©Journal of Sports Science and Medicine (2026) **25**, 84-111
http://www.jssm.org DOI: https://doi.org/10.52082/jssm.2026.84

Review article

Optimizing Agility Training in Team Sport Players—The Role of Perception-Action Coupling: A Systematic Review with Multi-Level Meta-Analysis

Zhiwei Zhao [1], Yuhang Liu [2] and Kai Xu [3,4]✉
[1] School of Physical Education and Sports, Central China Normal University, Wuhan, China; [2] School of Physical Education, Wuhan Sports University, Wuhan, China; [3] School of Athletic Performance, Shanghai University of Sport, Shanghai, China; [4] School of Medical and Health Sciences, Centre for Human Performance, Edith Cowan University, Joondalup, Australia

Abstract
Agility, characterized by rapid whole-body movement in response to external stimuli, is a key performance determinant in team sports owing to its reliance on perception-action coupling. Despite its importance, existing evidence on training effects is disproportionately focused on change-of-direction ability, overlooking the perceptual-cognitive demands of true agility. This systematic review therefore sought to quantify the effects of agility-specific training, examine moderators and establish dose-response model in team-sport players, while accounting for perception-action coupling. Following PRISMA 2020 guidelines, systematic searches were conducted in Web of Science, Scopus and PubMed for peer-reviewed English-language studies. Risk of bias was appraised using a modified Cochrane Collaboration's tool and study quality was evaluated via a tailored PEDro scale. Effect size (ES) was calculated using Hedge's g, and dependencies among multiple ESs within studies were addressed through three-level meta-analysis. Subgroup analyses and both linear and non-linear meta-regressions were performed to examine potential moderators and establish dose-response models. Twenty-eight studies met the inclusion criteria, of which 26 contributed 53 ESs to quantitative analyses. Agility training produced a large improvement in reactive agility test (RAT) performance (ES = 0.65, $p < 0.01$), a moderate improvement in pre-planned agility test (PAT) performance (ES = 0.55, $p < 0.01$), and a non-significant moderate effect in reaction test (RT) outcomes (ES = 0.52, $p > 0.05$). Training effects were moderated by participants' characteristics, with junior athletes (ES = 0.77, $p < 0.05$) and national level athletes (ES = 0.80, $p < 0.05$) demonstrating greater responses. However, female and mix-gender samples were underrepresented, and the evidence base was dominated by studies in soccer and basketball. Dose-response modelling revealed a curvilinear relationship between training duration and RAT (QM = 11.64, $p < 0.01$), peaking at 7-8 weeks and a positive linear association between training frequency and RAT ($\beta = 0.172$, $p < 0.01$). No significant relationship was observed between session time and RAT ($p > 0.05$), although most positive ESs clustered around 20-25 minutes per training session. Agility training exerts a large overall effect on RAT performance in team-sport players, with outcomes moderated by age and training status. Interventions of 7 - 8 weeks delivered at higher frequency (> 3 times/week) with 20 - 25 minutes session duration are frequently associated with favorable adaptions. These recommendations, however, should be interpreted cautiously given the moderate-to-low certainty of evidence, high within-study variability, dominance of soccer and basketball samples and potential risk of publication bias.

Key words: Agility, Team-sport, Perception-action coupling, Meta-analysis, Dose-response.

Introduction

Agility, defined as a rapid whole-body movement involving changes in velocity or direction in response to a stimulus (Sheppard and Young, 2006), is a key determinant of team-sport performance (Chaalali et al., 2016; Lucia et al., 2023). In contrast to change-of-direction (COD), which reflects pre-planned locomotor tasks, agility inherently requires perception-action coupling (Young et al., 2021), integrating perceptual-cognitive processes such as anticipation and decision-making with motor execution. This distinction underscores why agility cannot be fully captured by COD measures, although COD remains an important foundational component of agility performance (Sheppard and Young, 2006).

