity engagement (Jansen & Heil, 2010; Jansen, Lehmann, & Tafelmeier, 2018), while bodily-kinesthetic intelligence has been strongly linked to motor competence (Aguilar et al., 2021). Logical-mathematical intelligence contributes to anticipation and tactical decision-making in games (Armstrong, 2009), and interpersonal intelligence supports teamwork and cooperation (Mitchell & Kernodle, 2004). These findings suggest that MI profiles may help identify cognitive attributes that facilitate the acquisition of sport-specific skills such as dribbling, passing, shooting, or agility-based maneuvers.

Despite this potential, MI-based differentiation is still rarely applied in PE settings, especially in Arab educational systems where traditional PE instruction focuses heavily on repetitive drills and physical execution without acknowledging cognitive diversity (Al-Khayat & Al-Hayali, 2001). At the same time, elementary school years represent a critical phase for the development of physical fitness components such as strength, agility, flexibility, and speed, which are foundational for health and future sports participation (Bailey et al., 2009). Importantly, psychological and cognitive factors also contribute to motor proficiency and physical activity behaviors (Stodden et al., 2008), yet these dimensions remain under-examined in regional PE practice.

To measure MI profiles in applied school-based research, McKenzie's multiple intelligence inventory (MII) is among the most widely used tools and has been used consistently in academic research on education (McKenzie, 1999; Kutz et al., 2013). This survey's internal consistency, as reported by multiple researchers and utilized in earlier studies, ranges between 0.85 and 0.90 (Kutz et al., 2013; Razmjoo, 2008). The inventory contains 90 items covering nine distinct intelligence domains, with ten items representing each intelligence type.

However, in Iraq, there is still a lack of empirical research linking MI domains with children's sport-specific performance and physical fitness levels. This limits teachers' ability to design instruction that matches children's cognitive strengths and restricts the development of pedagogies that leverage MI theory to enhance motor learning.

Therefore, this study investigates the multiple intelligence profiles of fifth-grade Iraqi students and examines their associations with key physical fitness indicators and sport-specific technical skills. It focuses particularly on whether bodily-kinesthetic, visual-spatial, logical-mathematical, and interpersonal intelligences demonstrate meaningful correlations with performance outcomes. By identifying the most influential intelligence domains, this study seeks to inform PE pedagogy and curriculum development in Iraq, offering evidence-based recommendations to support more effective, individualized, and cognitively informed physical education.

Methods

Research design and Sample

During the 2021–2022 school year, 92 male fifth-graders (mean age = 10–12 years) were selected from Abu Hanifa Elementary School in Nineveh, Iraq. Power factors for correlational analyses were taken into account when determining the sample size. A sample size of 92 offers enough power to identify medium-sized relationships ($r = 0.29$–0.30) between several intelligence characteristics and measures of technical ability or physical fitness, with α set at 0.05 and power ($1 - \beta$) at 0.80. The current sample is sufficient to investigate the intermediate predictive associations that are the focus of this investigation, even though lower correlations ($r = 0.20$–0.25) would be underpowered.

Those who were officially enrolled in the fifth grade, were within the normal age range for this grade (10–12 years), and had received medical clearance to engage in regular physical education classes at school were eligible to participate. Only students who completed the Multiple Intelligences Inventory and the physical and technical competence tests were included, and all participants had informed parental approval and child assent.

Measurements and Procedures

The research sample at Abu Hanifa Intermediate School underwent a battery of tests including several intelligences tests, motor skill performance, and physical fitness between March 4 and May 2, 2022. First, under the guidance of the study team and assigned helpers, physical fitness tests were conducted outdoors between March 4 and April 17, 2022. In order to maintain consistency in the testing environment, motor skill performance tests were carried out in the same outdoor yard from April 18 to May 2, 2022, after this phase. Simultaneously, on April 27, 2022, the Multiple Intelligences Test was given in a specific classroom within the school, which offers a more regulated and calmer indoor environment appropriate for cognitive assessment. These sequential and structured assessments ensured comprehensive data collection across physical, motor, and cognitive domains. The research protocol was recognized in the Declaration of Helsinki (World Medical Association, 2013).

Physical Fitness Tests

Upper-Body Explosive Strength: Participants threw a 3-kg medicine ball as far as they could from their chest while sitting with their back supported. Meters were used to measure the distance. An accurate indicator of upper-limb explosive strength is this test (Stockbrugger & Haennel, 2001).

Lower-Body Explosive Strength: Students began standing and executed a maximum vertical jump test; the height of the jump was recorded. According to Markovic et al. (2004), this test is a common field evaluation of lower-limb power and explosive strength.

Linear running speed: Students performed a 30-m maximal sprint. They started standing and ran 30 meters, with timing recorded to the closest 0.1 seconds. The test is frequently used to assess acceleration and short distance running skills (Little & Williams, 2005).

Barrow Zig-Zag Agility Test: Students had to run in a zigzag pattern past cones as fast as they could, and their time was recorded. The test evaluates agility and multidirectional speed and is a component of the Barrow Motor Ability Test battery (Barrow, 1954).

Sit and Reach Flexibility Test: Students sat with their legs outstretched and reached as far forward as they could along a measuring box. Centimeters were used to mark the best of three efforts. According to Zanevskyy & Zanevska, (2017). this is a standardized test of lower-back and hamstring flexibility.

Technical skills tests

Basketball Free-Throw Accuracy Test (BFAT): Ten free throws from the designated free-throw line were attempted by each participant. The quantity of successful baskets was noted. Basketball shooting accuracy is frequently assessed using this test (Apostolidis & Emmanouil, 2015).

Basketball Dribbling Test with Direction Change (BDTDC): Students had to dribble a basketball through a zigzag pattern of cones, and the time it took them to finish was noted. Ball handling speed and control under directional changes are measured by this test (Apostolidis & Emmanouil, 2015).

Handball Shooting Accuracy Test (HSAT): Each student attempted ten goal shots from a predetermined distance; the number of successful attempts was recorded. This is a methodical evaluation of handball shot accuracy (Hoff & Almasbakk, 1995).

Soccer Zig-Zag Dribbling Test (SZDT): the best time of two attempts was recorded when dents dribbled a ball through five posts in a zigzag pattern as part of the football zig-zag dribbling test. This exam assesses football players' dribbling agility and ball control (Reilly et al., 2000).

Multiple Intelligences Test

MII was used since it has been used consistently in academic research on education (McKenzie, 1999; Kutz et al., 2013). This survey's internal consistency, as reported by multiple researchers and utilized in earlier studies, was 0.85 to 0.90 (Kutz et al., 2013; Razmjoo, 2008). There are nine different categories of intelligence represented by the 90 items in this test, with ten items for each type. Every sentence is written positively, and the contents are presented in a random order. Five response options are available to respondents: "Applies to me completely," "Applies to me a lot," "Applies to me sometimes," "Applies to me a little," and "Does not apply to me completely." These responses are scored on a five-point Likert scale ranging from 5 to 1, respectively. Each type of intelligence is scored independently, allowing for separate analysis of each intelligence dimension. The nine distinct types of intelligence are shown in Table 1.

Table1: Subscales of the Multiple Intelligence Inventory (MII) and Their Descriptions.

Subscale (Intelligence Type)	Description / What It Includes
1. Verbal-Linguistic Intelligence	Enjoys reading, writing, storytelling, and verbal games; good at speaking, debating, and remembering information.
2. Logical-Mathematical Intelligence	Solves puzzles and problems, works well with numbers, likes experiments and logical patterns, enjoys strategy games.
3. Visual-Spatial Intelligence	Thinks in images and pictures, enjoys drawing and design, good at puzzles and maps, visualizing objects in space.
4. Bodily-Kinesthetic Intelligence	Uses the body effectively; enjoys physical activity, sports, role-play, crafts, and learning through movement and touch.
5. Musical-Rhythmic Intelligence	Sensitive to sound, rhythm, and melody; enjoys singing, playing instruments, composing, and recognizing musical patterns.
6. Interpersonal Intelligence	Works well with others, good at communication and empathy, enjoys group work and social interaction.
7. Intrapersonal Intelligence	Self-aware and reflective; prefers working alone, sets goals, understands personal strengths and weaknesses.
8. Naturalistic Intelligence	Interested in nature, animals, plants, and the environment; classifies natural objects and enjoys outdoor activities.
9. Existential Intelligence	Thinks deeply about life, existence, and purpose; interested in philosophical questions and abstract thinking.

Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics software (version 20.0). Prior to conducting the main analyses, the normality of data distribution was assessed using the Shapiro–Wilk test. Descriptive statistics (means and standard deviations) were calculated for the three examined constructs: physical fitness, sport-specific technical skills, and multiple intelligences. Associations among these constructs were evaluated using Pearson's product–moment correlation analysis. Effect sizes for the correlation coefficients were interpreted according to Cohen's criteria, where values of approximately $r = 0.20$ indicate a small effect, $r = 0.50$ a medium effect, and $r \geq 0.80$ a large effect (Sullivan and Feinn, 2008). The level of statistical significance was set at $\alpha = 0.05$.

Results

Starting with the Physical Fitness tests, participants showed moderate performance in fitness components (Table 2). Mean scores were 4.5 m for upper limb strength, 4.62 cm for lower limb strength, 5.11 sec for sprinting, 21.8 sec for agility, and 17 cm for flexibility. Overall, the values reflect variability in motor capacities, with agility and flexibility showing relatively consistent results compared to explosive strength measures.

Table 2. Results of the Physical Fitness tests.

#	Physical Component	Test	Mean	SD (±)
1	Upper Limb Explosive Strength (m)	Medicine Ball Throw	4.5	3.21
2	Lower Limb Explosive Strength (cm)	Vertical Jump Test	4.62	2.92
3	Sprinting Speed (Sec.)	30-Meter Speed Test	5.11	2.84
4	Agility (Sec.)	Barrow Zig-Zag Agility Test	21.8	1.84
5	Flexibility (cm)	Sit and Reach Test	17	1.04

The participants' scores on sport-specific technical skill tests summarized in Table 3 indicated a moderate level of technical proficiency, with handball shooting showing the strongest relative performance and basketball free throws representing the most challenging skill.

Results of the Multiple Intelligences test were displayed in Table 4. These results showed that bodily-kinesthetic intelligence recorded the highest mean score (M = 3.6, SD = 1.62; total = 331.2), followed by existential (M = 3.1, SD = 2.11; total = 285.2) and naturalistic intelligence (M = 2.9, SD = 2.88; total = 266.8). These findings highlight bodily-kinesthetic intelligence as the most dominant domain among participants, aligning with their engagement in physical activities.

Table 3. Results of the sport-related technical skills tests.

#	Technical Skill	Test	Mean ($\bar{x}$)	SD (±)
1	Basketball Free Throw Shooting	Basketball Free-Throw Accuracy Test	3	1.02
2	Basketball Dribbling	Basketball Dribbling Test with Direction Change	11.32	1.83
3	Handball Shooting (#)	Handball Shooting Accuracy Test	17	1.60
4	Soccer Dribbling (Sec.)	Soccer Zig-Zag Dribbling Test	14.2	1.58

As for the relationships between Physical Fitness tests and Multiple Intelligences presented in Table 5, significant associations were found between visual-spatial intelligence and upper limb strength ($r = .247$, $p < .05$), bodily-kinesthetic intelligence and sprinting speed ($r = .209$, $p < .05$), and interpersonal intelligence with both lower limb strength ($r = .299$, $p < .01$) and sprinting speed ($r = .209$, $p < .05$).

Table 4. Results of the Multiple Intelligences test (MIT).

#	Multiple Intelligences Tests	Mean ($\bar{x}$)	SD (±)
1	Verbal-Linguistic Intelligence	2.7	2.62
2	Logical-Mathematical Intelligence	2.5	2.32
3	Visual-Spatial Intelligence	2.8	2.93
4	Bodily-Kinesthetic Intelligence	3.6	1.62
5	Musical-Rhythmic Intelligence	1.9	2.53
6	Interpersonal Intelligence	2.1	1.93
7	Intrapersonal Intelligence	2.8	1.67
8	Naturalistic Intelligence	2.9	2.88
9	Existential Intelligence	3.1	2.11

Table 5. Relationships between Physical Fitness Components and Multiple Intelligences.

#	Multiple Intelligences Tests	Upper Limb Explosive Strength	Lower Limb Explosive Strength	Sprinting Speed	Agility	Flexibility
1	Verbal-Linguistic Intelligence	-0.010	0.144	0.140	-0.015	0.164
2	Logical-Mathematical Intelligence	-0.017	0.024	0.060	-0.011	0.082
3	Visual-Spatial Intelligence	**.247***	-0.119	0.173	0.097	-0.121
4	Bodily-Kinesthetic Intelligence	0.116	-0.111	**.209***	-0.159	-0.048
5	Musical-Rhythmic Intelligence	0.083	0.191	-0.007	-0.043	-0.085
6	Interpersonal Intelligence	-0.088	**.299****	**.209***	-0.097	0.096
7	Intrapersonal Intelligence	0.109	-0.111	0.071	-0.090	0.143
8	Naturalistic Intelligence	0.107	-0.082	-0.078	0.093	-0.190
9	Existential Intelligence	-0.005	-0.096	-0.026	0.074	0.097

Note: *Correlation is significant at the 0.05 level; **Correlation is significant at the 0.01 level.

Finally, the relationships between technical skills and Multiple Intelligences presented in Table 6 indicated that bodily-kinesthetic intelligence was positively related to basketball dribbling ($r = .221$, $p < .05$) and handball shooting ($r = .212$, $p < .05$), while logical-mathematical intelligence correlated with handball shooting ($r = .245$, $p < .05$).

Table 6. Relationships between technical skills tests and multiple intelligences.

#	Multiple Intelligences Tests	*Free-Throw*	*Basketball Dribbling Test*	*Handball Shooting*	*Soccer Zig-Zag Dribbling*
1	Verbal-Linguistic Intelligence	-0.177	0.048	0.096	-0.045
2	Logical-Mathematical Intelligence	-0.140	-0.052	**.245***	-0.001
3	Visual-Spatial Intelligence	-0.104	-0.104	-0.171	0.027
4	Bodily-Kinesthetic Intelligence	0.094	**.221***	**.212***	-0.020
5	Musical-Rhythmic Intelligence	-0.043	0.075	-0.028	0.010
6	Interpersonal Intelligence	-0.087	0.010	-0.182	-0.053
7	Intrapersonal Intelligence	-0.067	0.049	0.060	0.051
8	Naturalistic Intelligence	0.040	-0.067	-0.049	-0.105
9	Existential Intelligence	0.011	0.090	0.120	0.051

Note: *Correlation is significant at the 0.05 level; **Correlation is significant at the 0.01 level.

Discussion

This study investigated the levels of multiple intelligence profiles, physical fitness, and sport-specific technical skills among elementary school students and examined the relationships between these three constructs to identify key cognitive predictors of motor and athletic development in young learners.

Results from the assessment of upper limb explosive strength showed lower values than the normative range of 6–7.5 meters for this age group (Faigenbaum et al., 2009; González-Devesa et al., 2025), Similarly this, the vertical jump revealed low lower limb explosive strength compared to anticipated 20–30 cm range (Milanese et al., 2010; Castro-Piñero et al., 2010). This suggests that programs targeted at explosiveness are necessary. On the other hand, children between the ages of 10 and 12 had moderate sprinting speed and agility (Ceylan, Saygin, & Irez, 2014; Sporiš et al., 2010). A healthy range of flexibility indicated sufficient joint mobility and a lower risk of injury (Nourbakhsh & Arab, 2002). These findings demonstrated the necessity of giving lower limb power development top priority in physical education programs.

Pupils who demonstrated a moderate level of proficiency in technical skills unique to their sport indicated that they could improve in particular areas. In line with similar earlier studies (Gallahue & Ozmun, 2006; Malina et al., 2004), basketball Free-throw performance averaged three successful attempts out of ten ($M = 3$, $SD = 1.02$). Basketball dribbling with direction change averaged 11.32 s, and soccer zigzag dribbling averaged 14.2 s, both of which were within normative bounds (Williams & Lacy, 2018; Sporiš et al., 2010). The best-performing skill was handball shooting, with 17 out of 30 shots, being successful, suggesting either excellent motor coordination or previous experience (Al-Khayat & Al-Hayali, 2001).

The Multiple Intelligence Inventory revealed that the preponderance of bodily-kinesthetic intelligence ($M = 3.6$) was indicative of the sport-related technical skills and vigorous physical exercise (Kutz et al., 2013; Shearer, 2004). Naturalistic ($M = 2.9$) and existential ($M = 3.1$) intelligences were also comparatively high, perhaps as a result of environmental and cultural influences (Armstrong, 2009; Gardner, 1999). Interpersonal ($M = 2.1$) and musical-rhythmic ($M = 1.9$) intelligences were lowest, presumably due to little exposure to or focus on these areas, but visual-spatial ($M = 2.8$) and intrapersonal ($M = 2.8$) intelligences were intermediate (Ekici, 2011; Rahayu et al., 2024). These results provide credence to Gardner's (1983, 1999) theory, suggesting that to promote inclusive and successful learning, education could encourage weaker domains while utilizing dominant intelligences.

Selective links between cognitive profiles and performance were demonstrated via correlation analysis. Upper limb explosive power and visual-spatial intelligence had a positive correlation ($r = .247$, $p < .05$), which could be attributed to the fact that visual abilities include visual sharpness, depth perception, contrast sensitivity, and visuomotor reaction speed (Ramyarangsi et al., 2024). These abilities enable athletes to effectively observe and respond to visual stimuli during sports activities. Its prominent role in both motor and skill proficiency was supported by the correlations between bodily-kinesthetic intelligence and sprinting speed ($r = .209$, $p < .05$), basketball dribbling ($r = .221$, $p < .05$), and handball shooting ($r = .212$, $p < .05$) (Shearer, 2004). Lower limb strength ($r = .299$, $p < .01$) and sprinting speed ($r = .209$, $p < .05$) were associated with interpersonal intelligence, suggesting the possible influence of social and motivational involvement on physical performance (Rahayu et al., 2024). Handball shooting was linked to logical-mathematical intelligence ($r = .245$, $p < .05$), which is in line with the importance of reasoning and judgement in strategic activities (Bracero-Malagón et al., 2022). Weak or non-significant relationships were found for other intelligences, such as verbal-linguistic, musical-rhythmic, and naturalistic, highlighting the task-specific influences of cognitive abilities.

While visual-spatial intelligence did not exhibit significant connections, perhaps because of task simplicity or developmental stage, bodily-kinesthetic and logical-mathematical intelligences were the most relevant predictors of technical skills. These findings support Gardner's (1999) claims and earlier research (Shearer, 2004; Ekici, 2011; Kutz et al., 2013; Ruiz Pérez et al., 2014; Bracero-Malagón et al., 2022) by showing that technical skill performance is dependent on both physical aptitude and cognitive involvement.

The study's main research question how multiple intelligences relate to physical fitness and technical skills was addressed: all nine intelligences were present, with bodily-kinesthetic dominance. Significant relationships were observed: visual-spatial, bodily-kinesthetic, and interpersonal intelligences with physical fitness; bodily-kinesthetic and logical-mathematical intelligences with technical skills. Hypotheses were partially supported: bodily-kinesthetic intelligence consistently predicted both domains, while logical-mathematical, interpersonal, and visual-spatial intelligences showed selective associations.

Several limitations need to be noted. The sample was limited to a single governorate, which decreased generalizability, and the cross-sectional methodology restricts causal inference. Socioeconomic and cultural elements were not under control. Furthermore, even while standardized tests are trustworthy, they cannot fully reflect all the subtleties of cognitive or motor performance.

Conclusion

According to the study, elementary-aged students have a wide variety of multiple intelligences, with bodily-kinesthetic intelligence being the most reliable indicator of both technical skill competency related to a sport and physical fitness. The selective contributions of interpersonal, visual-spatial, and logical-mathematical intelligences emphasize the significance of task-specific cognitive demands in motor and skill performance.

These findings have several practical implications. To maximize skill acquisition and motor development, physical education courses should first be customized to children's cognitive profiles, with an emphasis on bodily-kinesthetic capabilities. Second, to promote holistic athletic development, instructors can use strategy-based drills, spatial awareness exercises, and cooperative learning activities to engage and develop students' logical-mathematical, visual-spatial, and interpersonal intelligences. Third, by recognizing students' many intelligences, tailored interventions can be made to help weaker domains supplement the development of technical and physical skills. Lastly, these realizations might promote lifetime physical exercise and consistent participation in sports, which would improve mental and physical health. Future studies ought to use intervention-based or longitudinal designs, involve larger populations, and investigate the long-term effects of intelligence-informed training regimens on technical and physical development.

Acknowledgements: We would like to thank the principal of "Abu Hanifa Elementary School" in Nineveh, Iraq for all support and coordination to conduct this research.

***Declaration of interests*:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare no conflict of interest

***Data availability*:** All collected data are included in the manuscript. Raw data are available upon reasonable request to the corresponding author.

References

Aguilar, J., Garces-Jimenez, A., R-Moreno, M. D., & García, R. (2021). A systematic literature review on the use of artificial intelligence in energy self-management in smart buildings. *Renewable and Sustainable Energy Reviews, 151*, 111530. https://doi.org/10.1016/j.rser.2021.111530

Al-Khafaf, I. A. (2011). *Multiple intelligences: An applied program* (1st ed.). Dar Al-Manahj for Publishing and Distribution.

Al-Khatatbeh, M. A., & Shaalan, M. A. (2016). Multiple intelligences of students in the College of Sports Sciences at Mutah University. *Mutah Journal of Research and Studies, Humanities and Social Sciences Series, 31*(6).

Al-Khayat, M., & Al-Hayali, A. (2001). *Tests in team sports skills*. University of Mosul Press.

Al-Rawahi, N. Y., & Zayed, K. (2018). A comparative study of multiple intelligences according to gender and university attendance of students in the College of Physical Education at some Omani universities. *Al-Rafidain Journal of Sports Sciences, 21*(67). https://doi.org/10.33899/rajsport.1999.162936

Apostolidis, N., & Emmanouil, Z. (2015). The influence of the anthropometric characteristics and handgrip strength on the technical skills of young basketball players. *Journal of Physical Education and Sport*, *15*(2), 330. https://doi.org/10.7752/jpes.2015.02050

Armstrong, T. (2009). *Multiple intelligences in the classroom* (3rd ed.). ASCD.

Bailey, R., Hillman, C. H., Arent, S. M., & Petitpas, A. (2009). Physical activity: An underestimated investment in human capital? *Journal of Physical Activity and Health, 6*(3), 262–276. https://doi.org/10.1123/jpah.10.3.289

Barrow, H. M. (1954). *The Barrow General Motor Ability Test*. W.B. Saunders Company.

Bracero-Malagón, J., Juárez-Ruiz de Mier, R., Reigal, R. E., Caballero-Cerbán, M., Hernández-Mendo, A., & Morales-Sánchez, V. (2022). Logical intelligence and mathematical competence are determined by physical fitness in a sample of school children. *Frontiers in Psychology, 13*, 833844. https://doi.org/10.3389/fpsyg.2022.833844

Castro-Piñero, J., González-Montesinos, J. L., Mora, J., Keating, X. D., Girela-Rejón, M. J., Sjöström, M., & Ruiz, J. R. (2009). Percentile values for muscular strength field tests in children aged 6 to 17 years: influence of weight status. *Journal of strength and conditioning research*, *23*(8), 2295–2310. https://doi.org/10.1519/JSC.0b013e3181b8d5c1

Ceylan, H. I., Saygin, O., & Irez, G. B. (2014). The examining body composition, sprint and coordination characteristics of the children aged 7–12 years. *The Anthropologist, 18*(3), 859–867. https://doi.org/10.1080/09720073.2014.11891617

Demirel, O. (2000). *Teaching principles and methods*. Pegem A Publishing.

Ekici, S. (2011). Multiple intelligence levels of physical education and sports school students. *Educational Research and Reviews, 6*(21), 1018–1026.

Faigenbaum, A. D., et al. (2009). Youth resistance training: Updated position statement paper from the National Strength and Conditioning Association. *Journal of Strength and Conditioning Research, 23*(Suppl 5), S60–S79. https://doi.org/10.1519/JSC.0b013e31819df407

Gallahue, D. L., & Ozmun, J. C. (2006). *Understanding motor development: Infants, children, adolescents, and adults* (7th ed.). McGraw-Hill.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.

González-Devesa, D., Varela, S., Diz-Gómez, J. C., López-Amoedo, D., & Ayán-Pérez, C. (2025). Reliability and validity of the medicine ball throw test in children and adolescents: a systematic review and meta-analysis. *Kinesiology*, *57*(1), 122-135. https://doi.org/10.26582/k.57.1.11

Hoff, J., & Almasbakk, B. (1995). The effects of maximum strength training on throwing velocity and muscle strength in female team-handball players. *Journal of Strength and Conditioning Research, 9*(4), 255–258.

Jansen, P., & Heil, M. (2010). The relation between motor development and mental rotation ability in 5–6 years old children. European Journal of Developmental Science, 4, 66−74. https://doi.org/10.3233/DEV-2010-4105

Jansen, P., Lehmann, J., & Tafelmeier, C. (2018). Motor and visual-spatial cognition development in primary school-aged children in Cameroon and Germany. *The Journal of Genetic Psychology, 179*(1), 30–39. https://doi.org/10.1080/00221325.2017.1415201

Kutz, M., Dyer, S., & Campbell, B. (2013). Multiple intelligence profiles of athletic training students. *Internet Journal of Allied Health Sciences and Practice, 11*(1), 9. https://doi.org/10.46743/1540-580x%2F2013.1431

Kutz, M., Dyer, S., & Campbell, B. (2013). Multiple intelligence profiles of athletic training students. *Internet Journal of Allied Health Sciences and Practice*, *11*(1), 9. https://doi.org/10.46743/1540-580X/2013.1431

Little, T., & Williams, A. G. (2005). Specificity of acceleration, maximum speed, and agility in professional soccer players. *Journal of strength and conditioning research*, *19*(1), 76–78. https://doi.org/10.1519/14253.1

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). *Growth, maturation, and physical activity*. Human Kinetics.

Markovic, G., Dizdar, D., Jukić, I., & Cardinale, M. (2004). Reliability and factorial validity of squat and countermovement jump tests. *Journal of Strength and Conditioning Research, 18*(3), 551–555. https://doi.org/10.1519/00124278-200408000-00028

McKenzie, W. (1999). *Multiple intelligences survey*. http://surfaquarium.com/MI/inventory.htm

Milanese, C., et al. (2010). Anthropometry and motor fitness in children aged 6–12 years. *Journal of Sports Sciences, 29*(2), 125–133. https://doi.org/10.4100/jhse.2010.52.14

Mitchell, M., & Kernodle, M. (2004). Using multiple intelligences to teach tennis. *Journal of Physical Education, Recreation & Dance, 75*(8), 27–32. https://doi.org/10.1080/07303084.2004.10607286

Nourbakhsh, M. R., & Arab, A. M. (2002). Relationship between mechanical factors and incidence of low back pain. *The Journal of orthopedic and sports physical therapy*, *32*(9), 447–460. https://doi.org/10.2519/jospt.2002.32.9.447

Rahayu, P., Kriswanto, E. S., Pambudi, A. F., & Yuliarto, H. (2024). Interpersonal intelligence and emotional intelligence; Their effect on physical education learning outcomes in Grade V primary school students. *Fizjoterapia Polska, 24*(4). https://doi.org/10.56984/8ZG01A8A2C7

Ramyarangsi, P., Bennett, S. J., Siripornpanich, V., Nanbancha, A., Pokaisasawan, A., Noppongsakit, P., et al. (2024). Distinct visual processing patterns in female elite athletes: A comparative study of gymnastics, soccer, and esports using visual P300 event-related potentials. *International Journal of Exercise Science, 17*(5), 1595–1604. https://doi.org/10.70252/INCC1951

Razmjoo, S. A. (2008). On the relationship between multiple intelligences and language proficiency. *The Reading Matrix: An International Online Journal, 8*, 155–174.

Reilly, T., Williams, A. M., Nevill, A., & Franks, A. (2000). A multidisciplinary approach to talent identification in soccer. *Journal of Sports Sciences, 18*(9), 695–702. https://doi.org/10.1080/02640410050120041

Ruiz Pérez, L. M., Palomo Nieto, M., Ramón Otero, I., Ruiz Amengual, A., & Navia Manzano, J. A. (2014). Relationships among multiple intelligences, motor performance and academic achievement in secondary school children. *International Journal of Academic Research, 6*(6), 1–13. https://doi.org/10.7813/2075-4124.2014/6-6/B.10

Shearer, B. (2004). Multiple intelligences theory after 20 years. *Teachers College Record, 106*(1), 2–16. https://doi.org/10.1111/j.1467-9620.2004.00313.x

Sporiš, G., Jukic, I., Milanovic, L., & Vucetic, V. (2010). Reliability and factorial validity of agility tests for soccer players. *Journal of Strength and Conditioning Research, 24*(3), 679–686. https://doi.org/10.1519/jsc.0b013e3181c4d324

Stockbrugger, B. A., & Haennel, R. G. (2001). Validity and reliability of a medicine ball explosive power test. *Journal of Strength and Conditioning Research, 15*(4), 431–438.

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest, 60*(2), 290–306. https://doi.org/10.1080/00336297.2008.10483582

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. Journal of graduate medical education, 4(3), 279-282. https://doi.org/10.4300/jgme-d-12-00156.1

Williams, S. M., & Lacy, A. (2018). *Measurement and evaluation in physical education and exercise science*. Routledge.

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Zanevskyy, I., & Zanevska, L. (2017). Evaluation in the sit-and-reach flexibility test. *Journal of Testing and Evaluation, 45*(5), 346–355. https://doi.org/10.1520/JTE20150298

SPORTLOGIA
Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 17/08/2025
Accepted: 12/11/2025
Online first: 03/12/2025
UDK: 796.01/.09:159.947.5-057.87

Original article

Relationships Between Need-Supportive Teaching Style, Psychological Need Satisfaction, Motivation and Happiness Level in Secondary School Physical Education and Sports Course

[1,*]Öner Burak, [2]Karataş Özgür

[1]University of Munzur, Çemişgezek Vocational School, Tunceli, Turkey
[2]University of Inonu, Faculty of Sports Sciences, Malatya, Turkey
*Correspondence: burakoner@munzur.edu.tr

Abstract

From the perspective of Self-Determination Theory, this study aimed to investigate the relationships between middle school students' perceived need for support from their physical education (PE) teachers and their basic psychological need (BPN) satisfaction, motivation, and enjoyment through a structural equation modeling (SEM) approach. The study employed a cross-sectional design, with 1,266 public middle school students from various provinces in the Eastern Anatolia region of Türkiye participating in the research. Participants completed measurement instruments to assess the study variables. Preliminary analyses involved computing descriptive statistics for the study variables and conducting Pearson correlation analyses to evaluate their interrelations based on participants' self-reported data. SEM was employed to test the hypothesized mediation effects within the structural models. Our findings indicated that, within the context of PE classes, perceived need for support from teachers positively predicted middle school students' BPN satisfaction. In the model, perceived need support directly and positively predicted students' intrinsic and extrinsic motivation, while negatively predicting amotivation. Furthermore, BPN satisfaction partially mediated the relationships between perceived need for support and all three motivational regulations. Lastly, both BPN satisfaction and intrinsic motivation were found to mediate the effect of need support on enjoyment partially. The findings of the present study provide evidence for the assumption that, within the context of PE, teachers should adopt a need-supportive teaching style to create an environment where students initiate and engage in activities autonomously, experience enjoyment and satisfaction, and where amotivation is minimized or eliminated.

Keywords: basic psychological needs, enjoyment, motivation, physical education, self-determination theory, structural equation modeling

Citation: Öner, B. and Karataş, Ö. (2025). Relationships Between Need-Supportive Teaching Style, Psychological Need Satisfaction, Motivation and Happiness Level in Secondary School Physical Education and Sports Course. *Sportlogia. 21(*1), 34-47. https://doi.org/10.63356/spl.2025.012

Introduction

Movement is integral to childhood development, facilitating the advancement of physical, cognitive, and social skills (Cools et al., 2009; Lubans et al., 2010). The periods of childhood and adolescence are critical for establishing a foundation for sports and physical activity (Green, 2002). Within this developmental phase, the physical education (PE) classroom serves as a primary social context where children initially engage in physical activity and exercise. Schools, as fundamental social institutions, are tasked with promoting physical activity among youth through PE classes, thereby providing opportunities for physical engagement (Burgeson et al., 2001; Sallis et al., 1997). Although PE class schedules can enhance daily physical activity, there is a pressing need to develop specific curricula that support sustainability and generalization, alongside rigorous evaluation of these programs (Sallis et al., 1997). Consequently, it is crucial to comprehend the motivational, emotional, and cognitive processes that shape students' perceptions of PE classes (Ntoumanis, 2001; Okely et al., 2001). One factor contributing to this understanding is perceived teacher behavior, as teacher-student interactions are linked to children's school attachment and student motivation. Describing this bidirectional interaction is particularly important during middle school, a period characterized by a gradual decline in student motivation (Stroet et al., 2013).

One of the most suitable frameworks for describing the elements of a motivating teaching style is Self-Determination Theory (SDT) (Deci & Ryan, 1985, 2000). SDT addresses individuals' inherent developmental tendencies and psychological needs, serving as a macro theory that elucidates the intrinsic motivations driving behavior, independent of external influences. It emphasizes the nature of self-motivation and the degree of autonomy in human behavior. SDT delineates various types of motivation to explain the origins of behavioral drivers and their identifiable consequences in learning, performance, personal experience, and well-being. The theory establishes foundational principles regarding the development, maintenance, or diminishment of each type of motivation. Specifically, research based on SDT has explored factors that either enhance or diminish intrinsic motivation, self-regulation, and well-being (R. M. Ryan & Deci, 2000).

PE teachers can exert a profound and enduring influence on students' motivation and enjoyment (Bolter et al., 2018). The instructional styles employed by teachers are linked to student motivation and participation in PE classes (De Meyer et al., 2016). It has been suggested that PE teachers' instructional styles can significantly cultivate a supportive motivational climate, which predicts students' basic psychological needs (BPN) satisfaction, motivation quality, physical activity intentions, and engagement (Leo et al., 2022). It has been reported that students who perceive need-supportive behaviours from their teachers report more positive motivational experiences (Burgueño et al., 2024). In PE classes, need-supportive teaching behaviors have been positively associated with high-quality (autonomous) motivation and numerous positive outcomes, including enjoyment of PE classes, motor skill development, behavioral persistence, well-being, and increased levels of physical activity during leisure time (Ntoumanis, 2005; Standage et al., 2003, 2006). Conversely, need-thwarting teacher behaviors have been linked to negative affect, inadequate coping, low performance, controlled motivation, amotivation, and maladaptive student outcomes (R. M. Ryan & Deci, 2000; Van Den Berghe et al., 2013).

One of SDT's mini-theories, BPN Theory, states that teachers can either support or hinder students' BPN, such as autonomy, competence, and relatedness. According to SDT, it is extremely important for motivating teachers to succeed in supporting students' BPN such as autonomy, competence, and relatedness, which are seen as fundamental requirements for student development (Deci & Ryan, 2000). It is believed that PE teachers' teaching styles can greatly shape a supportive motivational climate that predicts students' BPN satisfaction, motivation quality, physical activity intentions, and engagement (Leo et al., 2022). The need-supportive teaching style is a concept that emphasizes the importance of meeting students' needs for autonomy, competence, and relatedness in the classroom. It is believed that positive learning outcomes will emerge in classrooms where these needs are supported (Stroet et al., 2015). In PE classes, BPN satisfaction and their possible outcomes have been examined in numerous previous studies involving students at different grade levels (Aibar et al., 2021; Burgueño et al., 2024; Cox & Williams, 2008; De Meyer et al., 2014; Franco & Coterón, 2017; Leo et al., 2022; Ntoumanis, 2005; Rutten et al., 2012; Sanchez-Oliva et al., 2014; Standage et al., 2005; Zhang et al., 2011). Research has consistently demonstrated that satisfaction of BPN is correlated with increased self-determined motivation (Cox & Williams, 2008; Ntoumanis, 2005; Rutten et al., 2012; Sanchez-Oliva et al., 2014; Zhang et al., 2011) and serves as a positive predictor of future intentions to engage in physical activity (Leo et al., 2022; Ntoumanis, 2005). Conversely, BPN satisfaction has been shown to have a negative association with amotivation (De Meyer et al., 2014; Leo et al., 2022; Shen et al., 2010; Vasconcellos et al., 2020). Evidence has been obtained that students with high BPN satisfaction enjoy PE classes more and feel more energetic (Franco & Coterón, 2017; Mouratidis et al., 2011; Standage et al., 2005).

In light of this information, it is imperative to comprehend the positive emotional characteristics of students in PE classes to achieve the desired educational outcomes. Cultivating positive attitudes towards PE and physical activity at an early age can contribute to individuals leading physically active lives and mitigating the adverse health consequences of inactivity. In alignment with this social benefit, the present study aims to elucidate the importance of a need-supportive teaching style and the satisfaction of needs in influencing students' motivation and happiness levels towards PE during the early stages, specifically the middle school period. Based on this information, the study aimed to examine the structural relationships between a supportive teaching style perceived by middle school students in PE classes, BPN satisfaction, motivation, and enjoyment.

Methods

Design and Participants

This research was conducted as a cross-sectional study utilizing the correlational survey model, a quantitative research methodology. The study involved 1,433 students from public secondary schools (grades 5 to 8) in the provinces of Malatya, Elazığ, Bingöl, and Tunceli, located in the Eastern Anatolia Region of Turkey, during the 2023–2024 academic year. Of these students, 533 (42.1%) were male and 733 (57.9%) were female. Regarding grade level, 273 (21.6%) were in fifth grade, 386 (30.5%) in sixth grade, 302 (23.9%) in seventh grade, and 305 (24.1%) in eighth grade. Ethical approval for the study was obtained from the Social and Human Sciences Research Ethics Committee at Inonu University.

Measurements and Procedures

Prior to the commencement of the research, parents were informed and provided written consent for their children's participation. After obtaining parental consent, students completed paper-and-pencil questionnaires in their classrooms under the supervision of the research team. Data were collected by the research team prior to regular PE classes in accordance with the principles of voluntariness and confidentiality. After invalid data were excluded, the final sample consisted of 1,266 public secondary school students. The sample group was selected using convenience sampling, a non-random sampling method.

Perceived Need-Supportive Behaviours

The Need-Supportive Teaching Style Scale in Physical Education-NSTSSPE Turkish version was used to measure the perceived need support that middle school students participating in this study received from their teachers. The original form of this scale was developed by Liu and Chung (2017). The adaptation of the scale to the Turkish form for the middle school sample was carried out by Sakallı and Kerkez (2021). The scale is evaluated using a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) based on participants' self-reports. Following the adaptation of the scale for the middle school sample, it was ascertained that the three sub-dimensions and the total of 15 items from the original form were retained. These sub-dimensions include autonomy support (e.g., ‘My teacher offers me many options’), competence support (e.g., 'My teacher shows me how to solve problems on my own'), and involvement support (e.g., 'My teacher spends time with me').

Basic Psychological Need Satisfaction in PE

The Basic Psychological Needs in Physical Education Scale developed by Vlachopoulos, Katartzi, and Kontou (2011) and adapted to the Turkish middle school and high school population by Turkay et al. (2019) was used to measure students' BPN satisfaction. The scale is evaluated using a 7-point Likert scale ranging from 1 (not at all appropriate) to 7 (completely appropriate) based on participants' self-reports. After the adaptation study, the 12-item structure of the scale with 3 sub-dimensions named autonomy (e.g., "I feel like I choose the activities we do"), competence (e.g., "I feel very good at tasks that most children find difficult"), and relatedness (e.g., "I feel like a valued member of a group of close friends") was preserved.

Motivation in PE

The Physical Education Motivation Scale (PEMS) developed by Sulz, Temple, and Gibbons (2016) and adapted for Turkish middle school and high school students by Akbulut and Öncü (2023) was used to determine the motivation levels of students participating in the current study in PE classes. PEMS is evaluated using a 5-point Likert scale ranging from 1 (I do not agree at all) to 5 (I agree completely) based on participants' self-reports. The scale consists of a 3-factor structure with a total of 9 items. The sub-dimensions of the scale consist of intrinsic motivation (e.g., 'I participate in PE classes because they are fun'), extrinsic motivation (e.g., 'I try very hard in PE classes because I want to get good grades'), and amotivation (e.g., 'I waste my time in PE classes').

Enjoyment in PE

In this study, the Happiness Level Scale for Physical Education Course (HSFPEC) developed by Uğraş and Serbes (2019) was used to determine participants' enjoyment of PE classes. The HSFPEC is a single-factor 5-point Likert scale measurement tool consisting of 9 items (e.g., 'I enjoy PE and sports classes'). The scale is calculated based on participants' self-reports, with scores ranging from 1 (strongly disagree) to 5 (strongly agree). The scale items do not contain any negative statements.

Statistical analyses

In the SEM analyses, two structural models were tested. Model 1 examined the relationships from perceived need support to basic psychological needs satisfaction, and subsequently to intrinsic motivation, extrinsic motivation, and amotivation (Figure 1). Model 2 tested the pathways from perceived need support to basic psychological needs satisfaction, then to intrinsic motivation, and finally to enjoyment in PE classes (Figure 2). In the current study, SPSS v25 statistical software was used for preliminary analyses and descriptive statistics, while AMOS v22 statistical software was used to test structural models. Before proceeding to the main analyses, confirmatory factor analyses (CFA) were conducted for all variables included in the study to validate the factor structures of the scales. In the next step, descriptive statistics (means, standard deviations) were calculated. The normality of the data was assessed using Skewness and Kurtosis values. The skewness (-1.169 to 1.145) and kurtosis (-0.943 to 0.579) values indicate that the data are normally distributed (George & Mallery, 2016). Cronbach's alpha coefficients were considered to evaluate the reliability of the measurement tools. Total indirect effects were evaluated using SEM for the two hypothesized models (Figure 1 and 2), with 95% bias-corrected bootstrap confidence intervals (95%CIBC), maximum likelihood estimation, and 5000 bootstrap samples. Model goodness was assessed using the chi-square statistic (x^2/sd) value, comparative fit index (CFI), goodness-of-fit index (GFI), incremental fit index (IFI), root mean squared error of approximation (RMSEA), and standardized root mean squared residuals (SRMR) indices. Fit indices of CFI, IFI, and GFI values close to or greater than 0.90, RMSEA values less than 0.06, and SRMR values less than 0.08 indicate an acceptable fit (Hu & Bentler, 1999). In the structural equation modeling (SEM) analyses, the structural models tested the relationships between perceived need support and students' basic psychological needs satisfaction, motivational regulations, and enjoyment. The level of relationship between variables was determined using Pearson's Correlation Test.

Results

Descriptive statistics, including means, standard deviations, skewness, kurtosis, and internal consistency values (Cronbach's alpha), were calculated for all study variables and are presented in Table 1. The perceived need-supportive teaching style, which includes the sub-dimensions of competence support, autonomy support, and involvement, was measured using a 7-point Likert-type scale. The internal consistency coefficients for the total scale and its sub-dimensions were high (α = 0.85 to 0.94), indicating good reliability. The mean scores for total support and its components ranged from 4.55 to 4.62, suggesting a moderately high perception of need-supportive teaching among students. Skewness values ranged from -0.60 to -0.47 and kurtosis values from -0.94 to -0.75, falling within the acceptable range for normal distribution (±2), as recommended by George and Mallery (2016). The BPN-PE, which includes competence, relatedness, and autonomy satisfaction, also utilized a 7-point response format. The internal consistency coefficients for the total score and sub-dimensions ranged from 0.74 to 0.85. The mean scores for these variables ranged between 4.43 and 5.02, indicating a moderate to high level of psychological need satisfaction. Skewness values (-0.63 to -0.26) and kurtosis values (-0.75 to -0.44) were within acceptable ranges, supporting the assumption of normality.

Motivation toward PE was assessed via PEMS, which includes intrinsic motivation, extrinsic motivation, and amotivation subscales, each measured on a 5-point Likert scale. Cronbach's alpha values ranged from 0.77 to 0.83, confirming adequate reliability. The mean scores across motivation types ranged from 2.01 (for amotivation) to 3.96 (for intrinsic motivation). Skewness and kurtosis values for motivation sub-dimensions were also within acceptable ranges (-1.17 to +1.14 for skewness and -0.57 to 0.29 for kurtosis), indicating that the distributions of the responses were not significantly deviated from normality.

The enjoyment level in PE was assessed via the HSFPEC scale on a 5-point Likert scale. The scale demonstrated excellent internal consistency (α = 0.93). The mean score was 3.81 (SD = 0.79), suggesting a relatively high level of enjoyment. The skewness (-1.08) and kurtosis (0.11) values were within the acceptable range, indicating a normal distribution.

Table 1. Mean scores, standard deviations, reliability coefficients, Skewness and Kurtosis values for all study variables.