The term "COD", "agility" and "reactive agility" were used interchangeably in agility reviews. However, according to the definition of agility, using the term "reactive" is redundant (Young et al., 2015). Moreover, studies were either evaluating COD ability (Asadi et al., 2017; Forster et al., 2022; Sun et al., 2025), or validity of agility tests (Morral-Yepes et al., 2022), incorporating outcomes such as T-test, Illinois test or pro-agility. While trainability of agility was qualitatively discussed (Thieschäfer and Büsch, 2022), quantitative evidence of aggregated agility training outcomes in team sports remained absent. As suggested by Young et al., (2021), findings on COD should not be extrapolated to on-field agility due to distinct perception-action coupling differences.

Furthermore, these reviews often did not account for hierarchical structure of data (Asadi et al., 2017; Sun et al., 2025), such as multiple effect sizes nested within studies, which can bias pooled estimates (Van den Noortgate et al., 2013). However, extracting only one effect size per study can also yield biased estimates. In addition, existing reviews have drawn heavily from soccer (Sun et al., 2025) and basketball (Zhang et al., 2025), with relatively fewer studies addressing other team sports, thereby limiting generalizability. A further limitation concerns the absence of dose-response modelling using linear and non-linear regression methods to identify optimal training parameters such as duration, frequency and session length. This is particularly important for practitioners, as understanding dose-response relationships enables coaches to prescribe training that maximizes adaptation while avoid overtraining. Other Factors such as participant characteristics (e.g.

Received: 27 August 2025 / Accepted: 20 November 2025 / Published (online): 01 March 2026

age, sex and competitive level) and stimulus type (e.g. light, video and human) have not been systematically explored in team sports.

Therefore, the objectives of this systematic review were to :

1) Quantify the effects of agility-specific training on reactive agility performance among team-sport players, with particular emphasis on perception-action coupling.
2) Examine moderating factors including participants' characteristics, training modalities and stimulus type, to provide tailored insights into athlete responsiveness.
3) Establish dose-response models to identify potential thresholds of training duration, frequency and session length that optimize agility development.

By addressing these objectives, this review aims to advance the evidence base on agility training and provide practical guidance for coaches and practitioners in designing valid training interventions.

Methods

Registration of systematic protocol

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) 2020 guidelines (Page et al., 2021) and was registered in PROSPERO (CRD420251060149). Registration occurred after article retrieval but before data extraction and analysis. Although this timing may increase the risk of selective reporting, analyses were conducted according to a pre-specified protocol to mitigate bias. Deviation from registered protocol was given in Supplementary Table S13.

Search strategy

Systematic searches were conducted on 20th April 2025 in Web of Science (WOS), Scopus and PubMed, with updates on 4th May 2025. Boolean search strings were developed iteratively from preliminary scoping, where key articles were identified (Sheppard and Young, 2006; Young et al., 2021; Young et al., 2015). Complete search strategies are provided in full for all databases in Supplementary Table S1. Identical syntax was adapted across platforms to ensure comparability of retrieval. In addition, backward and forward citation tracking was performed in Google Scholar. Records were imported into Endnote (Version 21; Clarivate, Philadelphia, PA, USA) and deduplicated prior to screening. A sample search strategy in WOS is shown below:

#1: ALL = (agility OR quickness OR change-of-direction OR reactive OR plyometric OR neuromuscular OR stimulu* OR SSG OR small-sided-game OR dual-task OR cognitive OR perceptual) AND (training OR intervention OR program*)

#2: ALL = (athlete OR player OR recreational OR elite OR professional OR trained) AND (soccer OR football OR rugby OR basketball OR volleyball OR team-sport)

#3: #1 AND #2

Study selection and eligibility criteria

Initial study retrieval was conducted by one author (Z.Z.). After identification and screening of articles, during the eligibility phase, two authors (Y.L. and Z.Z) reviewed the full texts based on the exclusion and inclusion criteria (Supplementary Table S2 and S2-1). When mutual agreement on studies cannot be achieved, decision was made in consultation with a third author (K.X.). All processes were conducted in Endnote 21.