Scale	Chronbach (α)	Scale range	Min.	Max.	$\bar{X}$	SD	Skew.	Kurt.
NSTSSPE	0.94	1-7	1.20	6.93	4.62	1.54	-0.558	-0.814
Competence support	0.85	1-7	1.00	7.00	4.69	1.62	-0.600	-0.753
Involvement support	0.88	1-7	1.00	7.00	4.63	1.68	-0.530	-0.854
Autonomy support	0.87	1-7	1.00	7.00	4.55	1.70	-0.471	-0.943
BPN-PE	0.85	1-7	1.25	6.92	4.66	1.25	-0.401	-0.441
Competence satisfaction	0.76	1-7	1.00	7.00	4.55	1.50	-0.261	-0.710
Autonomy satisfaction	0.74	1-7	1.00	7.00	4.43	1.58	-0.389	-0.759
Relatedness satisfaction	0.79	1-7	1.00	7.00	5.02	1.57	-0.637	-0.531
Intrinsic motivation	0.83	1-5	1.00	5.00	3.96	1.13	-1.169	0.422
Extrinsic motivation	0.77	1-5	1.00	5.00	3.71	1.13	-0.780	-0.290
Amotivation	0.81	1-5	1.00	5.00	2.01	1.05	1.145	0.579
HSFPEC	0.93	1-5	1.11	5.00	3.81	1.04	-1.084	0.115

Table 2 displays the mutual correlations of all variables. The correlation coefficients indicate that total need support and its three subdimensions (competence, involvement, autonomy support) are all positively correlated with total need satisfaction (r values range = 0.46–0.51, all p values < 0.01). Total need support showed positive correlations with intrinsic ($r = 0.46$, $p < 0.01$) and extrinsic motivation ($r = 0.44$, $p < 0.01$), and negative correlations with amotivation ($r = -0.45$, $p < 0.01$). Similarly, total need satisfaction showed positive correlations with intrinsic ($r = 0.55$, $p < 0.01$) and extrinsic motivation ($r = 0.48$, $p < 0.01$) and negative correlations with amotivation ($r = -0.44$, $p < 0.01$). Finally, the table shows that the relationship between total need support and enjoyment ($r = 0.53$) and the relationship between total need satisfaction and enjoyment ($r = 0.58$) are positive and significant at the $p < 0.01$ level. The mutual correlation coefficients between all variables shown in the table indicate that advanced mediation analyses can be performed.

Table 2. Pearson correlation matrix between NSTSSPE, BPN-PE, PEMS and HSFPEC subscales.

Scale	1	2	3	4	5	6	7	8	9	10	11
1. NSTSSPE	-										
2. Competence support	0.90**	-									
3. Involvement support	0.92**	0.73**	-								
4. Autonomy support	0.93**	0.77**	0.81**	-							
5. BPN-PE	0.51**	0.46**	0.46**	0.49**	-						
6. Competence satis.	0.40**	0.36**	0.36**	0.38**	0.81**	-					
7. Relatedness satis.	0.30**	0.26**	0.28**	0.28**	0.78**	0.42**	-				
8. Autonomy satis.	0.55**	0.50**	0.49**	0.53**	0.83**	0.56**	0.46**	-			
9. Intrinsic motivation	0.46**	0.41**	0.43**	0.44**	0.55**	0.43**	0.34**	0.55**	-		
10. Extrinsic motivation	0.44**	0.41**	0.38**	0.42**	0.48**	0.41**	0.27**	0.49**	0.54**	-	
11. Amotivation	-.45**	-.40**	-.42**	-.43**	-.44**	-.38**	-.25**	-.45**	-.75**	-.46**	-
12. HSFPEC	0.53**	0.47**	0.49**	0.50**	0.58**	0.47**	0.36**	0.57**	0.78**	0.52**	-,68**

Note: **significant at p<.01

Preliminary Analyses

Before evaluating the proposed structural model, confirmatory factor analyses (CFA) were performed to assess the factorial validity of the scales utilized in the research. These analyses were executed using AMOS v22 with maximum likelihood estimation. The NSTSSPE, which comprises three latent dimensions—competence support, autonomy support, and involvement support—showed a satisfactory model fit: $\chi^2/df = 331.78/84 = 3.95$, GFI = 0.96, CFI = 0.97, IFI = 0.97, RMSEA = 0.052, SRMR = 0.02. The factor loadings for the observed items varied from 0.68 to 0.77 for competence support, from 0.73 to 0.80 for autonomy support, and from 0.73 to 0.85 for involvement support. The BPN-PE was also structured with three latent factors—competence satisfaction, autonomy satisfaction, and relatedness (involvement) satisfaction. The CFA results indicated an acceptable fit: $\chi^2/df = 192.08/50 = 3.84$, GFI = 0.97, CFI = 0.96, IFI = 0.97, RMSEA = 0.05, SRMR = 0.03. Factor loadings for the BPN-PE items ranged from 0.48 to 0.82 across the three subscales. The PEMS was evaluated as a three-factor model representing intrinsic motivation, extrinsic motivation, and amotivation. The CFA findings showed a good model fit: $\chi^2/df = 64.11/24 = 2.67$, GFI = 0.98, CFI = 0.99, IFI = 0.99, RMSEA = 0.03, SRMR = 0.02. The factor loadings ranged from 0.63 to 0.90 for intrinsic motivation, 0.63 to 0.80 for extrinsic motivation, and 0.68 to 0.85 for amotivation. Finally, the unidimensional structure of the HSFPEC was confirmed with acceptable fit indices: $\chi^2/df = 61.61/24 = 2.56$, GFI = 0.95, CFI = 0.98, IFI = 0.98, RMSEA = 0.07, SRMR = 0.02. All item factor loadings for the HSFPEC ranged from 0.66 to 0.90. In conclusion, all measurement models exhibited satisfactory levels of model fit and standardized factor loadings, supporting the construct validity of the scales before their inclusion in the full structural model. .

Structural Equation Modeling (SEM)

Firstly, the aim was to determine the predictive effect of the perceived need support that students receive from their PE teachers on BPN satisfaction and their motivation regulations in class. In the structural model (Figure 1), perceived need support is an exogenous variable, whereas BPN satisfaction, intrinsic motivation, extrinsic motivation, and amotivation are endogenous variables. In this context, the goodness-of-fit values for the tested model were found to be χ^2/df (167.10/80) = 2.089; GFI = 0.98; CFI = 0.99; SRMR = 0.02; RMSEA = 0.029. All fit indices obtained from the analysis were within the good fit reference range. The tested structural model showed a good fit with the data.

The structural model tested and the standardized path coefficients obtained from the analysis are shown in Figure 1. According to the analysis results shown in the figure, the direct path coefficient from need-supportive teaching style to BPN satisfaction (β=0.66; C.R=18.440; p<.001), intrinsic motivation (β=0.13; C.R =3.336; p<.001), extrinsic motivation (β=0.16; C.R=3.739; p<.001) was found to be positive, while the direct path coefficient to amotivation (β=-0.23; C.R= -5.541; p<.001) was found to be negative and statistically significant. In addition, the direct path coefficient from BPN satisfaction to intrinsic motivation (β=0.61; C.R=13.365; p<.001) and extrinsic motivation (β=0.54; C.R=11.067; p<.001) was found to be positive, while the direct path coefficient from amotivation (β=-0.44; C.R=-9.642; p<.001) was negative and statistically significant. Furthermore, when examining the r^2 values explaining the percentage of variance explained by independent variables, 43% of BPN satisfaction is explained by perceived need support (r^2= 0.429). Furthermore, 49% of intrinsic motivation (r^2 = 0.492), 43% of extrinsic motivation (r^2 = 0.430), and 38% of amotivation (r^2 = 0.382) were explained by BPN satisfaction and need support.

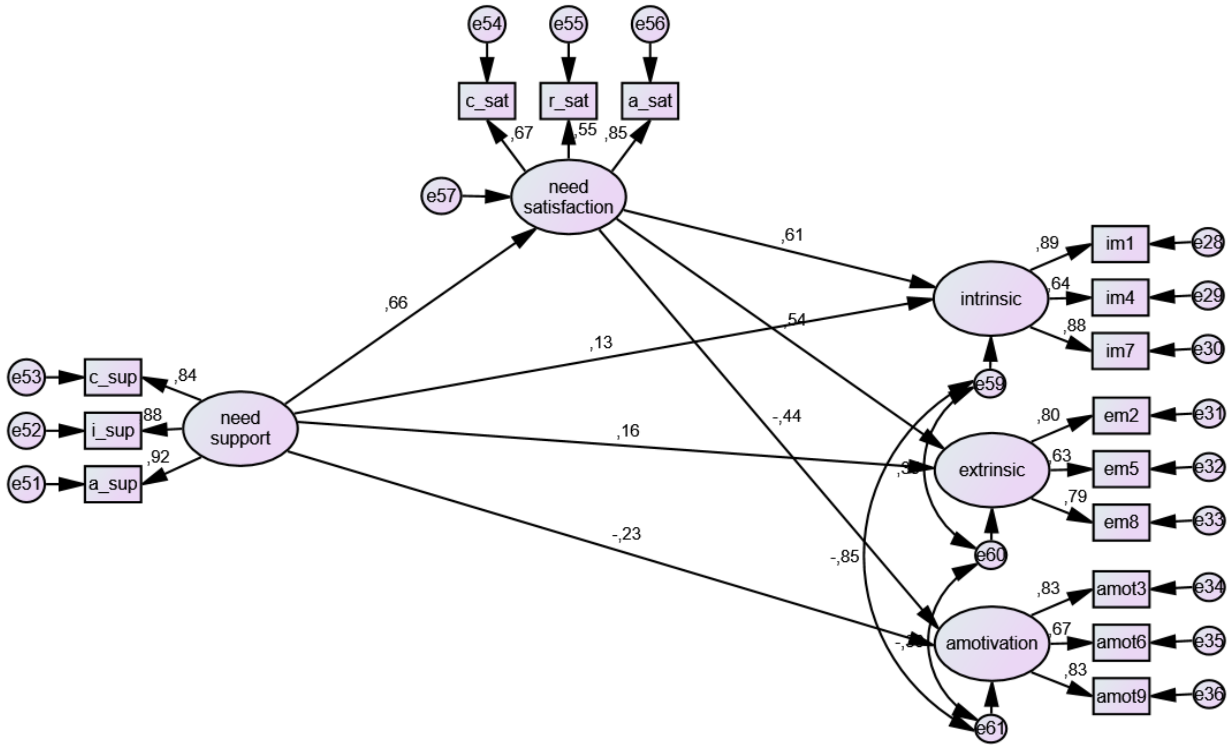


Figure 1. Standardized path coefficients related to the structural model

Note: c_sup - competence support, i_sup - involvement support, a_sup - autonomy support, c_sat - competence satisfaction, r_sat - relatedness satisfaction, a_sat - autonomy satisfaction, im - intrinsic motivation, em - extrinsic motivation, amot – amotivation.

In the current study, the mediating role of BPN satisfaction in the effect of perceived need support from PE teachers on motivational regulation and enjoyment was tested. Bootstrap 5000 resampling was preferred for the mediating effect analysis. It is known that the Bootstrap method provides highly reliable results in mediator variable analyses. Additionally, to conclude that a variable has a mediating effect, the values obtained from the analysis must not include zero (0) within the 95% confidence interval (Gürbüz, 2021). According to the Bootstrap results shown in Table 3, it was determined that the perceived need support has a significant indirect effect on intrinsic motivation through need satisfaction (β= 0.399; 95% Cl [0.335-0.473]). It can be seen that the lower and upper limits of the confidence interval do not include the value zero. These results indicate that perceived need-supportive teaching style partially mediates the effect of need satisfaction on intrinsic motivation.

Table 3. The mediating role of need satisfaction in the effect of perceived need support on intrinsic motivation.

Path	Effect Type	Standardized Coefficient (β)	SE	95% CI (Lower–Upper)	p-value	Significant
NsT → BPNs	Direct	0.65	0.027	[0.60, 0.70]	< .001	Yes
BPNs → IM	Direct	0.61	0.043	[0.52, 0.69]	< .001	Yes
NsT → IM	Direct	0.13	0.043	[0.04, 0.21]	.004	Yes
NsT → BPNs → IM	Indirect	0.40	0.035	[0.33, 0.47]	< .001	Yes
NsT → IM	Total	0.53	0.026	[0.47, 0.58]	< .001	Yes

According to the Bootstrap results presented in Table 4, it was determined that the perceived need support had a significant indirect effect on extrinsic motivation through need satisfaction (β= 0.354; 95% Cl [0.290-0.424]). It can be seen that the lower and upper limits of the confidence interval do not include the value zero. These results indicate that perceived need-supportive teaching style has a partial mediating effect on external motivation through need satisfaction.

Table 4. The mediating role of need satisfaction in the effect of perceived need support on extrinsic motivation.

Path	Effect Type	Standardized Coefficient (β)	SE	95% CI (Lower–Upper)	p-value	Significant
NsT → BPNs	Direct	0.65	0.027	[0.60, 0.70]	< .001	Yes
BPNs → EM	Direct	0.54	0.045	[0.45, 0.63]	< .001	Yes
NsT → EM	Direct	0.16	0.046	[0.06, 0.24]	.001	Yes
NsT → BPNs → EM	Indirect	0.35	0.034	[0.29, 0.42]	< .001	Yes
NsT → EM	Total	0.51	0.026	[0.45, 0.56]	< .001	Yes

NsT: Need-supportive Teaching BPNs: Basic Psychological Need Satisfaction EM: Extrinsic Motivation SE: Standard Error CI: Confidence Interval.

According to the Bootstrap results shown in Table 5, it was determined that the indirect effect of perceived need support on amotivation through need satisfaction was significant (β= -0.291; 95% Cl [-0.365, 0.224]). It can be seen that the lower and upper limits of the confidence interval do not include the zero value. These results indicate that perceived need-supportive teaching style partial mediates the effect of need satisfaction on amotivation.

Table 5. The mediating role of need satisfaction in the effect of perceived need support on amotivation

Path	Effect Type	Standardized Coefficient (β)	SE	95% CI (Lower–Upper)	p-value	Significant
NsT → BPNs	Direct	0.65	0.027	[0.60, 0.70]	< .001	Yes
BPNs → AM	Direct	-0.44	0.048	[-0.53, -0.35]	< .001	Yes
NsT → AM	Direct	-0.23	0.047	[-0.32, -0.13]	< .001	Yes
NsT → BPNs → AM	Indirect	-0.29	0.036	[-0.36, -0.22]	< .001	Yes
NsT → AM	Total	-0.52	0.027	[-0.57, -0.46]	< .001	Yes

NsT: Need-supportive Teaching BPNs: Basic Psychological Need Satisfaction AM: Amotivation SE: Standard Error CI: Confidence Interval.

Secondly, the structural model was created to determine the predictive effect of the perceived need support from PE teachers on students' enjoyment through BPN satisfaction and intrinsic motivation. In this context, the goodness-of-fit values for the tested model were found to be $\chi 2/df$ (78.929/32) = 2.467; GFI = 0.98; CFI = 0.99; SRMR = 0.02; RMSEA = 0.034. The goodness-of-fit values obtained were between good fit and acceptable fit values.

The standardized path coefficients obtained from the tested structural model are shown in Figure 2. According to the analysis results presented in the figure, there is a significant relationship from need support to enjoyment (β=0.15; C.R.=7.608; p<.001), and BPN satisfaction (β=0.67; C.R.=19.168; p<.001), from BPN satisfaction to intrinsic motivation (β=0.72; CR:=19.981; p<.001) and from intrinsic motivation to enjoyment (β=0.79; C.R.=34.685; p<.001) were found to be positive and statistically significant. Additionally, when examining the r^2 values, which indicate the percentage of variance explained by the independent variables, 45% of BPN satisfaction is explained by the need-supportive teaching style (r^2=0.455). Furthermore, 75% of enjoyment in PE class (r^2=0.757) was explained by need support, need satisfaction, and intrinsic motivation.

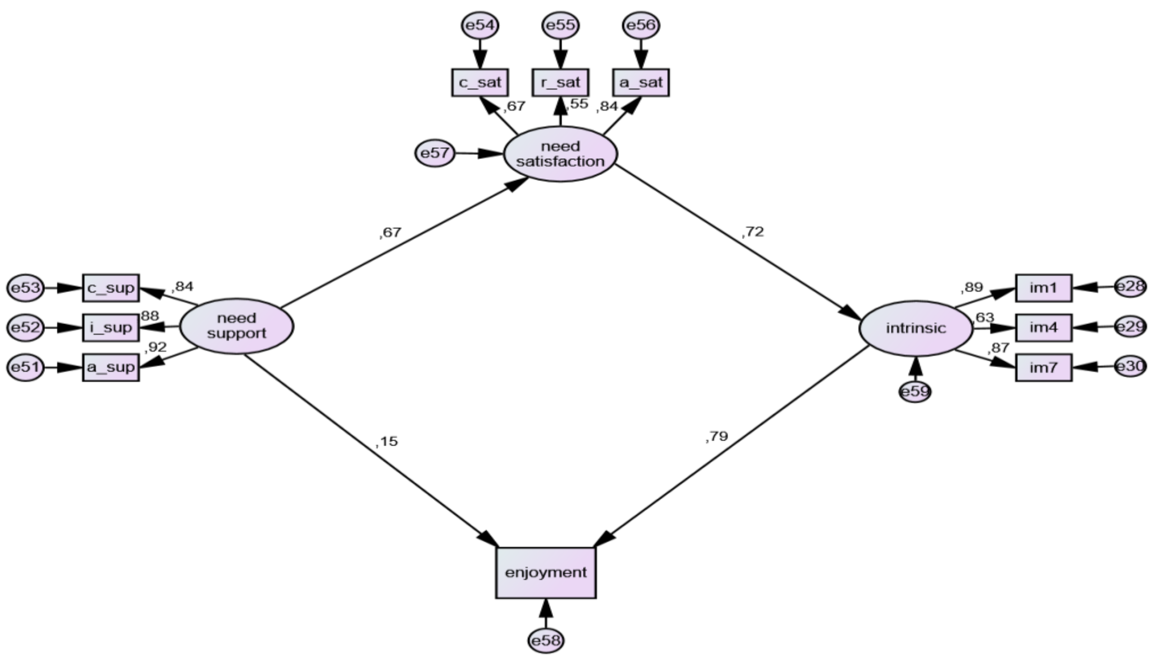


Figure 2. Standardized path coefficients related to the structural model
c_sup: competence support ***i_sup:*** involvement support ***a_sup:*** autonomy support ***c_sat:*** competence satisfaction ***r_sat:*** relatedness satisfaction ***a_sat:*** autonomy satisfaction ***im:*** intrinsic motivation

According to the Bootstrap results in Table 6, it was determined that the perceived need support had a significant indirect effect on enjoyment through need satisfaction and intrinsic motivation (β= 0.38; 95% CI [0.33-0.43]). It can be seen that the lower and upper limits of the confidence interval do not include the value zero. These results indicate that perceived need-supportive teaching style has a serial partial mediating effect on enjoyment through need satisfaction and intrinsic motivation.

Table 6. The mediating role of need satisfaction and intrinsic motivation in the effect of perceived need support on enjoyment.

Path	Effect Type	Standardized Coefficient (β)	SE	95% CI (Lower–Upper)	p-value	Significant
NsT → BPNs	Direct	0.67	0.025	0.62 – 0.72	< .001	Yes
BPNs → IM	Direct	0.72	0.022	0.68 – 0.76	< .001	Yes
IM → ENJ	Direct	0.78	0.026	0.73 – 0.83	< .001	Yes
NsT → ENJ	Direct	0.15	0.026	0.10 – 0.20	< .001	Yes
NsT → BPNs → IM → ENJ	Indirect	0.38	0.025	0.33 – 0.43	< .001	Yes
NsT → ENJ	Total	0.53	0.023	0.49 – 0.58	< .001	Yes

NsT: Need-supportive Teaching BPNs: Basic Psychological Need Satisfaction IM: Intrinsic Motivation ENJ: Enjoyment SE: Standard Error CI: Confidence Interval

Discussion

The present study aimed to investigate the structural relationships between teacher-provided need support and student need satisfaction, motivational regulations, and enjoyment. It was hypothesized that students who perceive their needs as supported would exhibit higher intrinsic motivation and greater enjoyment in PE classes while experiencing reduced amotivation. Two distinct structural models were developed within the scope of the research. The first structural model (Figure 1) confirmed the hypothesis that a supportive teaching style in PE lessons influences students' BPN satisfaction and motivational regulations, with the tested model demonstrating a good fit. The research found that middle school students' views on the autonomy, competence, and relatedness support they receive from teachers in PE classes positively influenced their overall satisfaction of psychological needs. Students' perception of need support had a direct and positive impact on both their intrinsic and extrinsic motivation, while it negatively impacted their amotivation. In the structural model, the satisfaction of basic psychological needs was shown to partially mediate the relationship between need-supportive teaching and all three types of motivation. As BPN satisfaction increased, so did students' intrinsic and extrinsic motivation, while their amotivation decreased. Overall, these results highlight the theoretical relevance of need-supportive teaching practices in promoting adaptive motivational outcomes among students in PE contexts.

Taken together, these findings align with the fundamental propositions of SDT (Deci & Ryan, 2000), suggesting that when individuals' BPN are met, they are more likely to develop autonomous forms of motivation, particularly intrinsic motivation. Conversely, when students experience low need satisfaction, they are more prone to amotivation or adopting less self-determined (i.e., lower-quality) motivational orientations. According to SDT, BPN satisfaction are crucial for developing and maintaining intrinsic motivation (Deci & Ryan, 1985; R. M. Ryan & Deci, 2000). In line with previous research, the current findings further illustrate the motivational and behavioral implications of need satisfaction in educational settings. Based on these findings, it can be assumed that prolonged amotivation may eventually lead to students' lack of participation in physical activity outside of PE classes. Indeed, students who experience a sustained lack of motivation toward physical activity are at a higher risk of developing a sedentary lifestyle (Ntoumanis et al., 2009).

According to Standage et al. (2005), students who perceived their PE classes as need-supportive reported higher levels of BPN satisfaction. Their findings showed that need-supportive teaching positively predicted intrinsic motivation and introjected regulation (a form of extrinsic motivation) through need satisfaction, while negatively predicting amotivation. Similarly, Ntoumanis (2005) found that PE teachers' support for students' needs for competence, relatedness, and autonomy significantly predicted their overall need satisfaction. This, in turn, was associated with more self-determined forms of motivation, which were linked to positive emotional, behavioral, and cognitive outcomes, as well as a greater intention to enroll in elective PE classes the following academic year. Zhang et al. (2011) identified that the link between perceived need support and physical activity is mediated by BPN satisfaction and intrinsic motivation. Their research demonstrated strong positive correlations among teacher-provided need support, need satisfaction, and intrinsic motivation. Rutten et al. (2012) found that BPN satisfaction mediates the relationship between students' perceptions of need-supportive teaching in PE classes and their autonomous motivation. Similarly, Leo et al. (2022) reported that students who perceived their teachers as engaging in need-supportive behaviors demonstrated stronger intentions to be physically active in the future, and this relationship was positively mediated by BPN satisfaction and autonomous motivation. In contrast, students who perceived their teachers as need-thwarting were more likely to report amotivation, which, in turn, negatively predicted their intentions to be physically active. Cox and Williams (2008) showed that the effect of perceived teacher need support on autonomous motivation was partially mediated by autonomy, competence and relationship needs satisfaction. Shen et al. (2010) reported that decreased need support from the teacher is an important factor resulting in students' amotivation. On the other hand, Franco and Coterón (2017) reported that a need-supportive intervention curriculum slightly increased students' intrinsic motivation, although the results were not statistically significant. The results of the present study are consistent with a growing body of literature suggesting that perceived need-supportive behaviors from teachers significantly enhance students' BPN satisfaction (Aibar et al., 2021; Burgueño et al., 2024; Franco & Coterón, 2017; Leo et al., 2022; Sanchez-Oliva et al., 2014; Vasconcellos et al., 2020). Furthermore, need-supportive teaching has been found to positively predict intrinsic motivation (Burgueño et al., 2024; Leo et al., 2022; Mouratidis et al., 2011; Sanchez-Oliva et al., 2014) and certain types of extrinsic motivation (Leo et al., 2022; Sanchez-Oliva et al., 2014) while negatively predicting amotivation (Burgueño et al., 2024; De Meyer et al., 2014; Leo et al., 2022; Vasconcellos et al., 2020). Collectively, these findings reinforce the theoretical proposition of SDT that the social environment, particularly through need-supportive teaching, plays a crucial role in the quality and regulation of students' motivation.

Secondly, considering the model (Figure 2), the hypothesis that a need-supportive teaching style affects students' enjoyment in PE courses through BPN satisfaction and intrinsic motivation was confirmed. The present study concluded that students' perceived need support has a significant positive effect on enjoyment, both directly and through BPN satisfaction and intrinsic motivation. According to the results of the quasi-experimental study conducted by Mouratidis et al. (2011), students enjoyed the lessons more and felt more energized when their teachers employed a need-supportive teaching style compared to a typical teaching style. The experimental group exposed to need-supportive practices differed positively and significantly from the students exposed to direct instructional practices in terms of enjoyment, interest, and vitality. Similarly, Franco and Coterón (2017) reported that students who participated in a need-supportive intervention program experienced greater enjoyment during PE classes. In research conducted with a middle school group, Standage et al. (2005) supported our conclusions by showing that teaching methods that support students' needs can indirectly boost enjoyment in PE by fulfilling BPN. Similarly, Huhtiniemi et al. (2019) emphasized the importance of students' experiences of need satisfaction and self-directed motivation in enhancing enjoyment during PE classes.

The results of the current study are largely consistent with SDT, with a few exceptions. Specifically, the positive impact of need-supportive teaching and overall need satisfaction on extrinsic motivation seems to contradict the theoretical assumptions of SDT. Nonetheless, these findings align with previous studies in the PE context (e.g., Behzadnia et al., 2018; Leo et al., 2022; Sanchez-Oliva et al., 2014; Standage et al., 2005). One possible explanation is the nature of introjected regulation, a type of extrinsic motivation that is only partially internalized and often involves internal conflict. As a result, it may show positive links with both beneficial and detrimental outcomes. In conclusion, more autonomous forms of extrinsic motivation are generally linked to positive outcomes, while less autonomous forms are associated with less favorable consequences (Vasconcellos et al., 2020). According to SDT, extrinsic motivation is not a unitary construct but consists of four qualitatively distinct, regulatory styles. Furthermore, the strong predictive role of relatedness satisfaction on external regulation, as reported by Vasconcellos et al. (2020), may offer a theoretical basis for our findings. Indeed, students' desire to maintain positive relationships with their teachers might enhance their levels of extrinsic motivation, particularly when driven by externally referenced goals. Ryan and Deci (2000) noted that externally motivated behavior is often performed to satisfy significant others in one's social environment, and that such motivation is not particularly conducive to sustained or persistent engagement. In line with this, the finding in our study that need support and need satisfaction positively influence external motivation reinforces our belief that students may be motivated to participate in PE classes in order to please their teachers.

Conclusion

The findings of the present study emphasize the importance of cultivating a supportive social environment in which educators deliberately foster the satisfaction of students' fundamental psychological needs, thereby enhancing their intrinsic motivation and enjoyment. Specifically, implementing a needs-based supportive teaching approach proves particularly effective in establishing an environment that enables students to engage autonomously and derive satisfaction from participating in PE classes. The validation of the structural models demonstrates that the motivational behaviors of PE teachers who employ a need-supportive instructional style—characterized by strategies that promote competence, autonomy, and relatedness—enhance students' BPN satisfaction, which subsequently encourages their voluntary participation in PE classes, driven by enjoyment and well-being. Moreover, this need-supportive style appears to inversely correlate with amotivation, indicating its potential as an effective tool for mitigating students' motivational deficits.

In light of these findings, the study largely corroborates the theoretical assumptions of SDT and affirms its applicability to Turkish middle school students. The results suggest that SDT provides a valuable theoretical framework for understanding teaching behaviors and motivational processes in PE and may serve as a foundation for designing effective intervention programmes aimed at enhancing learning outcomes in PE settings. Conversely, controlling teaching behaviors should be minimized or avoided, as even low levels of perceived teacher control can adversely affect students' motivation and affective outcomes (De Meyer et al., 2014). In summary, the current findings underscore that the teaching style adopted by educators plays a critical role in shaping positive outcomes among PE students and should therefore be a central consideration in teacher training and curriculum development efforts.

Acknowledgements: This article was produced from my doctoral dissertation titled *'Relationships Between Need-Supportive Teaching Style, Psychological Need Satisfaction, Motivation and Happiness Level in Secondary School Physical Education and Sports Course'* and is part of that dissertation.

***Declaration of interests*:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare no conflict of interest

***Data availability*:** All collected data are included in the manuscript. Raw data are available upon reasonable request to the corresponding author.

References

Aibar, A., Abós, Á., García-González, L., González-Cutre, D., & Sevil-Serrano, J. (2021). Understanding students' novelty satisfaction in physical education: Associations with need-supportive teaching style and physical activity intention. *European Physical Education Review*, *27*(4), 779-797. https://doi.org/10.1177/1356336X21992791

Akbulut, V., & Öncü, E. (2023). The Adaptation of Physical Education Motivation Scale into Turkish: A Validity and Reliability Study. *Sportive*, *6*(2), 139-152.

Behzadnia, B., Adachi, P. J. C., Deci, E. L., & Mohammadzadeh, H. (2018). Associations between students' perceptions of physical education teachers' interpersonal styles and students' wellness, knowledge, performance, and intentions to persist at physical activity: A self-determination theory approach. *Psychology of Sport and Exercise*, *39*, 10-19. https://doi.org/10.1016/j.psychsport.2018.07.003

Bolter, N. D., Kipp, L., & Johnson, T. (2018). Teaching Sportsmanship in Physical Education and Youth Sport: Comparing Perceptions of Teachers With Students and Coaches With Athletes. *Journal of Teaching in Physical Education*, *37*(2), 209-217. https://doi.org/10.1123/jtpe.2017-0038

Burgeson, C. R., Wechsler, H., Brener, N. D., Young, J. C., & Spain, C. G. (2001). Physical Education and Activity: Results from the School Health Policies and Programs Study 2000. *Journal of School Health*, *71*(7), 279-293. https://doi.org/10.1111/j.1746-1561.2001.tb03505.x

Burgueño, R., García-González, L., Abós, Á., & Sevil-Serrano, J. (2024). Students' motivational experiences across profiles of perceived need-supportive and need-thwarting teaching behaviors in physical education. *Physical Education and Sport Pedagogy*, *29*(1), 82-96. https://doi.org/10.1080/17408989.2022.2028757

Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of sports science & medicine*, *8*(2), 154.

Cox, A., & Williams, L. (2008). The roles of perceived teacher support, motivational climate, and psychological need satisfaction in students' physical education motivation. *Journal of sport and exercise psychology*, *30*(2), 222-239.

De Meyer, J., Soenens, B., Vansteenkiste, M., Aelterman, N., Van Petegem, S., & Haerens, L. (2016). Do students with different motives for physical education respond differently to autonomy-supportive and controlling teaching? *Psychology of Sport and Exercise*, *22*, 72-82. https://doi.org/10.1016/j.psychsport.2015.06.001

De Meyer, J., Tallir, I. B., Soenens, B., Vansteenkiste, M., Aelterman, N., Van den Berghe, L., Speleers, L., & Haerens, L. (2014). Does observed controlling teaching behavior relate to students' motivation in physical education? *Journal of Educational Psychology*, *106*(2), 541-554. https://doi.org/10.1037/a0034399

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Franco, E., & Coterón, J. (2017). The Effects of a Physical Education Intervention to Support the Satisfaction of Basic Psychological Needs on the Motivation and Intentions to Be Physically Active. *Journal of Human Kinetics*, *59*(1), 5-15. https://doi.org/10.1515/hukin-2017-0143

George, D., & Mallery, P. (2016). *IBM SPSS statistics 23 step by step: A simple guide and reference* (Fourteenth edition). Routledge, Taylor & Francis Group.

Green, K. (2002). Lifelong Participation, Physical Education and the Work of Ken Roberts. *Sport, Education and Society*, *7*(2), 167-182. https://doi.org/10.1080/1357332022000018850

Gürbüz, S. (2021). *AMOS ile Yapısal Eşitlik Modellemesi* (2. ed). Seçkin Publishing.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huhtiniemi, M., Sääkslahti, A., Watt, A., & Jaakkola, T. (2019). Associations among Basic Psychological Needs, Motivation and Enjoyment within Finnish Physical Education Students. *Journal of Sports Science & Medicine*, *18*(2), 239-247.

Leo, F. M., Mouratidis, A., Pulido, J. J., López-Gajardo, M. A., & Sánchez-Oliva, D. (2022). Perceived teachers' behavior and students' engagement in physical education: The mediating role of basic psychological needs and self-determined motivation. *Physical Education and Sport Pedagogy*, *27*(1), 59-76. https://doi.org/10.1080/17408989.2020.1850667

Liu, J.-D., & Chung, P.-K. (2017). Factor structure and measurement invariance of the Need-Supportive Teaching Style Scale for Physical Education. *Perceptual and Motor Skills*, *124*(4), 864-879.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports medicine*, *40*, 1019-1035.

Mouratidis, A. A., Vansteenkiste, M., Sideridis, G., & Lens, W. (2011). Vitality and interest–enjoyment as a function of class-to-class variation in need-supportive teaching and pupils' autonomous motivation. *Journal of Educational Psychology*, *103*(2), 353-366. https://doi.org/10.1037/a0022773

Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *British Journal of Educational Psychology*, *71*(2), 225-242. https://doi.org/10.1348/000709901158497

Ntoumanis, N. (2005). A Prospective Study of Participation in Optional School Physical Education Using a Self-Determination Theory Framework. *Journal of Educational Psychology*, *97*(3), 444-453. https://doi.org/10.1037/0022-0663.97.3.444

Ntoumanis, N., Barkoukis, V., & Thøgersen-Ntoumani, C. (2009). Developmental trajectories of motivation in physical education: Course, demographic differences, and antecedents. *Journal of Educational Psychology*, *101*(3), 717-728. https://doi.org/10.1037/a0014696

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of physical activity to fundamental movement skills among adolescents: *Medicine and Science in Sports and Exercise*, *33*(11), 1899-1904. https://doi.org/10.1097/00005768-200111000-00015

Rutten, C., Boen, F., & Seghers, J. (2012). How School Social and Physical Environments Relate to Autonomous Motivation in Physical Education: The Mediating Role of Need Satisfaction. *Journal of Teaching in Physical Education*, *31*(3), 216-230. https://doi.org/10.1123/jtpe.31.3.216

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Sakalli, D., & Kerkez, F. İ. (2021). Adaptation of the Need Supportive Teaching Style Scale for Physical Education to Turkish. *MSKU Journal of Education*, *8*(2), 655-673.

Sallis, J. F., McKenzie, T. L., Alcaraz, J. E., Kolody, B., Faucette, N., & Hovell, M. F. (1997). The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. *American Journal of Public Health*, *87*(8), 1328-1334. https://doi.org/10.2105/AJPH.87.8.1328

Sanchez-Oliva, D., Sanchez-Miguel, P. A., Leo, F. M., Kinnafick, F.-E., & García-Calvo, T. (2014). Physical Education Lessons and Physical Activity Intentions Within Spanish Secondary Schools: A Self-Determination Perspective. *Journal of Teaching in Physical Education*, *33*(2), 232-249. https://doi.org/10.1123/jtpe.2013-0043

Shen, B., Li, W., Sun, H., & Rukavina, P. B. (2010). The Influence of Inadequate Teacher-to-Student Social Support on Amotivation of Physical Education Students. *Journal of Teaching in Physical Education*, *29*(4), 417-432. https://doi.org/10.1123/jtpe.29.4.417

Standage, M., Duda, J. L., & Ntoumanis, N. (2003). A model of contextual motivation in physical education: Using constructs from self-determination and achievement goal theories to predict physical activity intentions. *Journal of Educational Psychology*, *95*(1), 97.

Standage, M., Duda, J. L., & Ntoumanis, N. (2005). A test of self-determination theory in school physical education. *British Journal of Educational Psychology*, *75*(3), 411-433. https://doi.org/10.1348/000709904X22359

Standage, M., Duda, J. L., & Ntoumanis, N. (2006). Students' Motivational Processes and Their Relationship to Teacher Ratings in School Physical Education: A Self-Determination Theory Approach. *Research Quarterly for Exercise and Sport*, *77*(1), 100-110. https://doi.org/10.1080/02701367.2006.10599336

Stroet, K., Opdenakker, M.-C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, *9*, 65-87. https://doi.org/10.1016/j.edurev.2012.11.003

Stroet, K., Opdenakker, M.-C., & Minnaert, A. (2015). Need supportive teaching in practice: A narrative analysis in schools with contrasting educational approaches. *Social Psychology of Education*, *18*(3), 585-613. https://doi.org/10.1007/s11218-015-9290-1

Sulz, L., Temple, V., & Gibbons, S. (2016). Measuring student motivation in high school physical education: Development and validation of two self-report questionnaires. *Physical Educator*, *73*(3), 530.

Turkay, H., Mumcu, H. E., Çeviker, A., Güngöz, E., & Kurtuluş, Ö. (2019). Adaptation of the Basic Psychological Needs in Physical Education Scale into Turkish. *Gazi University Gazi Faculty of Education Journal*, *39*(2), 1135-1155.

Uğraş, S., & Serbes, Ş. (2019). Validity and Reliability Study for Physical Education Course Happiness Level Scale. *Journal of Global Sport and Education Research*, *2*(2), 1-10.

Van Den Berghe, L., Soenens, B., Vansteenkiste, M., Aelterman, N., Cardon, G., Tallir, I. B., & Haerens, L. (2013). Observed need-supportive and need-thwarting teaching behavior in physical education: Do teachers' motivational orientations matter? *Psychology of Sport and Exercise*, *14*(5), 650-661. https://doi.org/10.1016/j.psychsport.2013.04.006

Vasconcellos, D., Parker, P. D., Hilland, T., Cinelli, R., Owen, K. B., Kapsal, N., Lee, J., Antczak, D., Ntoumanis, N., Ryan, R. M., & Lonsdale, C. (2020). Self-determination theory applied to physical education: A systematic review and meta-analysis. *Journal of Educational Psychology*, *112*(7), 1444-1469. https://doi.org/10.1037/edu0000420

Vlachopoulos, S. P., Katartzi, E. S., & Kontou, M. G. (2011). The basic psychological needs in physical education scale. *Journal of Teaching in Physical Education*, *30*(3), 263-280.

Zhang, T., Solmon, M. A., Kosma, M., Carson, R. L., & Gu, X. (2011). Need Support, Need Satisfaction, Intrinsic Motivation, and Physical Activity Participation among Middle School Students. *Journal of Teaching in Physical Education*, *30*(1), 51-68. https://doi.org/10.1123/jtpe.30.1.51

SPORTLOGIA
Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 09/09/2025
Accepted: 10/10/2025
Online first: 24/10/2025
UDK: 616.831-009.11-085:615.838

Review article

Effects of Aquatic therapy in the rehabilitation of spinal cord injury: A systematic review

[1]*Ajibade Omodoyinsola, [2]Ognjen Milić, [3]*Veljko Vukićević, [4]Ognjen Kukić,

[1]Universitair Sportcentrum KU Leuven, Faculty of Kinesiology and Rehabilitation Sciences, Leuven,Belgium
[2] University of Serbia, Faculty for Sport and Physical Education, Niš, Serbia
[3] University of Serbia, Faculty for Sport and Physical Education, Novi Sad, Serbia
*Correspondence: vukicevicveljko9@gmail.com

Abstract

Spinal cord injury (SCI) is a complex neurological condition that results in persistent physical, neurological, and physiological impairments, with secondary complications, particularly respiratory and cardiovascular, significantly contributing to morbidity and mortality. Integrated rehabilitation enables simultaneous improvements across multiple functional domains. Aquatic therapy, utilizing buoyancy, viscosity, and hydrostatic pressure, provides a unique environment to enhance motor, sensorimotor, and cardiorespiratory functions. This systematic review was conducted according to guidelines for transparent reporting of systematic reviews and meta-analyses (PRISMA). Included studies comprised randomized and non-randomized controlled trials, quasi-experimental studies, and case-control studies, involving over 300 adult participants of both sexes with complete or incomplete traumatic or non-traumatic SCI at all lesion levels according to the ASIA classification. Only studies with clearly defined primary outcomes of aquatic therapy were included, without restriction by publication date. Outcomes were analyzed using the International Classification of Functioning, Disability, and Health framework. Only primary outcomes were included in the meta-analysis, while other outcomes were presented narratively. Results demonstrated improvements in muscle strength, balance, and cardiorespiratory function, as well as enhancements in coordination, endurance, and quality of life. Aquatic therapy shows potential as an effective complementary rehabilitation method for adults with SCI, supporting the standardization of clinical practice and informing future evidence-based research.

Keywords: aquatic exercise, aquatics, aquatic therapy, aquatic physiotherapy, hydrotherapy

Citation: Omodoyinsola, A., Milić, O., Vukićević, V., Kukić, O. (2025). Effects of Aquatic therapy in the rehabilitation of spinal cord injury: A systematic review. *SportLogia 21(*2), 48-64
https://doi.org/10.63356/spl.2025.010

Introduction

Spinal cord injury (SCI) has an estimated global incidence of about 10.4 and 83 million inhabitants per year (Wyndaele & Wyndaele, 2006). Clinically, damage to the spinal cord may be complete or incomplete depending on the extent of damage to the cord. A complete injury generally means that both motor and sensory functions are completely lost at the lowest sacral segments, while incomplete spinal cord injury means a partial preservation of sensory or motor functions below the neurological level of injury, including the lowest sacral segments (Kirshblum et al., 2011). In clinical practice, completeness of injury is defined according to the American Spinal Injury Association (ASIA) rehabilitation and treatment guides, which is based on a classification ranging from A to D. A refers to complete injury with no motor or sensory skills preserved; B is an incomplete injury where sensory but not motor functions are preserved; C is an incomplete injury where motor functions are preserved below the neurological level and more than half of key muscles below the neurological level have a muscle power grade less than 3; D is an incomplete injury where motor function is preserved below the neurological level and at least half of key muscles below the neurological level have a muscle power grade of 3 or more; and E refers to normal preservation of motor and sensory functions (Kirshblum et al., 2011).

The resulting disability is defined as paraplegia or tetraplegia, referring to the number of limbs involved. Tetraplegia refers to damage at the cervical segment of the cord, resulting in an impairment in the function of the arms, trunk, legs, and pelvic organs, whereas paraplegia refers to damage in the thoracic, lumbar, or sacral segments of the cord, resulting in an impairment of the lower limbs while the upper limb is spared; trunk and pelvic organs may or may not be involved depending on the level of injury (Jacobs & Nash, 2004). Following a spinal cord injury, several health complications develop, beginning from the acute to the chronic phase of injury. Individuals experience complications and impairment in their physical, neuro-muscular, musculoskeletal, cardiovascular, respiratory, physiological functions, and other co-morbidities, which increase morbidity, mortality, and decrease quality of life (Jacobs & Nash, 2004; McKinley et al., 1999; McKinley et al., 2002). Pulmonary and cardiovascular complications are cited as a major cause of death in chronic SCI (Jacobs & Nash, 2004; McKinley et al., 2002; West et al., 2013) and have become a main focus during rehabilitation.

Management of resulting impairments is based on the clinical presentation of each individual (Gittler, 2003). These typically include improvement of spasticity, cardiovascular function, respiratory function, gait, muscle weakness, joint range of motion, impaired coordination, postural problems, and sensory dysfunction (Nas et al., 2015). Rehabilitation of specific goals are often separated and worked on one at a time. In a systematic review on non-pharmacological and pharmacological interventions for spasticity management, a non-pharmacological approach (TENS) was supported by Level 1 evidence and hydrotherapy by Level 2 evidence (Hsieh et al., 2007). Two reviews of strategies for gait rehabilitation indicated that approaches promoting repeated practice offer the highest benefit, including over-ground training and bodyweight-supported treadmill training, Locomat, and functional electrical stimulation (Lam et al., 2007; Wessels et al., 2010). Improvement in cardiovascular health through resistance training and arm ergometry has been shown to decrease the risk for cardiovascular disease in SCI (Warburton et al., 2007), whereas respiratory goals in the acute phase are achieved by mechanical ventilation with the aim to optimize respiratory status (Wong et al., 2012), and in the chronic phase, respiratory strengthening is recommended for optimization (Mueller et al., 2012). A systematic review showed respiratory training may improve respiratory function in SCI (Van Houtte et al., 2006).

As is evident, individual goals are often separated and require specific interventions, locations, and clinicians to be achieved. Therefore, an approach that offers an environment where treatment goals can be achieved simultaneously may be of benefit to optimize the duration of SCI rehabilitation. Aquatic therapy also offers such an environment by using warm water for healing and rehabilitation purposes (Hall et al., 2008), utilizing the physical properties of water (buoyancy, density, hydrostatic pressure, thermodynamics, and viscosity) to effect performance, responses, as well as changes in body structure and function (Becker, 2009). A systematic review on the effects of aquatic therapy in 2008 showed its effectiveness in several domains such as muscle strengthening, balance, gait, and cardiovascular and respiratory functions across a wide range of conditions, including stroke, rheumatoid arthritis, spinal cord injury, and many more (Geytenbeek, 2002). In recent years, aquatic therapy has become more popular, and more evidence has surfaced. Hence, the aim of this review is to quantify the evidence on aquatic therapy in the rehabilitation of SCI in adults, and categorize outcomes where its effects are demonstrated or yet to be demonstrated within the domains of the International Classification of Functioning, Disability and Health (ICF) for individuals with SCI, through a systematic review and meta-analysis in an attempt to guide clinical practice, future research, and health care policy.

The aim of this research is to systematically evaluate and quantify the existing scientific evidence on the effectiveness of aquatic therapy in the rehabilitation of adults with spinal cord injury (SCI). The research aims to identify, analyze and categorize the outcomes of the application of aquatic therapy according to the framework of the ICF, with special reference to its effects on muscle strength, balance and cardiorespiratory functions. In addition, the aim is to use the obtained findings to formulate recommendations for clinical practice, future research and health policy, taking into account that until now no systematic review of the literature focusing exclusively on the application of aquatic therapy in people with SCI has been conducted.

Methods

Protocol and Registration

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were used to report this systematic review in a transparent and clear manner. PRISMA is a statement of evidence-based minimum collection of 27 items developed for reporting systematic reviews and meta-analyses (Moher et al., 2010).