Data extraction

Data were extracted by two authors (Z.Z. and Y.L.) independently and validated by the third author (K.X.) after cross checked by the first two authors. Data extracted were organized into Excel tables (Microsoft Corporation, Redmond, WA, USA) for the following information: 1) study authors and year; 2) number of participants in each group, participants' age, sex , training status, and types of sport; 3) intervention duration (weeks), frequency (times/week), session time (minutes/session), and types of stimulus; 4) agility performance outcomes including reactive agility test (RAT), pre-planned agility test (PAT), and reaction test (RT) using means and standard deviations (SD) of pre- and post- agility training in the experimental and control groups. When data were missing, the corresponding authors were contacted for raw data, if there was no response the study was excluded. Specifically, one study was excluded for no response (Martin-Niedecken et al., 2023).

Coding of studies

Outcomes were classified according to Sheppard & Young, (2006), agility measurement was coded into simple agility (tests with no temporal or spatial uncertainty), temporal agility (tests with temporal uncertainty but pre-planned movement), spatial agility (tests with spatial uncertainty, but timing of movement is pre-planned), and universal agility in which the test has both spatial and temporal uncertainty, restricting athletes from anticipating where or when to react. However, simple agility and temporal agility do not require movement change and perceptual response at the same time. Like shuttle run (simple agility) and 100m race (temporal agility). Therefore, these two were not included for RAT. Spatial agility and universal agility tests were included in RAT since they require external stimuli and movement simultaneously. In addition, we included the outcomes from PAT and RT as secondary outcome.

Age was grouped based on Cobley's classification (Cobley et al., 2009), as it reflects key developmental stages like pre-puberty, early puberty, late puberty/adolescence and adulthood. Participants were categorized into child (< 11 years), junior (11 - 14 years), adolescent (15-18 years) and senior (>18 years).

Sex was coded into three categories: 1) male, where studies recruited male participants only; 2) female, where studies recruited female participants only and 3) mix, where studies recruited both male and female participants.

Training status classification was made based on McKay et al., (2022). Tier 1 and Tier 2 classification were not included since they are not identified as athletes. Therefore, we categorized athletes as: 1) developmental: local-level representation, identified with a specific team-sport and training with a purpose to compete; 2) national level: team-sport players competing in national and/or state leagues/tournaments, NCAA second or third division athletes; 3) International level: team-sport players on a national team and NCAA first division athletes. For those studies that did not mention competitive level, but identified to a specific sport and trained, we categorized those participants into developmental level (Arede et al., 2021; Dunton et al., 2020; Engelbrecht et al., 2016; Hassan et al., 2022; Trecroci et al., 2016; T. Zwierko et al., 2024).

Training duration was categorized into short term (< 6 weeks) and long term (≥ 6 weeks) using the median-split method (Ramirez-Campillo et al., 2022). Training frequency and duration per training session were also coded using the same method into low frequency (< 3 times/week) and high frequency (≥ 3 times/ week) and short session (< 25 minutes/session) and long session (≥ 25 minutes/session). Training load was coded into training volume and training intensity respectively. For training volume, data was coded based on the total sets/repetitions performed in running-based and SSG-based exercises respectively. While training intensity was coded according to heart rate. Although most studies (approximately 69.23% of the total sample) reported sets or repetitions performed (Supplementary Table S4), the reporting varied greatly in unit, such as number of COD maneuvers (Chaouachi et al., 2014; Young and Rogers, 2014), soccer passes (Dunton et al., 2020; Hicheur et al., 2020), or sets duration (Vogt et al., 2018; Zago et al., 2016). The lack of standardized training volume reporting making it impractical to synthesize data. Also, only three studies reported training intensity quantitively using heart rate or rating of perceived exertion (Chaouachi et al., 2014; Friebe et al., 2024; Vogt et al., 2018). Therefore, coding of training load data was not employed.