Eligibility Criteria

The following eligibility criteria were used:

Study Characteristics: Randomized controlled trials (RCTs), controlled trials without randomization, case-control studies, and quasi-experimental studies. Only studies with participants from an adult population (18 years and above), male and/or female, were included. This allows for generalizability of the results of this review to the injury-prone demographic as described by the World Health Organization (2013): males (20-70 years), females (15-60). Complete or incomplete injuries of traumatic or non-traumatic etiology were included. SCI of all lesion levels classified according to the ASIA classification (A to D), regardless of the duration of the trauma, were included (American Spinal Injury Association, 2015). Studies with clearly defined primary outcomes for aquatic therapy interventions were included. Being the first review of SCI and aquatic therapy, inclusion of studies irrespective of their publication date was done to have a broader selection of studies and to document historically the empirical evidence for the effectiveness of aquatic therapy in SCI, as in previous studies (Lam et al., 2007; Tamburella et al., 2013). All studies considered eligible were those reported in English Language.

Intervention Protocol: Exclusively, only studies with a clearly defined protocol of aquatic therapy were deemed eligible to be included – partial or complete immersion in water, single or group activities, swimming, and all methods of aquatic therapy. Intervention protocols that combined land and aquatic therapy without control or underwater treadmill and body weight support techniques were not considered eligible for this review.

Report Characteristics: Studies were considered eligible if the intervention reported was clearly described by its total length, duration of each session, and type of activity performed.

Information Sources

Information sources used for the search include databases and a manual hand search of references in selected articles and relevant journals. An electronic search of publications was done by two independent researchers, also referred to in this study as reviewers (A and B), using the following databases: EMBASE, CINAHL, Cochrane, SPORTDiscus, Web of Science, and PubMed from inception to the last day of the search. The aim of the search was to find as many potentially relevant works as possible. The electronic search took place on January 8, 2015. The search strategy contained concepts directly related to aquatic therapy and SCI using the following keywords: "aquatic exercise," "aquatics," "aquatic therapy," "aquatic physiotherapy," "hydrotherapy," "underwater exercise," and "spinal cord injury." Keywords were combined to foster maximal retrieval of literature. The search was not limited to publications in English language nor to any publication date. Below is an example of the search strategy used in one of the databases:

EMBASE: "aquatic exercise" or "aquatics" or "aquatic therapy" or "aquatic physiotherapy" or "hydrotherapy" or "underwater exercise" AND "spinal cord injury." The same input was made in the same order by two researchers in each database.

Study Selection

Titles retrieved from the search of the database were compiled and duplicates eliminated by reviewer A. Titles and abstracts of the residual studies were then read and assessed by reviewer B for relevance. Titles and abstracts were considered relevant if any of the concepts of aquatic therapy, hydrotherapy, aquatics, underwater exercise, water, aquatic physiotherapy, aquatic exercise, immersion, and spinal cord injury, cervical injury, thoracic injury, lumbar injury, neurologic disease, neurologic injury, were addressed, and irrelevant titles and abstracts were eliminated. The resulting studies were then examined by reviewer B for confirmation. Full texts of the articles were then obtained, and two reviewers proceeded to independently evaluate if the eligibility criteria described previously were met. A manual search of the reference list of selected articles, relevant peer-review journals, and conference proceedings was done by reviewer B to identify potential studies that fit the eligibility criteria. Additional studies were identified and screened similarly to the electronically retrieved articles. Any disagreement on selection was resolved by discussion between both reviewers.

Statistical analyses

A data extraction sheet developed by reviewer A, based on the template of the Cochrane Consumers and Communication Review Group data extraction template (Moher et al., 2009), was used. The extraction sheet was pilot-tested independently by both reviewers on 3 randomly selected studies; any disagreement was resolved by a discussion between both parties, and the sheet modified accordingly. Reviewer A independently extracted data from the articles included in the review. No authors were contacted for additional raw data of their study. Variables for which data were sought were: (1) characteristics of the participants (age, sex, height, weight, ASIA classification, level of paralysis, and onset) (2) inclusion and exclusion criteria of individual studies (Table 1), (3) type of aquatic therapy (type, water temperature, intensity and frequency, administrator and comparator) (4) outcomes (mobility, cardiovascular, respiratory, hematological, participation, self-efficacy) (Table 2).

Assessment for Bias in Individual Studies

The Scottish Intercollegiate Guidelines Network (SIGN) checklist was used to assess risk of bias (network Sig., 2001). Any bias, if encountered, is reported descriptively (Table 3).

Summary Measures

The means as reported in each study are used to assess and describe the summary effect measure for each outcome. Standardized differences in means and confidence interval were calculated from reported means.

Synthesis of Results

The ICF framework was used to guide the syntheses of the various outcomes assessed across studies for a standard description of health-related states and outcomes (World Health Organisation, 2001). Due to the large variation in study designs, interventions, and reported outcome measures, not all outcomes were included in the meta-analysis. Primary outcomes (Thomaz et al., 2005; Leal et al., 2010; Jung et al., 2014; Kesiktas et al., 2004; Silva et al., 2005) were reported in a meta-analysis (Fig 2, Fig 3) to estimate overall effect size of the intervention in studies with identical outcome measures. Effect size is calculated using standardized difference in means. All statistical analysis was conducted using Comprehensive Meta-Analysis software (meta-analysis C., 2015). Other outcomes are reported in a narrative manner (Tamburella et al., 2013; Bosch & Wells, 1991), similar to the approach in previous reviews (West et al., 2013). Risk of bias across studies was assessed by testing for heterogeneity.

Results

The PRISMA flow diagram illustrates the process of literature search and study selection. A total of 1,011 records were identified through searches of relevant electronic databases, along with four additional records retrieved from other sources. After removing duplicates and screening titles and abstracts, 15 studies were assessed in full text, of which seven met the inclusion criteria for the synthesis. The remaining studies were excluded due to unavailability of full text, inadequate population, unpublished data, or lack of a control group. Part of the included studies was incorporated into the meta-analysis.

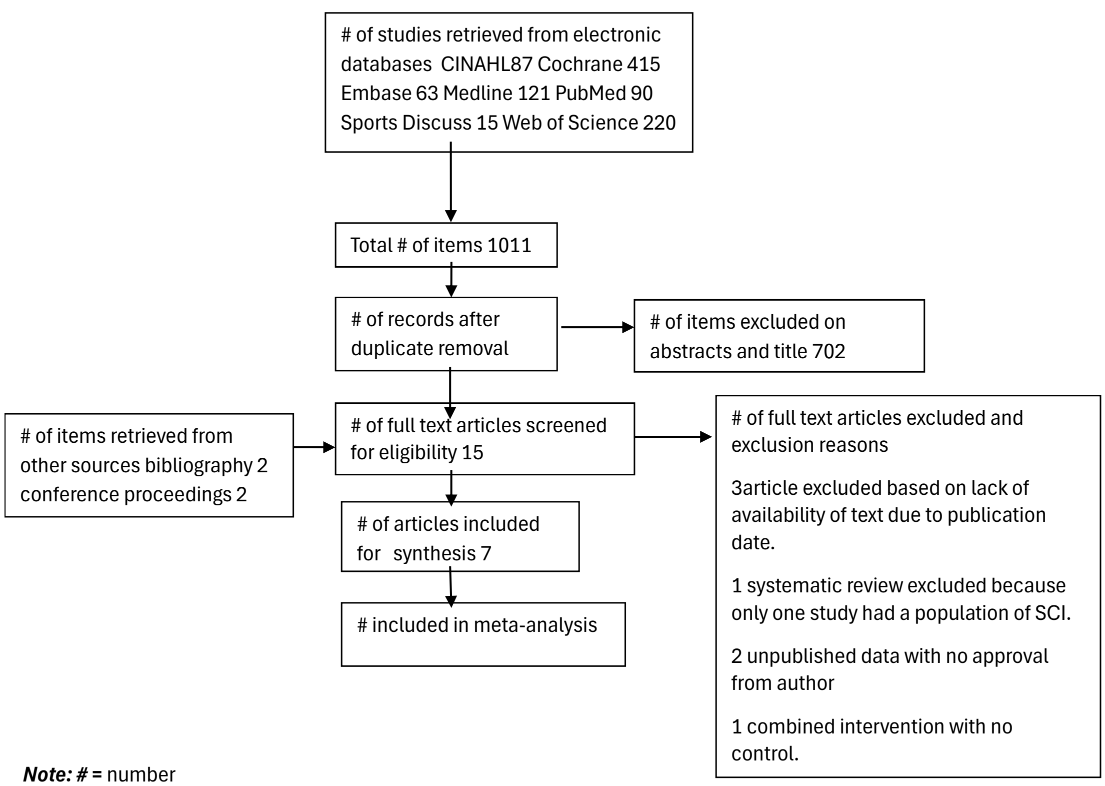


Figure 1. PRISMA Flow Diagram of Study Selection for Aquatic Therapy in Spinal Cord Injury Rehabilitation.

Table 1. Study characteristics.

Study	Outcome investigated	Purpose	No	Mean age	Gender	Asia impairment / Level of injury	Post injury months	Exclusion criteria	Mean height(cm)/weight(Kg)
Bosch et al 1991	Pulmonary function	Effect of immersion on residual volume of SCI and able-bodied	24	28.5	M	C4-C7, T3-T12	>6	medically unstable, onset <6months	67.7kg
Kesiktas et al. 2004	Mobility function	Effect of hydrotherapy on spasticity in SCI	20	32.13	M,F	A-D	8.6	Outpatient of rehab facility, no spasticity	-
Thomaz et al 2005	Pulmonary function	Effect of immersion on spirometry parameters of SCI and able bodied	23	27 median	M	A or B C4-C8	9	Prior tracheostomy, active pulmonary complications, clinical instability	-
DaSilva et al. 2005	Mobility, Selfcare	Effect of swimming on functional independence of SCI	16	25.75	M, F	A C5 L3	14-40	Contra indication to swimming, post injury >4yrs, ASIA >A, participation in rehab program <4weeks	-
Leal et al 2009	Pulmonary function, Hematologic function	Effect of graded immersion on VC in SCI	23	30.4	M	A or B / C4-C7	2-15	Clinical instability, history of tracheostomy, active respiratory disorder, conditions that prevent form partaking in water activities (skin infection and incontinence)	-
Tamburella et al 2013	Mobility	Walking in water VS Walking on land after SCI	30	43.5	M,F	D C4-T12		Cardiac or respiratory failure infective skin, excessively low, high or uncontrolled blood pressure, urinary tract infection, incontinence, morbid hydrophobia.	173.93cm 66.6Kg
Jung et al. 2014	Pulmonary function	Effect of aquatic exercise on pulmonary function in SCI	20	47.1	M,F	B-D	4-12	ASIA A, injury level <C8 >L5, outpatient of the rehab facility.	170cm/ 64Kg

Note: no= Number, M =Male, F =Female, cm =Centimeters, Kg= Kilograms

Study characteristics are shown in table 1. A total of 121 SCI individuals were investigated across all studies with no dropouts reported. Studies investigating pulmonary functions recruited 52 SCI participants in total. Two studies investigated only males (Thomaz et al., 2005; Leal et al., 2010), two investigated both males and females with mean age ranging from 27 to 51 years (Bosch & Wells, 1991; Jung et al., 2014). Severity of injury ranged from ASIA A-B in two studies (Thomaz et al., 2005; Leal et al., 2010), unspecified in one (Bosch & Wells, 1991), and ASIA A-D in one (Jung et al., 2014). Injury level ranged from C4-T12 across all four studies, with post-injury duration between 2 to 89 months. Two studies reported height and weight (Jung et al., 2014; Bosch & Wells, 1991), one reported weight only (Leal et al., 2010), and one reported neither (Thomaz et al., 2005). Exclusion criteria were not described in one study (Jung et al., 2014), while inclusion criteria were described in all four studies.

In the analyzed studies, aquatic interventions, including immersion (Bosch et al., 1990; Thomaz et al., 2005; Leal et al., 2009), swimming (DaSilva et al., 2009), water walking (Tamburella et al., 2013), and structured aquatic exercises (Jung et al., 2014), were conducted in water temperatures of 33–35°C. Session duration ranged from 5 to 60 minutes, with a frequency of 2–3 times per week over periods of 8 weeks to 4 months. Most interventions required assistance or the use of aids (e.g., wheelchairs, flotation devices, parallel bars, mechanical lifts), whereas activities in the control groups were performed on land (Kesiktas et al., 2004; DaSilva et al., 2009; Jung et al., 2014). A detailed overview of the interventions, including protocols, temperature, duration, frequency, administrators, and aids, is presented in Table 2**.** These protocols demonstrate that aquatic activities can be successfully adapted to meet different functional and respiratory needs of participants.

Table 2. Summary of interventions in individual studies.

Study	Intervention	Protocol	Temp(°C)	Intensity	Frequency	Administrator	Aids
Bosch et al. 1990	Immersion EG and CG	Spirometry measurement of residual volume in water in an upright seated position in an underwater weight tank. Spirometry measurement of residual volume on land in a seated position leaning forward from the waist, hydrostatic weighing, dry weight measurement first on land followed by underwater weighing in the tank in a seated position.	34	5 trials	-	Author	Wheelchair for transfers.
Thomaz et al 2005	Immersion	Immersion to shoulder level in an upright seated position, followed by spirometry in the water	33.5- 34.5	5-15mins	shoulder	Author	-
Kesiktas et al. 2004	(EG) CG	Underwater exercise, oral baclofen, PROM on land and psychotherapy Oral baclofen, PROM on land and psychotherapy	21	20mins 3Xweekly X10weeks	1.7m	-	Flotation devices, paddles, parallel bars, weighted stools or chairs
DaSilva et al. 2009	Swimming (EG) (CG)	Warm up: muscular stretching, transference training from wheelchair to the ground to swimming pool tested using different techniques. Main part: exercises aimed at independence change on the balance point and movement referential, crawl back and breast stroke exercises. Cool down: relaxation, fluctuation and respiration, techniques to exit the pool and transference from ground to wheelchair were trained. Locomotion training in wheelchair or gait training and sportive physical activities on land	-	2Xweekly 45mins 4months	-	prof. physical education	-

Leal et al 2009	Immersion EG (a) CG(a) EG (b) CG (b)	Spirometry performance at pool side followed by spirometry during graded immersion in an upright seated position in the pool to the pelvis, xiphoid and neck Spirometry on land with pressure cuffs around legs and groin deflated on land, then inflated to one-half of the pulse pressure followed by a repeat during graded immersion in water.	33.5- 34.5	5mins at each depth of immersion	pelvis, xiphoid and neck	Author	Mechanical crane
Tamburella et al 2013	Walking in water EG CG	Land walking at self- selected speed in a walkway. Walking at self- selected speed in the water with both arms placed above the surface of the water	35		xiphoid	-	-
Jung et al. 2014	Aquatic exercise (EG) CG	Warm up: ROM and breathing exercise, upper extremity exercises for function and weight bearing to left and right with arms at 90° in sitting, forward and backward movements in sitting position, with clasped hands, moving a heavy sensory ball forward, backward, side. Cool down: ROM, flexibility and breathing Same activities and protocol on land.	-	10min each phase, total 60min 3Xweekly for 8 weeks	-	-	-

Note: EG= experimental group, CG = control group

In a meta-analysis, the effects of aquatic therapy on respiratory functions in adults with spinal cord injury (SCI) were evaluated, through the four most frequently investigated parameters: forced vital capacity (FVC), vital capacity (VC), expiratory reserve volume (ERV) and inspiratory capacity (IC). Analysis of the combined results of two studies examining FVC (Thomaz et al., 2005; Jung et al., 2014) showed a statistically significant advantage of aquatic therapy compared to control interventions ($Z = 4.255$, $p < 0.001$), with complete homogeneity of results ($I^2 = 0\%$). A similar trend was observed in VC, where the results of two studies (Thomaz et al., 2005; Leal et al., 2010) indicated a significant increase in vital capacity after the application of aquatic therapy ($Z = 4.004$, $p < 0.001$, $I^2 = 0\%$).For ERV (Leal et al., 2010; Thomaz et al., 2005) a positive effect in favor of aquatic therapy was also recorded ($Z = 2.831$, $p = 0.005$), with moderate heterogeneity ($I^2 = 34.83\%$). In contrast, the results for IC (Leal et al., 2010; Thomaz et al., 2005) showed no statistically significant difference between the aquatic and control groups ($Z = 0.760$, $p = 0.447$, $I^2 = 0\%$). Overall, these findings indicate that aquatic therapy has a significant positive effect on key parameters of respiratory function (FVC, VC and ERV), while the effects on inspiratory capacity remain insufficiently confirmed. The low value of heterogeneity in most analyzes testifies to the consistency of findings among studies and confirms the potential of aquatic therapy as an effective rehabilitation intervention in the population of persons with spinal cord injury.

The methodological quality of the included randomized and controlled trials (RCTs and CCTs) was assessed using the SIGN checklist (Table 3a). All studies clearly stated their research objectives and included groups that were comparable in baseline characteristics. Standardized and reliable methods were employed to measure outcomes. Only one study implemented randomization, and procedures for allocation concealment and investigator blinding were often not reported, which could increase the risk of bias. The dropout rate was low (0%), and where feasible, analyses were conducted on an intention-to-treat basis. Based on the SIGN criteria, all studies were considered to have a minimal risk of bias (+) (Kesiktas et al., 2004; Jung et al., 2014; DaSilva et al., 2009).

Table 3. Risk of assessment in individual studies: SIGN scoring system (RCTs and CCTs).

Risk	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	2.1
Studies	**clear focused question**	**randomization**	**concealment**	**treatment blinding**	**Baseline groups similarity**	**Groups only differ in treatment under investigation**	**Standard and reliable outcome measures**	**Drop out (%)**	**Intention to treat analysis**	**Results comparable for all sites**	**Quality assessment based on bias minimality**
Kesiktas et al. 2004	yes	no	no	no	yes	yes	yes	0%	yes	N/A	+
Jung et al. 2014	yes	yes	can't say	can't say	yes	yes	yes	0%	yes	N/A	+
DaSilva et al. 2009	yes	no	no	no	yes	yes	yes	0%	yes	N/A	+

Risk of bias in individual studies was assessed using the SIGN system for RCTs and CCTs (Table 4). All studies had a clearly focused research question and comparable populations, with mostly consistent exclusion criteria. Participation rates ranged from 32% to 52%, and participants were generally compared with non-participants to reduce bias. Case-control differentiation was clear, and controls were confirmed as non-cases in most studies. Exposure measurement was generally valid and reliable. All studies were rated as having minimal risk of bias (+).

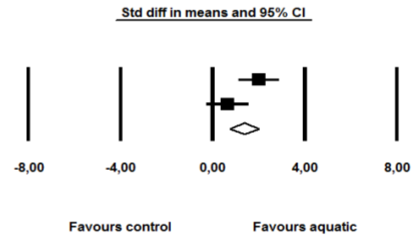
Table 4. Risk of assessment in individual studies: SIGN scoring system (RCTs and CCTs).

Risk	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	2.1
Studies	**Clear focused question**	**Comparable populations**	**Same exclusion criteria**	**Percentage participation (case and control respectively)**	**Participants and non-participants compared**	**Clearly defined differentiation of cases from controls**	**Controls are non-cases established**	**Case ascertainment bias prevented**	**Standard, valid, reliable measurement of**	**Potential confounders addressed**	**Set confidence intervals**	**Quality assessment based on bias minimality**
Bosch et al. 1991	yes	Yes	yes	50% 50%	yes	yes	yes	can't say	yes	yes	no	+
Thomaz et al. 2005	yes	No	yes	68% 32%	yes	yes	yes	no	yes	yes	yes	+
Leal et at. 2009	yes	Yes	yes	48% 52%	yes	Yes	yes	can't say	yes	yes	no	+
Tamburella et al. 2013	yes	Yes	yes	50% 50%	yes	yes	yes	Yes	yes	yes	no	**+**

Figure 2 presents the results of a meta-analysis showing standardized mean differences in lung function measures (FVC, VC, ERV, IC) across multiple studies, including confidence intervals and the weight of each study. It visually assesses whether the effects favor the control or aquatic intervention groups, along with the statistical significance and heterogeneity of the data.

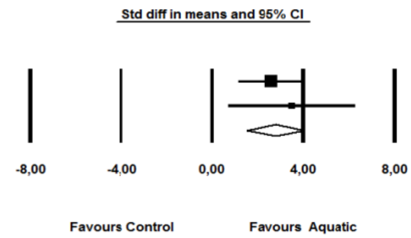
Study name	Outcome	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Relative weight
Thomaz et al. 2005	FVC	2,000	0,440	0,193	1,138	2,862	4,550	0,000	52,12
Jung et al. 2014 .	FVC	0,643	0,459	0,210	-0,256	1,542	1,402	0,161	47,88
		1,350	0,317	0,101	0,728	1,972	4,255	0,000	

I^2=78.090



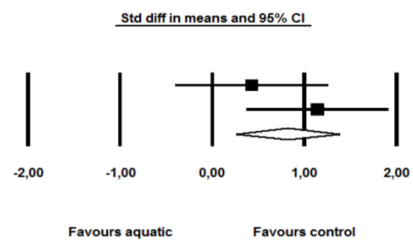
Study name	Outcome	Std diff in means	Standard error	Variance	Lower limit	Z-Value	Upper limit	p-Value	Relative weight
Thomaz et al. 2005	VC	2,577	0,729	0,532	1,147	3.533	4,007	0,000	79,22
Leal et al 2010	VC	3,496	1,424	2,028	0,705	2.455	6,287	0,014	20,78
		2,768	0,649	0,421	1,496	4.264	4,040	0,000	

I^2= 0.000



Study name	Outcome	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Relative weight
Leal et al 2010	ERV	0,430	0,422	0,178	-0,398	1,257	1,018	0,309	46,29
Thomaz et al 2005	ERV	1,144	0,392	0,154	0,375	1,912	2,918	0,004	53,71
		0,813	0,287	0,083	0,250	1,376	2,831	0,005	

I^2 =34.828



Study name	Outcome	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Relative weight
Leal et al 2010	IC	0,127	0,418	0,175	-0,692	0,946	0,305	0,760	43,80
Thomaz et al 2005	IC	0,339	0,369	0,136	-0,384	1,062	0,920	0,357	56,20
		0,247	0,277	0,076	-0,295	0,789	0,892	0,373	

I^2=0.00

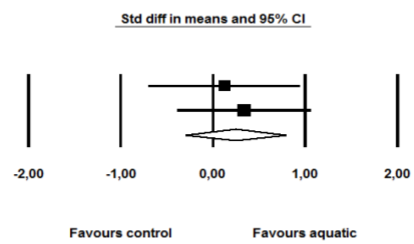


Figure 2. Meta-analysis showing summary of effect of aquatic therapy on respiratory functions in forest plot. **Note:** Standardized difference in means plot for effect, diamond =overall effect size of studies combined, squares = study effect size, size of box represents weight of study, horizontal lines represent CI=confidence Interval. Heterogenty= I^2

Figure 3 shows the overall effect of hydrotherapy on functional mobility, where squares represent individual studies, diamonds the combined effect of all studies. Y^2 = 0 and p = 1.000 indicate that there is no heterogeneity among studies, i.e. the results are consistent.

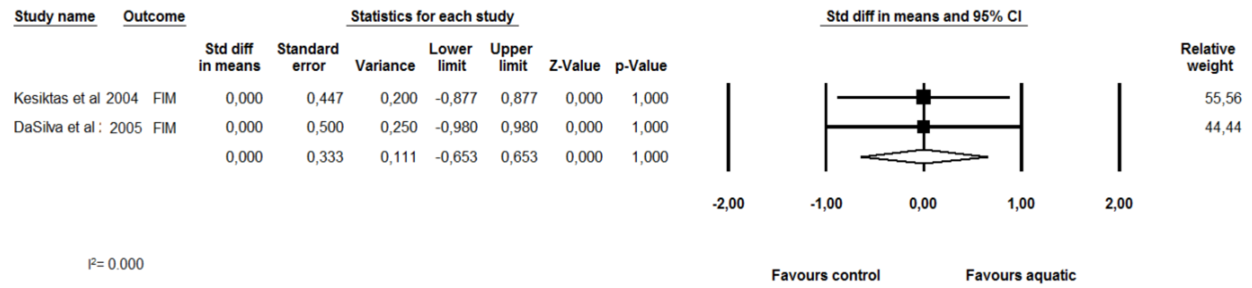


Figure 3. Meta-analysis showing summary of overall result for the effect of aquatic therapy on functional mobility in forest plot. **Note:** Standardized difference in means plot for effect, diamond =overall effect size of studies combined, squares = study effect size, size of box represents weight of study, horizontal lines represent Confidence interval = CI. Heterogenty I^2= 0 , p= 1.000.

Table 5 summarizes the effects of various interventions across several studies. Bosch et al. (1990) reported reductions in body density measures (RVW, QUAD) in the experimental group (*P<0.001*). Kesiktas et al. (2004) observed significant decreases in spasm severity and improvements in FIM scores with oral baclofen treatment (*p<0.05–0.0001*). Thomaz et al. (2005) found notable improvements in pulmonary function (FVC, FEV1, %FVC, %VC, %IC, %ERV) in the experimental group. DaSilva et al. (2009) reported significant enhancements in FIM sub-scores, including self-care, transference, communication, and social integration (*p<0.05*). Leal et al. (2010) demonstrated improvements in lung volumes, inspiratory and expiratory capacities, and hematocrit levels (p<0.008–0.017). Tamburella et al. (2013) observed significant gains in gait parameters (speed, stride length, stance phase, gait cycle time) (*p<0.05–0.005*). Jung et al. (2014) reported significant improvements in FVC, FER, FEV1, and FEV1/FVC ratios (**p<0.001–0.01*). Overall, the table highlights that the experimental interventions led to statistically significant improvements in functional, motor, and respiratory outcomes compared to control groups.

Table 5. Summary measures of individual studies.

Study	Primary outcome measure(s)	EG intervention values (SD)		CG intervention values (SD)		p
		Pre	Post	Pre	Post	
Bosch et al. 1990	RVW (L)	PARA 1.50(0.42)	1.24(0.19)	1.44(0.46)	1.40(0.41)	*P<0.001
		QUAD 2.43(0.34)	1.44(0.19)			
	Body density	PARA 1.049	1.043			
		QUAD 1.050	1.033	1.062	1.062	
Kesiktas et al. 2004	Ashworth	4.1	1.7	3.9	2.1	
	spasm severity	2.4	0.7**	2.3	1.4*	*p<0.05
	FIM score	52	94**	54.7	69.1*	**p<0.0001
	oral baclofen (mg)	100	45	96	96	
Thomaz et al. 2005	FVC (L)		3.3(0.7)**		4.7(0.7)**	
	FEV1(L)		2.9(-0.6)*		4.0(0.7)**	**p<0.001
	△ %FVC	2.9(14.6)	18.8(18.4)**	5.2(0.7)	-8.8(4.4)	*p<0.05
	△%VC	2.6(0.7)	24.7(28.1)	4.5(0.7)	-3.9(5.6)**	
	△ %IC		40.7(48.6) **		26.7(15.7)**	
	△% ERV		-6.4(58.4)		-62.7(15.9)**	

DaSilva et al. 2009	FIM total score	102.0	109.7*	100.4	103.1*	*p<0.05
	BODY CARE	34.7	38.0*	34.6	36.1*	
	to eat	6.7	6.7	6.6	6.4	
	to get ready	6.2	6.6	6.4	6.4	
	to shower	5.7	6.5*	5.1	5.7	
	to dress upper body	6.2	6.4*	6.4	6.4	
	to dress lower body	5.0	6.2*	5.1	5.5	
	to use bathroom	4.7	5.5	5.0	5.5*	
	TRANSFERENCE	14.9	19.2*	14.0	15.6*	
	bed, chair, wheelchair	5.1	6.4*	4.7	5.4	
	bathroom	4.9	6.5*	4.7	5.4	
	bathtub, shower	4.9	6.4*	4.5	4.9	
	Motor score: Subtotal	67.0	74.7*	65.2	68.4*	
	COMMUNICATION	14.0	14.0	14.0	14.0	
	comprehension	7.0	7.0	7.0	7.0	
	expression	7.0	7.0	7.0	7.0	
	SOCIAL INTEGRATION	21.0	21.0	20.9	21.0	
	Social interaction	7.0	7.0	6.9	7.0	
	resolution of problems	7.0	7.0	7.0	7.0	
	memory	7.0	7.0	7.0	7.0	
	Cognitive Score subtotal	35.0	35.0	34.9	35.0	
Leal et al. 2010	VC land (L)					***p<0.008
	pelvis	2.80(0.94)	-6.8 (-12.5)	5.20(0.74)	-1.9(5.7)	**p<0.017
	xiphoid		8.6(26.2)		-4.3(5.5)	
	neck		27.2(25.8)***		-6.3(5.0)***	
	IC (L)					
	IC neck		36.1(23.14)**		15.7(15.6)**	
	ERV (L)	2.32(0.06)		3.74(0.80)		
	ERV xiphoid	0.61(0.31)	40.3(51.5)**	1.46(0.55)	-27.6(21.9)**	
	ERV neck	NR		NR	56.4(16.3)**	
	Hematocrit level (%)		4		4	
Tamburella et al. 2013	Speed (m/s)	0.27(0.22)	0.17(0.09) *	0.95(0.21)	0.33(0.05)***	*p<0.05
	Stride length(cm)	0.67(0.21)	0.69(0.20)	1.25(0.24)	1.02(0.15)***	**p<0.005
	Stance phase (%)	71.49(7.52	60.60(15.48)*	65.13(2.71)	63.98(3.25)	
	Gait cycle time (s)	2.75(3.78)	3.78(1.11)**	1.34(0.2)	3.15(0.73)***	
Jung et al. 2014	FVC (L)	2.5(0.7)	4.3 (1.4)**	3.0(0.9)	3.4 (1.4)	***p<0.001
	FER (L)	80.5(15.5)	90.5(17.0)*	85.2(18.0)	90.6(18.0)*	**p<0.01
	FEV1(L)	2.1(0.9)	3.2(1.2)*	2.7(1.0)	2.9 (1.0)	*p<0.05
	FEV1/FVC	89.3(3.8)	93.0 (3.6)**	88.3 (4.6)	90.4 (3.2	

Note: EG – Experimental Group, CG – Control Group, RVW – Residual Volume in Water (L), FVC – Forced Vital Capacity (L), FEV1 – Forced Expiratory Volume in 1 second (L), VC – Vital Capacity (L), IC – Inspiratory Capacity (L), ERV – Expiratory Reserve Volume (L), FIM – Functional Independence Measure (total score and subscales), PARA – Paraplegic, TETRA – Tetraplegic, p – p-value (statistical significance).

Discussion

Results show that aquatic therapy improves respiratory function in people with SCI, especially forced vital capacity (FVC) and vital capacity (VC). Immersion up to neck level is beneficial due to hydrostatic pressure and strengthening of respiratory muscles (Thomaz et al., 2005; Jung et al., 2014). Respiratory parameters were assessed mainly by spirometry, including FVC, VC, inspiratory capacity (IC) and expiratory reserve volume (ERV), while mobility was assessed by Walking Index for Spinal Cord Injury II (WISCI-II) and Functional Independence Measure (FIM). While respiratory parameters improved, functional independence showed no significant improvement.

The body of evidence included in the meta-analysis is limited, therefore results are interpreted with great caution. Findings related to body structure and functioning suggest that aquatic therapy significantly improves respiratory functions in SCI individuals, especially when immersed to the neck (Thomaz et al., 2005; Jung et al., 2014). Forced Vital Capacity (FVC) and Vital Capacity (VC), major determinants of inspiratory strength (McKinley et al., 2002), show improvement during immersion and aquatic exercise. In the meta-analysis of studies investigating FVC (Figure 2a), results indicate that overall, aquatic therapy is highly effective in improving FVC during immersion to the neck as well as during aquatic exercise compared to land. However, the high effect size is accompanied by similarly high heterogeneity, and one study shows effectiveness but with no significance, suggesting a disparity between both studies. The effect of aquatic therapy on increasing VC (Thomaz et al., 2005; Leal et al., 2010) was even much higher than that of FVC, with a wide confidence interval and high homogeneity. Analysis on Inspiratory Capacity (IC) (Thomaz et al., 2005; Leal et al., 2010) showed immersion had a small effect in increasing IC that was not statistically significant, as well as a wide confidence interval demonstrating variability in the true effect. A decrease in Expiratory Reserve Volume (ERV) (Thomaz et al., 2005; Leal et al., 2010) during immersion was also highly effective with similarly wide confidence intervals as seen in IC increase and with high heterogeneity.

A significant level of morbidity and mortality is associated with respiratory dysfunction in persons with chronic SCI (Brown et al., 2006; Zimmer et al., 2007), which impacts respiratory efficiency and quality of life, especially in high lesion levels. Restriction of pulmonary volumes and respiratory capacity, including FVC, VC, and IC in SCI, has been reported in other studies (Schilero et al., 2009; Linn et al., 2000; Stepp et al., 2008) as consequences of respiratory dysfunction resulting from the injury. The resulting restriction of respiratory efficiency has been linked to respiratory muscle paralysis (Zimmer et al., 2007), hence the need for effective management to reduce impairment and fatality. Respiratory training, based on two systematic reviews of RCTs (Van Houtte et al., 2006; Sheel et al., 2008), has not demonstrated superior effectiveness on land. Whereas, findings in this review indicate that the use of aquatic therapy—either one-time immersion or exercise—is highly effective in the management of respiratory complications by improving lung volumes such as FVC, VC, and IC. This benefit may stem from respiratory muscle strengthening and hydrostatic pressure, similar to abdominal binding on land that has been demonstrated to improve lung volumes in persons with SCI (Langbein et al., 2001). One of the studies presented hematocrit level changes during immersion in both tetraplegic and able-bodied persons as a means to determine the underlying mechanism behind the increase in VC in its complementary study (Leal et al., 2010). It concluded that despite blood volume changes, hydrostatic pressure was responsible for the observed increase in VC in tetraplegics. This was supported by the finding that when pressure cuffs fitted around the leg at the groin were used to evoke a shift in blood volume on land, VC remained unchanged, in contrast to during immersion where inflation or deflation of cuffs increased VC in the tetraplegic group, but its extent did not differ based on inflation or deflation of the cuffs in water, thereby suggesting hydrostatic pressure as the true cause. Regardless of the mechanism involved (hydrostatic pressure or respiratory muscle strengthening), SCI patients with respiratory restrictions can certainly benefit from aquatic exercise or immersion to the neck level in a pool temperature between 33-34.5°C to enhance short-term respiratory functions. Short-term is specified because only one of the studies in our meta-analysis used a long-term intervention as opposed to immersion in improving lung volumes.

Similarly, immersion to the neck is specified, as it remains the common depth where significant improvement in lung volumes were recorded for all SCI individuals. Further research on the long-term benefits of immersion on respiratory restrictions in SCI will be of great value. In this interest, considerations for the assessment of respiratory function in SCI should also be of concern; it was observed during this review that spirometry testing equation of percentage predicted values were only modified to the SCI group in one study (Leal et al., 2010), similarly familiarization sessions were only conducted in two studies (Jung et al., 2014; Bosch & Wells, 1991). While we consider that this may not have influenced the results as gains were clearly recorded, it follows that specific volumes reported of the predicted range may not be accurate due to the flawed equation. In support of this, a cross-sectional study with 278 participants recommended the modification of spirometry testing standards to include excessive back-extrapolated volume to reduce the potential for bias during testing (Kelley et al., 2003). Also observed were the reported outcome measures. Studies that investigated respiratory responses or changes inconsistently reported VC, FVC, IC, ERV, RV, FEV1, FEV1/FEV. Studies tended to report only 3 out of all volumes even when the study aim referred to respiratory or pulmonary functions as a whole and not specific outcomes such as VC or FVC changes. This may be attributed to selective reporting across all studies, perhaps because outcomes were not in the interest of the author(s). Especially when armed with the knowledge that measurement of FVC, a measure of total lung capacity, is clinically accompanied by the measurement of RV (Langbein et al., 2001). Only one study reported change in RV during immersion (Bosch & Wells, 1991); it focused on changes in RV and its impact as a determinant of body density and body weight. It concluded that using RV as a derivative, body weight and body density differed if RV was measured on land or in water. The study concluded that body weight and density for tetraplegics should be derived from RV in water when it is reduced, but not for paraplegics or healthy population, as water or land measurements had no impact on their hydrostatic weight. These considerations should be taken into account for future studies.

Changes in mobility function differed in the results of this review. The inability or difficulty to walk is one of the most visible disabilities in SCI individuals (Dobkin et al., 2006), depending on lesion type. Muscle paralysis and spasticity are often the most common contributing factors (McKinley et al., 2002). Consequently, reducing spasticity and optimizing mobility independence and ambulation are crucial. A decrease in spasticity severity as well as oral baclofen intake was reported by one of the studies included in this review (Kesiktas et al., 2004). While a finite conclusion of spasticity reduction in SCI resulting from the use of hydrotherapy cannot be based on one study, as at the time of this review, it remained the only existing evidence to this effect. Nonetheless, similar findings of spasticity reduction in neurological diseases as a result of aquatic therapy have been reported (Becker, 1997; Beresneva et al., 2009; Geytenbeek, 2008), credited to the temperature of the water and the ease of conducting passive stretching in this environment. Future research in this area will provide much-needed concrete evidence, as land management of spasticity remains insufficient, with only pharmacological agents and transcutaneous electrical nerve stimulation (TENS) cited as possible management techniques (Hsieh et al., 2007). Functional independence on the other hand appeared to show no significant improvement based on our meta-analysis (Figure 2b). Both the aquatic intervention and conventional intervention did not effectively improve FIM scores, neither was more superior to the other as a 0 effect was attained. Yet, the two studies in the review reported a significant increase in overall FIM scores, although all increases were limited to transference and shower alone. It is plausible the flaw in both studies included in the meta-analysis evaluated improvement in independence with an inappropriate tool (FIM), as studies have shown the Spinal Cord Independence Measure (SCIMIII), which has demonstrated extensive validity and reliability, to be a more acceptable tool for measurement of independence in SCI individuals (Itzkovich et al., 2007). Previously, Oakley et al. (2013) in a single case study reported a 10% increase in SCIMIII scores of an SCI individual following aquatic activity-based restoration therapy. This study was not included in this review because it is a case study that combined both aquatic and land treatment without a control. Regardless, the strides made through the intervention may be of clinical consideration.

Changes in gait pattern during walking in water were explored in one of the studies included in this review (Tamburella et al., 2013). The study compared walking on land to walking in water. A decrease in gait speed and stance phase yielded values similar to those of healthy individuals. Similarly, hyper-flexion of the knee and hip were also reported in water. Ideally, clinical considerations for gait rehabilitation in SCI include reduction in speed, stance phase duration, and increased hip and knee flexion in order to allow transition to a more physiologic gait (Nudo, 2003), supporting the outcome of this study. Considering that gait rehabilitation of SCI on land is guided by the principles of task-acquisition, retention, and transfer principles of motor learning, i.e., functional repetition of a movement task to strengthen neural connectivity (Nudo, 2003; Marsh et al., 2011), gait training in water in a simple walkway as in the study of Tamburella et al. (2013) may be the key to providing the environment where optimal physiologic walking can be practiced through aquatic therapy. However, further research is needed to solidify these findings. Similarly, a case study (Rotondo et al., 2013) identified during our search, but not included in this review due to the study design, assessed gait-related functions in an individual with SCI following an aquatic therapy program including underwater treadmill adjunct to over-ground training. The results of the study revealed no change in SCIM III, however, the individual's walking index for spinal cord injury increased from 0-8, demonstrating an improvement in gait function similar to our reports above. As is evident in our search, underwater treadmill training in general was not included in this review as it is considered a body weight supported treadmill training technique (BWSTT), which is beyond the scope of this review and the approach hinges on its own specific principles similar to BWSTT on land.

The two additional unpublished studies obtained from conference abstracts identified during our search but not included in this review, as the authors did not permit publication of their results, investigated the effect of aquatic therapy on abdominal adiposity and insulin resistance in two people with chronic motor incomplete spinal cord injury with positive results (Geigle et al., 2015), while the other investigated cardiovascular and functional effects with results in favor of aquatic therapy (Geigle et al., 2015).

Conclusion

The evidence reviewed in this study indicates that aquatic therapy holds considerable promise for improving functional, motor, and respiratory outcomes in individuals with spinal cord injury (SCI). Across the included studies, participants undergoing aquatic interventions consistently demonstrated improvements in mobility, pulmonary function, muscle strength, body composition, and performance in activities of daily living, as reflected in measures such as FIM scores, FVC, FEV1, gait parameters, and body density. These findings underscore the unique advantages of aquatic therapy, which leverages buoyancy, resistance, and hydrostatic pressure to facilitate movement, reduce spasticity, and promote cardiovascular and muscular conditioning. However, the overall methodological quality of the existing literature is moderate, with only two controlled trials (one RCT and one CCT) and the remaining studies being case-control in design. Limitations such as small sample sizes, unequal group distributions, heterogeneous intervention protocols, and diverse outcome measures restrict the generalizability of findings and complicate meta-analytic synthesis. To address these gaps, future research should implement standardized reporting practices, including detailed participant characteristics, SCI-adapted spirometry, validated mobility assessments, and objective measures such as the Walking Index for Spinal Cord Injury (WISCI). Moreover, long-term outcomes, including quality of life, social participation, psychological well-being, and functional independence, remain largely unexplored and should be a focus of upcoming studies. High-quality, adequately powered randomized controlled trials are essential to confirm the efficacy of aquatic therapy, identify optimal intervention parameters, and inform evidence-based clinical guidelines. Overall, while current evidence is limited, aquatic therapy appears to be a safe, effective, and versatile intervention with the potential to significantly enhance rehabilitation outcomes and overall quality of life for individuals with SCI.

Acknowledgements: The authors acknowledge no conflicts of interest.

References

Becker, A. (1997). A Bad Ragaz Ring Method variation for use with the cervical spine. *J Aquatic Physical Therapy*, *3*. https://doi.org/10.1002/pmrj.12435

Becker, B. E. (2009). Aquatic therapy: scientific foundations and clinical rehabilitation applications. *PM&R*, *1*(9), 859–872. https://doi.org/10.1016/j.pmrj.2009.05.017

Beresneva, J., Stirane, D., Kiukucane, E., & Vetra, A. (2009). The use of aquatic therapy in stroke patients for the management of spasticity. *International Journal of Rehabilitation Research*, *32*, S110. https://doi.org/10.1051/shsconf/20120200033

Bosch, P. R., & Wells, C. L. (1991). Effect of immersion on residual volume of able-bodied and spinal cord injured males. *Medicine and Science in Sports and Exercise*, *23*(3), 384–388. https://doi.org/10.1249/00005768-199103000-00019

Brown, R., DiMarco, A. F., Hoit, J. D., & Garshick, E. (2006). Respiratory dysfunction and management in spinal cord injury. *Respiratory Care*, *51*(8), 853–870. https://doi.org/10.4187/respcare.05108%E2%80%91853

Dobkin, B., Apple, D., Barbeau, H., Basso, M., Behrman, A., DeForge, D., Ditunno, J., Dudley, G., Elashoff, R., Fugate, L., Harkema, S., Saulino, M., & Scott, M.; Spinal Cord Injury Locomotor Trial Group. *Weight-supported treadmill vs over-ground training for walking after acute incomplete SCI*. *Neurology.* 2006;66(4):484-493. https://doi.org/10.1212/01.wnl.0000202600.72018.39

Geigle, P., Kochunov, P., Perreault, J., & Scott, W. (2015). Abdominal adiposity, insulin resistance and prescribed aquatic exercise for two people with chronic motor incomplete spinal cord injury (CMISCI): A case report. Paper presented at: Aqua-Leuven 2015. https://doi.org/10.1179/2045772312Y.0000000033

Geigle, P., WS, VanHeil, L., Sweatman, M., Keith, T., & Gorman, P. (2015). Aquatic therapy and robotic assisted body weight supported treadmill training (RABWSTT) for individuals with chronic motor incomplete spinal cord injury (CMISCI): A case report. Paper presented at: Aqua-Leuven 2015. https://doi.org/10.1002/pmrj.12435

Geytenbeek, J. (2002). Evidence for effective hydrotherapy. *Physiotherapy*, *88*(9), 514–529. https://doi.org/10.1016/S0031%E2%80%919406(05)60134%E2%80%914

Geytenbeek, J. (2008). Aquatic physiotherapy evidence-based practice guide. National Aquatic Physiotherapy Group: Australian Physiotherapy Association: Sidney. https://doi.org/10.1016/S0031%E2%80%919406(05)60134%E2%80%914

Gittler, M. S. (2003). Acute Rehabilitation in Cervical Spinal Cord Injury. *Topics in Spinal Cord Injury Rehabilitation*, *9*, 60–73.