Stimuli type was coded into visual, visual-auditory and human stimuli modified from (Paul et al., 2016). We used these intercepts to better conclude various stimulus types as there are some verbal or sound instructions during agility training.

Quality evaluation and risk of bias evaluation

Risk of bias was assessed using a modified version of the Cochrane Collaboration tool (Higgins et al., 2011). Blinding of participants is critical when results are highly subjective, such as pain (Higgins et al., 2011). However, in current research, results were mostly assessed subjectively. Therefore, we excluded this domain and incorporated additional items on test familiarization, measurement reliability and sample size estimation, to assure a holistic assessment of risk of bias.

For familiarization of tests, as recommended by Paul et al., (2016), athletes should be appropriately familiarized with tests before commencing data collection. Therefore, studies were categorized into "low risk", if study reported that participants were well familiarized with the test protocols or achieved stable performance with trivial detectable change before data collection; "unclear" if familiarization was allowed but degree of familiarity was unknown; "high risk" if no familiarization was allowed. For measurement bias, tests for agility are often modified to reflect on-field performance. Therefore, studies were categorized into "low risk" if two or more measurement reliability were reported (e.g. Intraclass correlation coefficient (ICC) and coefficient variation (CV)) (Currell and Jeukendrup, 2008); "unclear" if only one reliability was reported and "high risk" if no reliability was reported. For sample size bias, studies were categorized into "low risk", if they reported prior sample size estimation with appropriate parameters (e.g. G*power using accurate estimates including α probability, power and effect size); "unclear" if they didn't report sample size estimation with appropriate parameters or software used; "high risk" if they didn't report sample size estimation and included less than 10 participants.

Study quality was evaluated using a modified PEDro tailored for strength and conditioning (Brughelli et al., 2008), to assess the completeness of research methodology process and reporting. The scale includes 10 items (total score ranging from 0-20) with each item rated as: 0 = clearly no; 1 = maybe; and 2 = clearly yes. The items included: 1) inclusion criteria were clearly stated; 2) subjects were randomly allocated to groups; 3) intervention was clearly defined; 4) groups were tested for similarity at baseline; 5) use of a control group; 6) outcome variables were clearly defined; 7) assessments were practically useful; 8) duration of intervention practically useful; 9) between-group statistical analysis appropriate; 10) point measures of variability were provided. The modified version of quality and risk of bias assessment tools were discussed and agreed by the three authors (Z.Z., Y.L. and K.X.).

Certainty of evidence was graded using GRADE (Grading of Recommendations Assessment, Development, and Evaluation) framework (Guyatt et al., 2008). The system is used to rate the quality of evidence in reviews and meta-analyses, offering structured assessments of our confidence that the true effect is near the estimated effect. GRADE categorizes the quality of evidence into four tiers: high, moderate, low and very low based on risk of bias, inconsistency, indirectness, imprecision and publication bias.

The assessments were then carried out by the three authors independently, discrepancies after evaluation were discussed and resolved.

Statistical analysis

Data synthesis and effect measures

In this study, between-group comparisons were made to calculate effect sizes (ES). Given the typically small sample size in sport science, Hedge's g ES was utilized for its bias correction using means and SDs from experimental and control group (Hedges and Olkin, 2014). Specific formulas used are as follows:

$$ES = \frac{M_1 - M_2}{SD_{pooled}} \times (1 - \frac{3}{4(n_1 + n_2) - 9})$$

$$SD_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1{}^2 + (n_2 - 1)SD_2{}^2}{n_1 + n_2 - 2}}$$

Where M stands for mean, SD stands for standard deviation and n stands for sample size, and the subscripts 1 and 2 represent experimental and control group, respectively.