Hall, J., Swinkels, A., Briddon, J., & McCabe, C. S. (2008). Does aquatic exercise relieve pain in adults with neurologic or musculoskeletal disease? A systematic review and meta-analysis of randomized controlled trials. *Archives of Physical Medicine and Rehabilitation*, *89*(5), 873–883. https://doi.org/10.1016/j.apmr.2007.09.054

Hsieh, J., Wolfe, D., Connolly, S., Eng, J., Haddad, S., & Krassioukov, A. (2007). Spasticity after spinal cord injury: an evidence-based review of current interventions. *Topics in Spinal Cord Injury Rehabilitation*, *13*(1), 81. https://doi.org/10.5312/wjo.v6.i1.8

Itzkovich, M., Gelernter, I., Biering-Sorensen, F., Weeks, J., Laramee, M. T., Craven, B. C., Tonack, M., Hitzig, S. L., Glaser, E. H., & Zeilig, G. (2007). The Spinal Cord Independence Measure (SCIM) version III: reliability and validity in a multi-center international study. *Disability & Rehabilitation*, *29*(24), 1926–1933. https://doi.org/10.1080/09638280601046302

Jacobs, P. L., & Nash, M. S. (2004). Exercise recommendations for individuals with spinal cord injury. *Sports Medicine*, *34*(11), 727–751. https://doi.org/10.2165/00007256-200434110-00003

Jung, J., Chung, E., Kim, K., Lee, B-H., & Lee, J. (2014). The Effects of Aquatic Exercise on Pulmonary Function in Patients with Spinal Cord Injury. *Journal of Physical Therapy Science*, *26*(5), 707. https://doi.org/10.1589/jpts.26.707

Kelley, A., Garshick, E., Gross, E. R., Lieberman, S. L., Tun, C. G., & Brown, R. (2003). Spirometry testing standards in spinal cord injury*. *Chest*, *123*(3), 725–730. https://doi.org/10.1378/chest.123.3.725

Kesiktas, N., Paker, N., Erdogan, N., Gülsen, G., Bicki, D., & Yilmaz, H. (2004). The use of hydrotherapy for the management of spasticity. *Neurorehabilitation and Neural Repair*, *18*(4), 268–273. https://doi.org/10.1177/1545968304270002

Kirshblum, S. C., Burns, S. P., Biering-Sorensen, F., Donovan, W., Graves, D. E., Jha, A., Marino, R. J., Waring, W. P., & Young, W. (2011). International standards for neurological classification of spinal cord injury (revised 2011). *The Journal of Spinal Cord Medicine*, *34*(6), 535–546. https://doi.org/10.1179/204577211X13207446293695

Lam, T., Eng, J. J., Wolfe, D. L., Hsieh, J. T., & Whittaker, M. (2007). A systematic review of the efficacy of gait rehabilitation strategies for spinal cord injury. *Topics in Spinal Cord Injury Rehabilitation*, *13*(1), 32–57. https://doi.org/10.1310/sci1301-32

Langbein, W. E., Maloney, C., Kandare, F., Stanic, U., Nemchausky, B., & Jaeger, R. J. (2001). Pulmonary function testing in spinal cord injury: effects of abdominal muscle stimulation. *Journal of Rehabilitation Research and Development*, *38*(5), 591–598. https://doi.org/10.1097/MD.0000000000006328

Leal, J., Mateus, S., Horan, T., & Beraldo, P. (2010). Effect of graded water immersion on vital capacity and plasma volume in patients with cervical spinal cord injury. *Spinal Cord*, *48*(5), 375–379. https://pubmed.ncbi.nlm.nih.gov/19859079/

Linn, W. S., Adkins, R. H., Gong, H., & Waters, R. L. (2000). Pulmonary function in chronic spinal cord injury: a cross-sectional survey of 222 southern California adult outpatients. *Archives of Physical Medicine and Rehabilitation*, *81*(6), 757–763. DOI: 10.1016/s0003-9993(00)90107-2

Marsh, B. C., Astill, S. L., Utley, A., & Ichiyama, R. M. (2011). Movement rehabilitation after spinal cord injuries: emerging concepts and future directions. *Brain Research Bulletin*, *84*(4), 327–336. DOI: 10.1016/j.brainresbull.2010.07.011

McKinley, W. O., Gittler, M. S., Kirshblum, S. C., Stiens, S. A., & Groah, S. L. (2002). 2. Medical complications after spinal cord injury: Identification and management. *Archives of Physical Medicine and Rehabilitation*, *83*, Supplement 1(0), S58–S64. DOI: 10.1053/apmr.2002.32159

McKinley, W. O., Jackson, A. B., Cardenas, D. D., & Michael, J. (1999). Long-term medical complications after traumatic spinal cord injury: a regional model systems analysis. *Archives of Physical Medicine and Rehabilitation*, *80*(11), 1402–1410. DOI: 10.1016/s0003-9993(99)90251-4

Meta-analysis C. (2015). Comprehensive meta-analysis software. http://www.meta-analysis.com/index.php. Accessed 18 Jun 2015.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Physical Therapy*, *89*(9), 873–880. DOI: 10.1371/journal.pmed.1000097

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. DOI: 10.1016/j.ijsu.2010.02.007

Mueller, G., Hopman, M. T., & Perret, C. (2012). Comparison of respiratory muscle training methods in individuals with motor complete tetraplegia. *Topics in Spinal Cord Injury Rehabilitation*, *18*(2), 118–121. DOI: 10.2340/16501977-1097

Nas, K., Yazmalar, L., Şah, V., Aydın, A., & Öneş, K. (2015). Rehabilitation of spinal cord injuries. *World Journal of Orthopedics*, *6*(1), 8. DOI: 10.5312/wjo.v6.i1.8

Network Sig. (2001). Critical appraisal: notes and checklists. http://www.sign.ac.uk/methodology/checklists.html. Accessed 18 Jun, 2015.

Nudo, R. J. (2003). Functional and structural plasticity in motor cortex: implications for stroke recovery. *Physical Medicine and Rehabilitation Clinics of North America*, *14*(1), S57–S76. DOI: 10.1016/s1047-9651(02)00054-2

Oakley, A., & Becker, D. (2013). Aquatic activity-based restorative therapy in an individual with spinal cord injury. *J Spinal Cord Med*, *35*(5), 524–527. https://doi.org/10.1310/sci1801-34

Rotondo, K., Martin, R., & Sadowsky, C. (2013). Aquatic locomotor training improves over-ground gait in a patient with tetraplegia. *J Spinal Cord Med*, *36*(5), 524–567. DOI: 10.1016/j.apmr.2016.10.022

Schilero, G. J., Spungen, A. M., Bauman, W. A., Radulovic, M., & Lesser, M. (2009). Pulmonary function and spinal cord injury. *Respiratory Physiology & Neurobiology*, *166*(3), 129–141. DOI: 10.1016/j.resp.2009.04.002

Sheel, A. W., Reid, W. D., Townson, A. F., Ayas, N. T., Konnyu, K. J., & Team, S. C. R. E. R. (2008). Effects of exercise training and inspiratory muscle training in spinal cord injury: a systematic review. *The Journal of Spinal Cord Medicine*, *31*(5), 500. DOI: 10.1080/10790268.2008.11753645

Silva, M. C. R. d., Oliveira, R. J. d., & Conceição, M. I. G. (2005). Effects of swimming on the functional independence of patients with spinal cord injury. *Revista Brasileira de Medicina do Esporte*, *11*(4), 251–256. https://doi.org/10.1590/S1517-86922005000400010

Stepp, E. L., Brown, R., Tun, C. G., Gagnon, D. R., Jain, N. B., & Garshick, E. (2008). Determinants of Lung Volumes in Chronic Spinal Cord Injury. *Archives of Physical Medicine and Rehabilitation*, *89*(8), 1499–1506. DOI: 10.1016/j.apmr.2008.02.018

Tamburella, F., Scivoletto, G., Cosentino, E., & Molinari, M. (2013). Walking in water and on land after an incomplete spinal cord injury. *American Journal of Physical Medicine & Rehabilitation*, *92*(10), e4–e15. DOI: 10.1097/PHM.0b013e3182a1e6c3

Thomaz, S., Beraldo, P., Mateus, S., Horan, T., & Leal, J. C. (2005). Effects of partial isothermic immersion on the spirometry parameters of tetraplegic patients. *CHEST Journal*, *128*(1), 184–189. DOI: 10.1378/chest.128.1.184

Van Houtte, S., Vanlandewijck, Y., & Gosselink, R. (2006). Respiratory muscle training in persons with spinal cord injury: a systematic review. *Respiratory Medicine*, *100*(11), 1886–1895. DOI: 10.1016/j.rmed.2006.02.029

Warburton, D. E., Eng, J. J., Krassioukov, A., & Sproule, S. (2007). Cardiovascular health and exercise rehabilitation in spinal cord injury. *Topics in Spinal Cord Injury Rehabilitation*, *13*(1), 98–122. DOI: 10.1310/sci1301-98

West, C. R., Bellantoni, A., & Krassioukov, A. V. (2013). Cardiovascular function in individuals with incomplete spinal cord injury: a systematic review. *Topics in Spinal Cord Injury Rehabilitation*, *19*(4), 267–278. DOI: 10.1310/sci1904-267

Wong, S. L., Shem, K., & Crew, J. (2012). Specialized respiratory management for acute cervical spinal cord injury: a retrospective analysis. *Topics in Spinal Cord Injury Rehabilitation*, *18*(4), 283–290. DOI: 10.1310/sci1804-283

World Health Organisation. (2001). International classification of funtioning disability and health. http://www.who.int/classifications/icf/en/. Accessed 18 Jun, 2015.

World Health Organisation. (2013). Fact sheet spinal cord injury. http://www.who.int/mediacentre/factsheets/fs384/en/. Accessed 18 May, 2015.

Wyndaele, M., & Wyndaele, J. J. (2006). Incidence, prevalence and epidemiology of spinal cord injury: what learns a worldwide literature survey? *Spinal Cord*, *44*(9), 523–529. DOI: 10.1038/sj.sc.3101893

Zimmer, M. B., Nantwi, K., & Goshgarian, H. G. (2007). Effect of spinal cord injury on the respiratory system: basic research and current clinical treatment options. *The Journal of Spinal Cord Medicine*, *30*(4), 319. DOI: 10.1080/10790268.2007.11753947

SPORTLOGIA

Scientific-Expert Journal of Anthropological Aspects of Sports, Physical Education and Recreation

Submitted: 23/09/2025
Accepted: 18/11/2025
Online first: 08/12/2025
UDK: 616.831-004-092:796.01/.09

Review article

Evaluating Motor Deficits in Multiple Sclerosis Using Jump and Hop Tests: A Review of Current Evidence

[1,*]Eva Žura, [1]Adis Burzić, [2]Erik Paulin, [2,*]Žiga Kozinc

[1]Department of Neurology, University Medical Centre Ljubljana, Ljubljana, Slovenia
[2]University of Primorska, Faculty of Health Sciences, Izola, Slovenia
*Correspondence: ziga.kozinc@fvz.upr.si

Abstract

Multiple sclerosis (MS) is a chronic neurological disease that often leads to subtle motor impairments even in early stages. Traditional clinical assessments may fail to detect these early deficits. Jump and hop tasks, requiring complex neuromuscular coordination, have emerged as promising functional assessments in neurological populations.This review aims to synthesize current evidence on the use of jump- and hop-based assessments to evaluate motor performance in people with MS (pwMS). A systematic search of PubMed and Scopus was conducted in April 2025. Studies were included if they involved pwMS, used jump or hop tasks, and reported performance, kinetic, or kinematic outcomes. Nine studies met inclusion criteria. Countermovement jumps (CMJ) were most frequently used and consistently detected motor deficits in pwMS, including reduced flight time, jump height, and power output. Single-leg CMJ tests identified asymmetries correlated with disability scores. Bipedal hops revealed impaired coordination and anticipatory control. One study assessed motor fatigability through repetitive jumping, while another evaluated the patient experience of sensor-based jump testing. Jump and hop assessments provide potentially sensitive, non-invasive tools for detecting early motor impairments in MS. They offer potential for improving clinical monitoring and guiding individualized rehabilitation strategies.

Keywords: neuromuscular assessment, functional mobility, countermovement jump, motor fatigability, asymmetry detection

Citation: Žura, E., Burzić, A., Paulin, E., Kozinc, Ž. (2025). Evaluating Motor Deficits in Multiple Sclerosis Using Jump and Hop Tests: A Review of Current Evidence. *Sportlogia. 21(*1), 65-73.
https://doi.org/10.63356/spl.2025.013

Introduction

Multiple sclerosis (MS) is a progressive, inflammatory, and degenerative disease of the central nervous system (Compston and Coles, 2008; Garg and Smith, 2015). In the brain and spinal cord, there is damage and loss of axons (Garg and Smith, 2015), as well as the spread of demyelinating plaques (Reich et al., 2018). People with MS may report a wide range of symptoms in their medical history, such as sensory loss, paraesthesia, visual disturbances, weakness in the upper and/or lower limbs, facial muscle weakness, ataxia, dizziness, paroxysmal symptoms, heat sensitivity, pain, urological problems, and fatigue (Hauser and Cree, 2020). The disease manifests in various forms. It most commonly begins as a clinically isolated syndrome (CIS), which causes symptoms such as visual problems, tingling, difficulties with walking and balance, altered muscle tone, fatigue, muscle weakness, urinary disturbances, and dizziness (Milo and Miller, 2014; Fisniku et al., 2008). Later, the disease may progress to relapsing-remitting MS (RRMS), characterized by periods of relapses and remissions. A relapse leads to a worsening of the individual's general condition and affects autonomic, sensory, cognitive, and motor functions (Doshi and Cathaway, 2016; Rovaris et al., 2006). Within 10–15 years, RRMS usually develops into secondary progressive MS (SPMS), which is no longer characterized by relapses but instead shows a constant worsening of symptoms (Klineova and Lubin, 2018). About 10–20% of individuals develop the rarest form, primary progressive MS (PPMS), which is marked by gradual worsening from the onset (Compston and Coles, 2008).

The disease brings with it numerous symptoms and requires a multidisciplinary approach that, in addition to a neurologist, usually includes specialized nurses, physiotherapists, occupational therapists, speech therapists, and psychologists (Feinstein and Freeman, 2015; Uygunoglu et al., 2016; Soelberg et al., 2019). Intensive neurophysiotherapy is important both for therapeutic interventions and for precise and meaningful assessment, which forms the basis for identifying the patient's movement disorders, detecting changes, and monitoring the effectiveness of therapeutic procedures. It is reasonable to monitor the individual's motor skills from the time of diagnosis, as a good initial assessment can identify existing issues and help target them effectively. The assessment tools used for this purpose must be reliable, accurate, and sensitive enough to detect subtle changes and deviations from functional movement.

Jumping tasks represent a complex motor skill that integrates muscle strength, coordination, balance, and temporal precision (Klavora, 2000). Unlike isolated strength or balance tests, jump-based assessments require the individual to generate and control rapid force across multiple joints in a sequential and functional manner. These characteristics make jump tasks particularly valuable for detecting subtle neuromuscular impairments that might not be evident in simpler clinical tests. Given the increasing interest in their use for functional assessment, even in older adults (Santos et al. 2022) and populations with neurological conditions (Reina et al., 2018), jumps have emerged as promising tools for evaluating motor performance in people with MS. For example, Geßner et al. (2023) demonstrated that countermovement jumps could detect early motor deficits in individuals with MS even when conventional clinical assessments indicated minimal disability. Despite the growing use of jump-based assessments in neurological research, their application in MS remains relatively underexplored and lacks synthesis. Given the importance of early detection of motor impairments—and the limitations of conventional clinical scales in capturing subtle neuromuscular changes—there is a clear need to evaluate whether and how these complex motor tasks can serve as reliable, sensitive, and functionally meaningful indicators of motor decline. This review therefore aims to consolidate current findings, identify methodological gaps, and assess the potential of jump-based metrics to complement or enhance standard neurological evaluations in people with MS.

Methods

A structured literature search was conducted in April 2025 using PubMed and Scopus databases. The search was independently performed by two authors using the following Boolean search string: ("neuromuscular disorder*"[tiab] OR "neuromuscular disease*"[tiab] OR "motor neuron disease"[tiab] OR "muscular dystrophy"[tiab] OR "multiple sclerosis"[tiab] OR "Parkinson's disease"[tiab] OR "amyotrophic lateral sclerosis"[tiab] OR "spinal muscular atrophy"[tiab] OR "cerebral palsy"[tiab] OR "Guillain-Barré syndrome"[tiab] OR "myasthenia gravis"[tiab] OR "Charcot-Marie-Tooth disease"[tiab] OR "myopathy"[tiab] OR "neuropathy"[tiab]) AND ("vertical jump" OR "squat jump" OR "countermovement jump" OR "CMJ"[tiab] OR "SJ"[tiab]). No restrictions were applied regarding publication year, but only articles published in English and involving human subjects were considered. Full texts of eligible studies were retrieved and reviewed to confirm inclusion. In addition, backward (reviewing reference list of included papers) and forward (using "cited by" function in Google Scholar) citation tracking was performed on all included articles to identify any additional relevant studies.

After the removal of duplicates, the titles and abstracts were screened for relevance. Studies were included if they met the following criteria: (1) they involved participants with a diagnosis of multiple sclerosis, (2) they applied any type of jump or hop task as a motor assessment tool, and (3) they reported relevant performance, kinematic, or kinetic outcome measures. Full texts of eligible articles were then reviewed to confirm inclusion. Disagreements between the reviewers were resolved through discussion and consensus. Data extraction included study aims, participant characteristics, measurement tools, jump/hop protocols, and key findings. The methodological approach of each study was also noted, including the types of parameters analyzed (e.g., jump height, ground contact time, force, power, asymmetry). A narrative synthesis of results was conducted due to heterogeneity in study designs and outcome measures.

Results

We found nine studies that investigated hops or jumps measurements in different variations, published between 2007 (Pérez et al., 2007) and 2025 (Geßner et al., 2024). Details about individual studies are shown in Table 1, and the study selection process is shown in Figure 1.

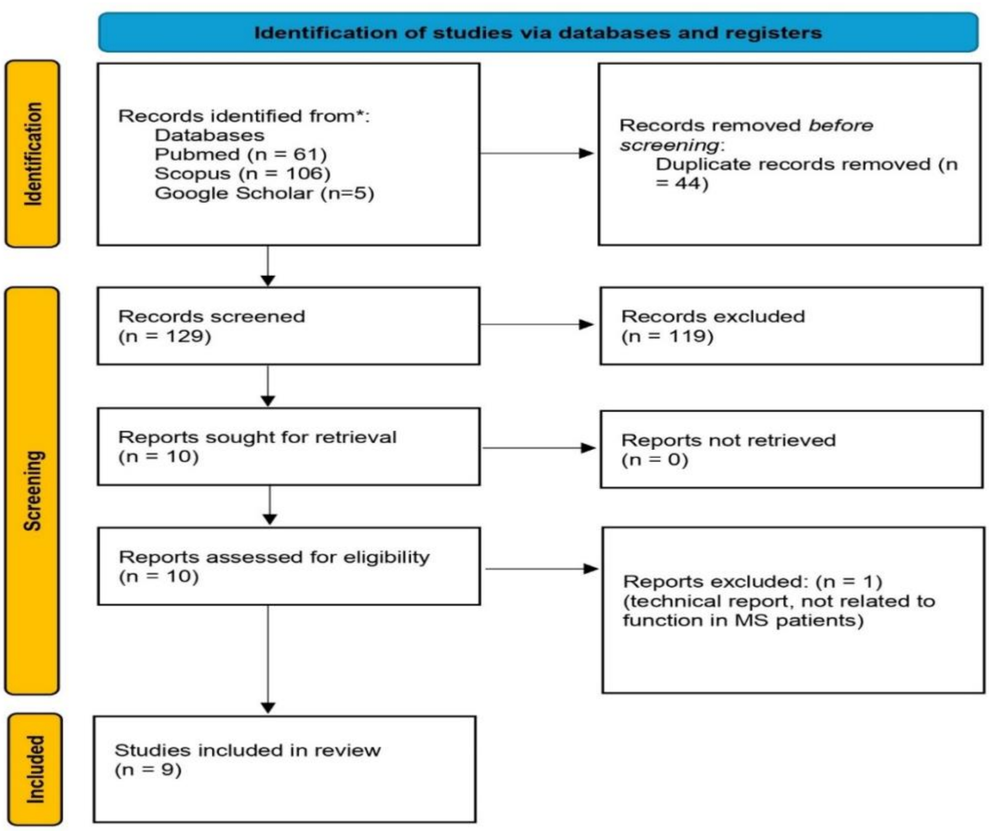


Figure 1. PRISMA flowchart diagram.

Table 1. Details of individual studies included in the review.

Author, Year, Article Title	**Purpose**	**Participant characteristics**	**Measurement methods**	**Results /Findings**
Geßner et al., 2021 (Quantification of motor fatigue by Jump analysis in people with Multiple Sclerosis)	To evaluate a 10-second jump test to detect muscle fatigue in people with MS.	- 30 individuals with MS (mean age: 33.8 ± 8.1 years; BMI: 23.6 ± 3.1; EDSS: 1.75) - 15 healthy controls (mean age: 39.3 ± 10.8 years; BMI: 24.1 ± 2.8)	Test: As many maximal jumps as possible in 10 seconds Tool: Force plate (AMTI, Accu Power-O) Measured parameters: ground contact time (GCT), reactive strength index (RSI), peak power (PP), push-off impulse (PI), number of jumps Additional: EDSS, neurological assessment, GLMM analysis	MS individuals performed fewer jumps (9.9 ± 2.9) than healthy controls (12.9 ± 2.4). RSI and GCT were negatively associated with functional scores (EDSS). Peak power was negatively associated with cerebellar function in EDSS.
Geßner et al., 2024 (The Association of Age, Sex, and BMI on Lower Limb Neuromuscular and Muscle Mehanical Function in People with Multiple Sclerosis)	To research the association of age, sex, and BMI with muscle mechanical function in people with MS using the Countermovement Jump (CMJ) test.	- 164 individuals with MS (18–65 years, EDSS 0–3.0, able to walk > 500 m unaided, perform basic motor tasks – walking on heels/toes) - 98 healthy controls	Tool: Force plate (AMTI, Accu Power-O), 1000 Hz Analysis: AccuPower Solutions (v1.5.4.2082) Parameters: positive/negative power, force at peak acceleration/deceleration, time to take-off, flight time, jump height, take-off velocity Statistical analysis: mean of 3 jumps, Shapiro Wilk, GLMM, Bonferroni correction, Cohen's d, Spearman correlation	Age, sex, and BMI significantly influenced all jump parameters (flight time, height, power). MS individuals showed poorer jump performance than healthy participants, especially in middle-aged (31–49), normal or overweight, and in both sexes. CMJ can detect subtle motor deficits not visible in standard neurological assessments.
Geßner et al., 2024 (Sensitive Identification of Asymmetries and Neuromuscular Deficits in Lower Limb Function in Early Multiple Sclerosis)	Early detection of neuromuscular deficits and asymmetries in lower limbs using single-leg jump (SLCMJ).	- 126 individuals with MS (mean age: 36.7 ± 8.5 years; 66.7% female; BMI: 24.8 ± 4.6; EDSS: 1.5) - 79 healthy controls (mean age: 37.9 ± 10.9 years; 60.1% female; BMI: 24.2 ± 3.4)	Tool: Force plate (AMTI, 1000 Hz) Test: 3 single-leg CMJ jumps Parameters: push-off time, flight time, landing time, force at zero velocity, GRF, negative/positive/peak power, jump height Analysis: ANCOVA (covariates: age, BMI, sex)	PwMS had shorter flight time, lower force, reduced power, and lower jump height. Greater inter-limb asymmetries in MS, correlated with higher EDSS. SLCMJ is suitable for detecting subclinical deficits and asymmetries in lower limb function in MS.
Geßner et al., 2023 (Countermovement Jumps Detect Subtle Motor Deficits in People with Multiple Sclerosis below the Clinical Threshold)	Using CMJ to identify early motor deficits in people with MS.	- 77 individuals with MS without walking difficulties (age: 35.9 ± 8.8; 70.7% female; BMI:24.6 ± 4.3; EDSS: 1.5) - 22 with MS and motor issues (age: 41.9 ± 10.6; 72.7% female; BMI: 25.6 ± 5.3; EDSS: 3.0) - 33 healthy controls (age: 34.8 ± 7.0; 63.6% female; BMI: 25.0 ± 5.0)	EDSS, GLTEQ questionnaire CMJ Temporal parameters: flight time, braking time, push-off time, FTCTR ratio Kinetics parameters: FZV (N/kg), max force (N/kg), negative/positive power (W/kg), BPIR ratio Performance: jump height (cm)	CMJ proved sensitive in detecting early neuromuscular and motor deficits in MS even at low EDSS scores. CMJ enables identification of eccentric and concentric muscle activity impairments.