Additionally, when standard errors (SE) were presented in studies, they were converted into SDs using $SD = SE \times \sqrt{n}$. Magnitude of ESs were categorized into small (<0.2), medium (0.2-0.5), and large (>0.5) (Cohen, 2013). I^2 and Q tests were utilized to evaluate heterogeneity, with I^2 of 25, 50 and 75% indicating low, moderate and high heterogeneity, respectively. The Q tests were considered significant at $p < 0.1$ (Xu et al., 2025). The performance evaluated in time units was multiplied by -1 to calculate the effect size in the analysis, in which the improvement of performance ES was adjusted to reflect a positive enhancement.

Three-level meta-analysis

Agility research often reported various performance outcomes (Chaouachi et al., 2014; Friebe et al., 2024), and assessments at multiple time points (Born et al., 2016; Zago et al., 2016). Different ESs within single study are correlated and including them simultaneously would violate the independence assumption of traditional meta-analysis, which leads to overlapped information and in turn potentially biased outcomes (Kadlec et al., 2023). However, including only one ES for each study may seem too conservative to reflect the true estimates (Van den Noortgate et al., 2013). Therefore, we utilized a three-level meta-analysis model following the methods of Xu et al., (2025), with analysis carried out using the open-source code provided by the same authors to account for dependence between effect sizes. It is assumed that multiple assessments and comparisons from the same study were nested within studies, therefore we decomposed the variance of the observed effect size into sampling variance, within-study variance (level 2) and between study variance (level 3) to account for within-study correlations (Cheung, 2019). The model was used to fit the hierarchical structure of the data (e.g. effect sizes are nested within studies), so that multiple ESs from each study add up to the statistical power and provide estimates closer to the true ES (Assink and Wibbelink, 2016). The following equation will be used in three-level model:

$$\widehat{\theta_{ij}} = \mu + \zeta_{(2)ij} + \zeta_{(3)ij} + \epsilon_{ij}$$

Here, $\widehat{\theta_{ij}}$ represents the estimated true effect size for i-th effect within the j-th study. The term μ denotes the overall pooled effect size. The random components $\zeta_{(2)ij}$ and $\zeta_{(3)ij}$ stand for within-study and between study variance respectively, while ϵ_{ij} represents the sampling variance associated with each effect size. Accordingly, the observed effect size can be expressed as the sum of the pooled effect size, the variance components at each level and the sampling error.

Cluster-robust variance estimation method with small-sample adjustment to adjust the within-study standard errors for correlations between ESs was utilized for sensitivity analysis (Hedges et al., 2010). Within-study correlation was set at 0.6 to reflect moderate dependence, consistent with previous multilevel syntheses (Xu et al., 2025), sensitivity analyses at 0.4 and 0.8 confirmed stability of results (Supplementary Table S10). In addition, results did not change after performing small sample size adjustment. The analyses were carried out using the restricted maximum-likelihood (REML) method as recommended by Harrer et al., (2021), and cross-verified using maximum-likelihood (ML). Three-level meta-analysis model was selected based on goodness-of-fit information (e.g. Akaike information criterion). Three-level meta-analysis was conducted using the *metafor* package in R (version 4.5.0; R core team, Vienna, Austria) (Viechtbauer, 2010).

Subgroup analysis and meta-regression

Subgroup analyses were conducted based on participants' characteristics (i.e. age, sex, training status and sport type) and training modalities (i.e. training period, frequency, time per session, and stimulus type) to explore the moderating effect on agility performance. Subgroups analyses were conducted if the number of ESs within a subgroup was greater than 10 ($k \geq 10$) to avoid biased outcomes. Subgroups analysis results will be kept descriptive and detailed analyses of moderators will be explored using dose-response modelling (meta-regression).