Perez et al., 2007 (Effects of a Resistance Training Program in Multiple Sclerosis Spanish Patients: A Pilot Study)	To evaluate the effectiveness of a training program to improve physical fitness in people with MS.	- 24 individuals (9 male, 15 female) - Mean age: 44.4 ± 9.5 years; height: 161.4 cm; weight: 61.5 kg; EDSS: 1–6	Walking speed: zig-zag test over 9 m Arm mobility and orientation: clapping test (CT), dynamic flexibility test (DF) Arm power: medicine ball throw Leg power: vertical jump (VJ) in cm Core strength: abdominal test (AT), back muscles test (BM), leg lifts (LL), Kraus-Weber (KW) test Balance: Flamingo balance test (FB)	Strength training is beneficial for MS rehabilitation; simple strength exercises should be done at home. Individualized exercise programs are recommended.
Kedar et al., 2018 (Comparison Between Common Performance-Based Tests and Self-Reports of Physical Function in People With Multiple Sclerosis: Does Sex or Gender Matter?)	To assess how gender differences affect the relationship between performance tests and self-reported functional outcomes.	- 188 individuals (140 female, 48 male) - MS types: RRMS, SPMS, PPMS, PRMS, CIS; EDSS: 2 (IQR: 1–3)	Tests: modified Canadian aerobic test, grip strength, vertical jump, push-up, curl-up, EQ 5D-3L, 6-min walk test, 9-hole peg test Self-reported activities: rotation, partial sit-up, standing balance, sit-to-stand, forward bend, floor pick-up, trunk twist, bed-making, lifting, changing light bulb, recreational arm/shoulder activities, hand movement activities, transport	Balance tests were highly correlated with daily function. Therapy should be tailored to gender, as physical response may differ.
Kirkland et al., 2017 (Bipedal Hopping Reveals Evidence of Advanced Neuromuscular Aging Among People With Mild Multiple Sclerosis)	Use of bipedal hopping to detect advanced changes in people with mild MS.	- 13 with MS (EDSS ≤ 3.5) - 9 healthy (18–64 years) - 13 older adults (> 70 years)	Tool: instrumented walkway (Protokinetics pressure plate) Tests: bipedal hops, 25-ft walk test (T25FWT) Parameters: T25FWT time, hop length/width, speed, stance time %, mean pressure, CoP path efficiency, variability and asymmetry between legs Analysis: SPSS v21.0	T25FWT: MS group slower than healthy; other speed metrics not significantly different MS group resembled older adults. Linear regression: EDSS linked to hop length. Hop length may be useful for assessing lower limb function in mild MS.
Kirkland et al., 2018 (Bipedal hopping timed to a metronome to detect impairments in anticipatory motor control in people with mild multiple sclerosis)	To assess anticipatory motor control using bipedal hops timed to a metronome.	- 13 with MS (EDSS ≤ 3.5) - 9 healthy (18–64 years) - 13 older adults (> 70 years)	Tool: instrumented walkway (Protokinetics, Zeno electronic walkway) Tests: T25FWT, bipedal hops in rhythm (40 & 60 bpm over 8.5 m), MoCA test Parameters: T25FWT time; hops: step length/width, speed, hop time, stance %, mean pressure, CoP efficiency, variability and asymmetry Analysis: SPSS v 21.0	T25FWT: MS slower than healthy. Shorter hops, higher asymmetry and variability in pressure predicted poorer T25FWT time. Hop delay predicted MoCA only in MS. Shorter hops predicted worse T25FWT time. Bipedal metronome hops can detect mild motor and anticipatory control deficits in MS.
Geßner et al., 2025 (Experiences of People with Multiple Sclerosis in Sensor-Based Jump Assessment)	To evaluate the patient-reported experience of sensor-based jump assessment in pwMS To identify factors influencing these experiences, finding barriers and facilitators for clinical implementation	-175 pwMS (EDSS 0 – 5.5) - age 18-50 - ability to walk more > 500 m and to perform heel stand, toe stand, squats	Tool: portable single force plate (AMTI) Tests: 10 s hop test, CMJ, single-leg CMJ, Patient-Reported Experience Measures (PREM) Questionnaire Parameters: three-dimensional ground reaction forces (Fx, Fy, Fz) and force moments (Mx, My, Mz) Analysis: AccuPower Solutions (Version 1.5.4.2082)	A positive experience with the sensor-based jump assessment by pwMS Able to detect subtle motor changes. Sensor-based jump assessment, especially CMJ, offers a tech-driven way to support early rehabilitation.

Between 24 (Pérez et al., 2007) and 262 (Geßner, Hartmann, Trentzsch, et al., 2024) subjects with a diagnosis of multiple sclerosis participated in the studies. To capture individuals, suitable for early detection of motor deficits, the mean of the subjects on The Expanded Disability Status Scale (EDSS) was a maximum of 2.04 (Kirkland et al., 2017, 2018), indicating minimal disability (Demir, 2022). One study examined the effectiveness of an exercise programme (Pérez et al, 2007), one study looked for sex differences in the relation between tests of physical performance and self-reports about function (Mate et al., 2019), one study looked for sex, age and BMI differences in countermovement jump (Geßner, Hartmann, Trentzsch, et al., 2024), one study looked for -patient-reported experience of sensor-based jump assessment and identified factors influencing these experiences and searched for barriers and facilitators for clinical implementation (Geßner et al., 2025) and one study looked for a suitable measurement tool to detect motor fatigue (Geßner et al., 2021). Three studies searched for a suitable measurement tool to detect early motor deficits in multiple sclerosis (Geßner, Hartmann, Vágó, et al., 2024; Geßner et al., 2023; Kirkland et al., 2017), and one study examined the coordination and anticipation of hopping in a metronome (Kirkland et al., 2018).

Among the included tests, bipedal hops on 8.5-meter-long walkaway (Kirkland et al., 2017, 2018), countermovement jump (Geßner et al., 2023; Geßner, Hartmann, Trentzsch, et al., 2024, Geßner et al., 2025) and vertical jump (Mate et al., 2019; Pérez et al., 2007) were used twice each, while single-leg countermovement jump (Geßner, Hartmann, Vágó, et al, 2024, Geßner et al., 2025) and vertical 10-second hopping (Geßner et al., 2021, Geßner et al., 2025) were used once each. Except in cases where researchers measured only the height of the jump (Mate et al., 2019; Pérez et al., 2007), the measuring tools were force plates (Geßner, Hartmann, Trentzsch, et al., 2024; Geßner, Hartmann, Vágó, et al., 2024; Geßner et al., 2021, 2023, 2025; Kirkland et al., 2017, 2018).

The included measured parameters were time in the air (Geßner, Hartmann, Vágó, et al., 2024; Geßner et al., 2021, 2023) and time on the ground (Geßner, Hartmann, Trentzsch, et al., 2024; Geßner, Hartmann, Vágó, et al., 2024; Geßner et al., 2021, 2023; Kirkland et al, 2017, 2018), force (Geßner, Hartmann, Trentzsch, et al., 2024; Geßner, Hartmann, Vágó, et al., 2024; Geßner, Hartmann, Trentzsch, et al., 2023; Kirkland, et al., 2017, 2018), push-off and landing power (Geßner, Hartmann, Trentzsch, et al., 2024; Geßner, Hartmann, Vágó, et al, 2024; Geßner et al., 2021, 2023), push-off speed (Geßner, Hartmann, Trentzsch, et al., 2024; Geßner, Hartmann, Vágó, et al., 2024; Geßner et al., 2023), jump height (Geßner, Hartmann, Trentzsch, et al., 2024; Mate et al., 2019; Pérez et al, 2007), hop length (Kirkland et al., 2017, 2018) and hopping speed (Kirkland et al., 2017, 2018), three-dimensional ground reaction forces (Fx, Fy, Fz) and force moments (Mx, My, Mz) (Geßner et al., 2025). Kirkland et al. (2017) and Kirkland et al. (2018) additionally looked for inter-hop variability and asymmetries between the lower limbs. Details of the purpose, subjects, measurement tools and findings are summarised in Table.

Discussion

The analysis of the nine included studies demonstrates that jump- and hop-based assessments are effective tools for evaluating motor performance in people with multiple sclerosis (PwMS), even in individuals with minimal or no overt clinical disability. Countermovement jumps (CMJ), single-leg CMJ (SLCMJ) and bipedal hops were used to assess neuromuscular function, revealing consistent impairments in PwMS compared to healthy controls. CMJ was the most frequently used test and showed high sensitivity in detecting subtle motor deficits. Studies by Geßner et al. (2023, 2024) found that individuals with MS had shorter flight times, reduced jump height, and lower push-off force and power compared to healthy controls, even at low EDSS scores. This suggests that CMJ can reveal early neuromuscular dysfunction that may not be captured by standard clinical scales. Single-leg CMJ (Geßner et al., 2024) further enhanced the detection of inter-limb asymmetries, which were significantly greater in PwMS and positively correlated with EDSS scores. This finding underscores the potential of unilateral tests to identify compensatory strategies and early unilateral weakness. These findings are consistent with prior evidence of pronounced torque asymmetries in the lower limbs already present in individuals with early MS (Kalron et al., 2011).

Bipedal hopping tasks used by Kirkland et al. (2017, 2018) revealed that PwMS performed shorter hops with greater asymmetry and variability, indicating deficits in coordination and anticipatory motor control. Similar gait abnormalities, such as increased variability in step length and timing, have also been reported in early MS, further supporting the presence of coordination deficits even at minimal disability levels (Kalron et al., 2011). Interestingly, PwMS showed hopping patterns more similar to older adults than to age-matched healthy controls, suggesting early neuromuscular aging in MS. The study by Geßner et al. (2021) introduced a fatigue-focused protocol using a 10-second maximal jump test, demonstrating that PwMS performed fewer jumps with lower peak power and reactive strength index. These metrics were negatively associated with EDSS scores, highlighting the relevance of this tast for assessing motor fatigability.

In addition, one early study (Pérez et al., 2007) used vertical jump height to track improvements following resistance training and provided evidence that simple strength exercises may improve jump performance in PwMS. Geßner et al., (2025) found that pwMS had a positive experience with sensor-based jump assessment. It can detect subtle motor changes and, especially through CMJ, supports early, tech-driven rehabilitation. Overall, these findings indicate that jump and hop tests offer valuable, sensitive, and non-invasive means for identifying motor dysfunction in MS. Their ability to detect asymmetries, fatigue, and coordination deficits makes them especially relevant for early-stage patients and for evaluating intervention outcomes. The observed motor deficits align with findings from other studies reporting reduced lower limb performance compared to healthy controls (Ramari et al., 2019). A meta-analysis by Jørgensen et al. (2017) confirms decreased muscle strength, power, and explosive strength, with the latter particularly impaired during rapid concentric contractions. Furthermore, Jonsdottir et al. (2020) highlighted a significantly diminished push-off phase during gait, attributing this to adaptations in neuromuscular control of the ankle.

All studies included in the systematic review identified motor deficits in PwMS across various jump parameters. These deficits manifest as shorter flight time, lower jump height, impaired push-off phase, asymmetries between lower limbs, shorter and more variable jumps, reduced coordination, increased fatigability, and decreased explosive power. A key question remains regarding the underlying pathophysiological mechanisms responsible for these changes. It is most likely a combination of neural and musculoskeletal factors (Maffiuletti et al., 2016). At the central nervous system and crucial sensorimotor pathways, demyelination and axonal loss occur (Reich et al., 2018; Garg & Smith, 2015), simultaneously reducing the capacity for motor unit activation (Ng et al., 2004). Jumping requires complex functional performance involving multiple motor components such as muscle strength and power, coordination, flexibility, and effective neuromuscular control (Cormie et al., 2011). Coordination impairments, frequently presenting as dysdiadochokinesia and ataxia, significantly affect functional abilities, particularly gait, balance, and fall risk (Manto & Marmolino, 2009). Although asymmetries in lower limb muscle performance have been infrequently studied, evidence indicates that greater asymmetry correlates with increased postural sway in mediolateral and anteroposterior directions (Winter, 1995), negatively impacting coordinated movement and balance maintenance during daily activities (Mansfield et al., 2011). Despite the relatively limited number of studies examining jumping tasks in the context of MS, these functional assessments have proven valuable for detecting subtle motor impairments. Their advantage lies in evaluating multiple motor components simultaneously within a functional context, reflecting everyday movement patterns. Moreover, they allow efficient data collection with minimal burden on participants, which is especially important given the characteristic physical and cognitive fatigue in this population (Bakshi, 2023). Such assessments are useful not only for diagnostic evaluation but also for therapy planning, goal setting, and monitoring rehabilitation progress.

Limitations

The search was conducted in two databases. The literature review included a relatively small number of studies due to the inclusion and exclusion criteria. This limits the generalizability of their findings to the broader population of PwMS. The studies examined different forms of jumps and measured different parameters, which hinders direct comparisons between them. The range of EDSS scores among participants varied significantly across studies, indicating a heterogeneous overall sample and consequently diverse motor conditions. Additionally, the studies differed in the measurement tools used to assess jumps, complicating the establishment of a standardized protocol suitable for wider application and potential integration into clinical practice.

Conclusion

A systematic literature review reveals that various types of jumps can serve as an effective assessment tool for detecting subtle motor changes in PwMS. Jumps require a combination of different motor skills and thus represent functional movement. Assessments that evaluate movements resembling those in daily life can contribute to more effective detection of impairments that affect the performance of everyday activities and, consequently, quality of life. This kind of evaluation can enable more specific goal setting, more structured rehabilitation planning and better progress monitoring. There are relatively few studies examining different types of jumps in relation to MS, so further research in this area will be necessary for broader clinical application in the future.

***Declaration of interests*:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Bakshi, R., 2003. Fatigue associated with multiple sclerosis: diagnosis, impact and management. Mult Scler 9, 219–227. https://doi.org/10.1191/1352458503ms904oa

Compston, A., Coles, A., 2008. Multiple sclerosis. The Lancet 372, 1502–1517. https://doi.org/10.1016/S0140-6736(08)61620-7

Cormie, P., McGuigan, M.R., Newton, R.U., 2011. Developing Maximal Neuromuscular Power: Part 1 – Biological Basis of Maximal Power Production. Sports Medicine 41, 17–38. https://doi.org/10.2165/11537690-000000000-00000

Doshi, A., Chataway, J., 2016. Multiple sclerosis, a treatable disease. Clinical Medicine 16, s53–s59. https://doi.org/10.7861/clinmedicine.16-6-s53

Feinstein, A., Freeman, J., Lo, A.C., 2015. Treatment of progressive multiple sclerosis: what works, what does not, and what is needed. The Lancet Neurology 14, 194–207. https://doi.org/10.1016/S1474-4422(14)70231-5

Fisniku, L.K., Brex, P.A., Altmann, D.R., Miszkiel, K.A., Benton, C.E., Lanyon, R., Thompson, A.J., Miller, D.H., 2008. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. Brain 131, 808–817. https://doi.org/10.1093/brain/awm329

Garg, N., Smith, T.W., 2015. An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. Brain and Behavior 5, e00362. https://doi.org/10.1002/brb3.362

Geßner, A., Hartmann, M., Trentzsch, K., Stölzer-Hutsch, H., Schriefer, D., Ziemssen, T., 2024a. The Association of Age, Sex, and BMI on Lower Limb Neuromuscular and Muscle Mechanical Function in People with Multiple Sclerosis. Biomedicines 12, 971. https://doi.org/10.3390/biomedicines12050971

Geßner, A., Hartmann, M., Vágó, A., Trentzsch, K., Schriefer, D., Mehrholz, J., Ziemssen, T., 2024b. Sensitive Identification of Asymmetries and Neuromuscular Deficits in Lower Limb Function in Early Multiple Sclerosis. Neurorehabil Neural Repair 38, 570–581. https://doi.org/10.1177/15459683241245964

Geßner, A., Stölzer-Hutsch, H., Trentzsch, K., Schriefer, D., Ziemssen, T., 2023. Countermovement Jumps Detect Subtle Motor Deficits in People with Multiple Sclerosis below the Clinical Threshold. Biomedicines 11, 774. https://doi.org/10.3390/biomedicines11030774

Geßner, A., Trentzsch, K., Schriefer, D., Hartmann, M., Ziemssen, T., 2021. Quantification of motor fatigue by Jump analysis in people with Multiple Sclerosis. Gait & Posture 90, 67–68. https://doi.org/10.1016/j.gaitpost.2021.09.034

Geßner, A., Vágó, A., Stölzer-Hutsch, H., Schriefer, D., Hartmann, M., Trentzsch, K., Ziemssen, T., 2025. Experiences of People with Multiple Sclerosis in Sensor-Based Jump Assessment. Bioengineering 12, 610. https://doi.org/10.3390/bioengineering12060610

Hauser, S.L., Cree, B.A.C., 2020. Treatment of Multiple Sclerosis: A Review. The American Journal of Medicine 133, 1380-1390.e2. https://doi.org/10.1016/j.amjmed.2020.05.049

Jørgensen, M., Dalgas, U., Wens, I., Hvid, L., 2017. Muscle strength and power in persons with multiple sclerosis – A systematic review and meta-analysis. Journal of the Neurological Sciences 376, 225–241. https://doi.org/10.1016/j.jns.2017.03.022

Kirkland, M.C., Chen, A., Downer, M.B., Holloway, B.J., Wallack, E.M., Lockyer, E.J., Buckle, N.C.M., Abbott, C.L., Ploughman, M., 2018. Bipedal hopping timed to a metronome to detect impairments in anticipatory motor control in people with mild multiple sclerosis. Clinical Biomechanics 55, 45–52. https://doi.org/10.1016/j.clinbiomech.2018.04.009

Kirkland, M.C., Downer, M.B., Holloway, B.J., Wallack, E.M., Lockyer, E.J., Buckle, N.C.M., Abbott, C.L., Ploughman, M., 2017. Bipedal Hopping Reveals Evidence of Advanced Neuromuscular Aging Among People With Mild Multiple Sclerosis. Journal of Motor Behavior 49, 505–513. https://doi.org/10.1080/00222895.2016.1241750

Klavora, P., 2000. Vertical-jump Tests: A Critical Review. J Strength Cond 22, 70. https://doi.org/10.1519/1533-4295(2000)022<0070:VJTACR>2.0.CO;2

Klineova, S., Lublin, F.D., 2018. Clinical Course of Multiple Sclerosis. Cold Spring Harb Perspect Med 8, a028928. https://doi.org/10.1101/cshperspect.a028928

Maffiuletti, N.A., Aagaard, P., Blazevich, A.J., Folland, J., Tillin, N., Duchateau, J., 2016. Rate of force development: physiological and methodological considerations. Eur J Appl Physiol 116, 1091–1116. https://doi.org/10.1007/s00421-016-3346-6

Mansfield, A., Danells, C.J., Inness, E., Mochizuki, G., McIlroy, W.E., 2011. Between-limb synchronization for control of standing balance in individuals with stroke. Clinical Biomechanics 26, 312–317. https://doi.org/10.1016/j.clinbiomech.2010.10.001

Manto, M., Marmolino, D., 2009. Cerebellar ataxias. Current Opinion in Neurology 22, 419–429. https://doi.org/10.1097/WCO.0b013e32832b9897

Mate, K.Kv., Kuspinar, A., Ahmed, S., Mayo, N.E., 2019. Comparison Between Common Performance-Based Tests and Self-Reports of Physical Function in People With Multiple Sclerosis: Does Sex or Gender Matter? Archives of Physical Medicine and Rehabilitation 100, 865-873.e5. https://doi.org/10.1016/j.apmr.2018.10.009

Milo, R., Miller, A., 2014. Revised diagnostic criteria of multiple sclerosis. Autoimmunity Reviews 13, 518–524. https://doi.org/10.1016/j.autrev.2014.01.012

Pérez, C.A., Sánchez, V.M., Teixeira, F.D.S., Fernández, J.A.D.P., 2007. Effects of a Resistance Training Program in Multiple Sclerosis Spanish Patients: A Pilot Study. Journal of Sport Rehabilitation 16, 143–153. https://doi.org/10.1123/jsr.16.2.143

Ramari, C., Hvid, L.G., David, A.C.D., Dalgas, U., 2020. The importance of lower-extremity muscle strength for lower-limb functional capacity in multiple sclerosis: Systematic review. Annals of Physical and Rehabilitation Medicine 63, 123–137. https://doi.org/10.1016/j.rehab.2019.11.005

Reich, D.S., Lucchinetti, C.F., Calabresi, P.A., 2018. Multiple Sclerosis. N Engl J Med 378, 169–180. https://doi.org/10.1056/NEJMra1401483

Reina, R., Iturricastillo, A., Sabido, R., Campayo-Piernas, M., Yanci, J., 2018. Vertical and Horizontal Jump Capacity in International Cerebral Palsy Football Players. International Journal of Sports Physiology and Performance 13, 597–603. https://doi.org/10.1123/ijspp.2017-0321

Rovaris, M., Confavreux, C., Furlan, R., Kappos, L., Comi, G., Filippi, M., 2006. Secondary progressive multiple sclerosis: current knowledge and future challenges. The Lancet Neurology 5, 343–354. https://doi.org/10.1016/S1474-4422(06)70410-0

Santos, C.A.F., Amirato, G.R., Jacinto, A.F., Pedrosa, A.V., Caldo-Silva, A., Sampaio, A.R., Pimenta, N., Santos, J.M.B., Pochini, A., Bachi, A.L.L., 2022. Vertical Jump Tests: A Safe Instrument to Improve the Accuracy of the Functional Capacity Assessment in Robust Older Women. Healthcare 10, 323. https://doi.org/10.3390/healthcare10020323

Soelberg Sorensen, P., Giovannoni, G., Montalban, X., Thalheim, C., Zaratin, P., Comi, G., 2019. The Multiple Sclerosis Care Unit. Mult Scler 25, 627–636. https://doi.org/10.1177/1352458518807082

Uygunoglu, U., Kantarci, O., Siva, A., 2016. Integrated multidisciplinary clinics should be the gold standard in managing progressive MS – YES. Mult Scler 22, 1126–1128. https://doi.org/10.1177/1352458516650526

Winter, D., 1995. Human balance and posture control during standing and walking. Gait & Posture 3, 193–214. https://doi.org/10.1016/0966-6362(96)82849-9