Linear and nonlinear relationships of different moderators and performance outcomes were explored using meta-regression models including simple linear, cubic polynomial, restricted cubic spline, natural cubic spline and thin plate spline models (Xu et al., 2025). The fitting model was than selected based on goodness-of-fit and practical considerations. Linear model reflected better fit (e.g. lower Akaike information criterion) for training frequency and RAT performance, but results can be misleading due to the complex nature of agility in team-sport. Therefore, non-linear model with a better fit was conducted to explore if potential non-linear relationship exists. Except for training frequency, non-linear models all have the better fit than linear model and can better reflect the complex nature of agility intervention. Therefore, non-linear models were used for meta-regression analyses.

Publication bias and sensitivity analysis

To assess potential publication bias or small-study effects in the presence of dependent ESs, we employed a multilevel extension of Egger's test, which models asymmetry in a three-level framework (Fernández-Castilla et al., 2021). The model extends the traditional two-level Egger's regression to account for sampling error (level 1), within-study variance (level 2) and between study variance (level 3), aligning with three-level structure in the current meta-analysis. And it has been proved to have adequate false-positive error control and a slight power advantage over robust variance estimation alternatives (Rodgers and Pustejovsky, 2021). Subsequently, Hat, Cook's distance and studentized residuals were employed for leverage, outliers, and influential case diagnosis at level 2 and level 3 respectively. Studies were flagged if their Hat and Cook's distance values were greater than three times their respective mean, which were further confirmed if they had an absolute Studentized residual value greater than 3. Then the three-level model was repeated after excluding the influential studies. Additionally, each study was excluded at level 2 and level 3 to determine the impact of individual study on overall effect size (Harrer et al., 2021).

Baseline difference for all interventions were calculated using unpaired t-test to ensure stability. Subsequently, sensitivity analysis was conducted by excluding significantly different baseline data. If no significant change in results was observed, the models were retained.

Results

Search results

The initial article retrieval yielded 11,840 records (PubMed = 1490, WOS = 5918, Scopus = 4432). After removal of duplicates, 8362 titles/abstracts were screened. After which, 361 records were identified as potential articles for full-text screening based on our pre-specified exclusion criteria. Reasons for full-text exclusion included:

1) Isolated physical or perceptual training (n = 173);
2) Absence of control group or non-randomized design (n = 28);
3) Lack of reactive agility test results (n = 108);
4) Population other than team-sport players (n = 23);
5) Full-text not available (n = 5)

Therefore, a final 25 articles met the inclusion criteria. In addition, a snowballing method was used for secondary search, which identified 3 additional articles. Subsequently, per pre-specified criteria, two studies were excluded due to influential outlier status, leaving 26 studies contributing 53 ESs to quantitative analysis (Figure 1). Thus, 28 studies were included for systematic review, while 26 studies entered meta-analysis.

Study characteristics

A summary of included studies can be found in (Supplementary Table S3), a total of 755 participants aged between 10.31 and 35 years old were recruited. Among them, 2 studies (Lee et al., 2024; Mancini et al., 2024) recruited female subjects (43 subjects), 6 studies (Arede et al., 2021; Ehmann et al., 2022; Lucia et al., 2023; Lucia et al., 2024; Zwierko et al., 2023; M. Zwierko et al., 2024) recruited both male and female subjects (209 subjects) and 3 studies (Dunton et al., 2020; Hassan et al., 2022; Trecroci et al., 2016) didn't report the gender of the subjects (75 subjects). Participants were all team-sport players, 14 studies recruited developmental athletes, 8 studies recruited national level athletes, and 6 studies recruited international level athletes. Training duration ranged from 1 to 22 weeks, training frequency ranged from 1 to 6 times per week, and time per session ranged from 12 to 60 minutes. Number of studies chose to use visual, visual-auditory, and human stimuli was 14, 5 and 7 respectively. Two studies (Hicheur et al., 2020; Lucia et al., 2024) utilized both visual and visual-auditory stimuli in their research. All included studies reported at least one indicator of reactive agility and/or open-skill performance.

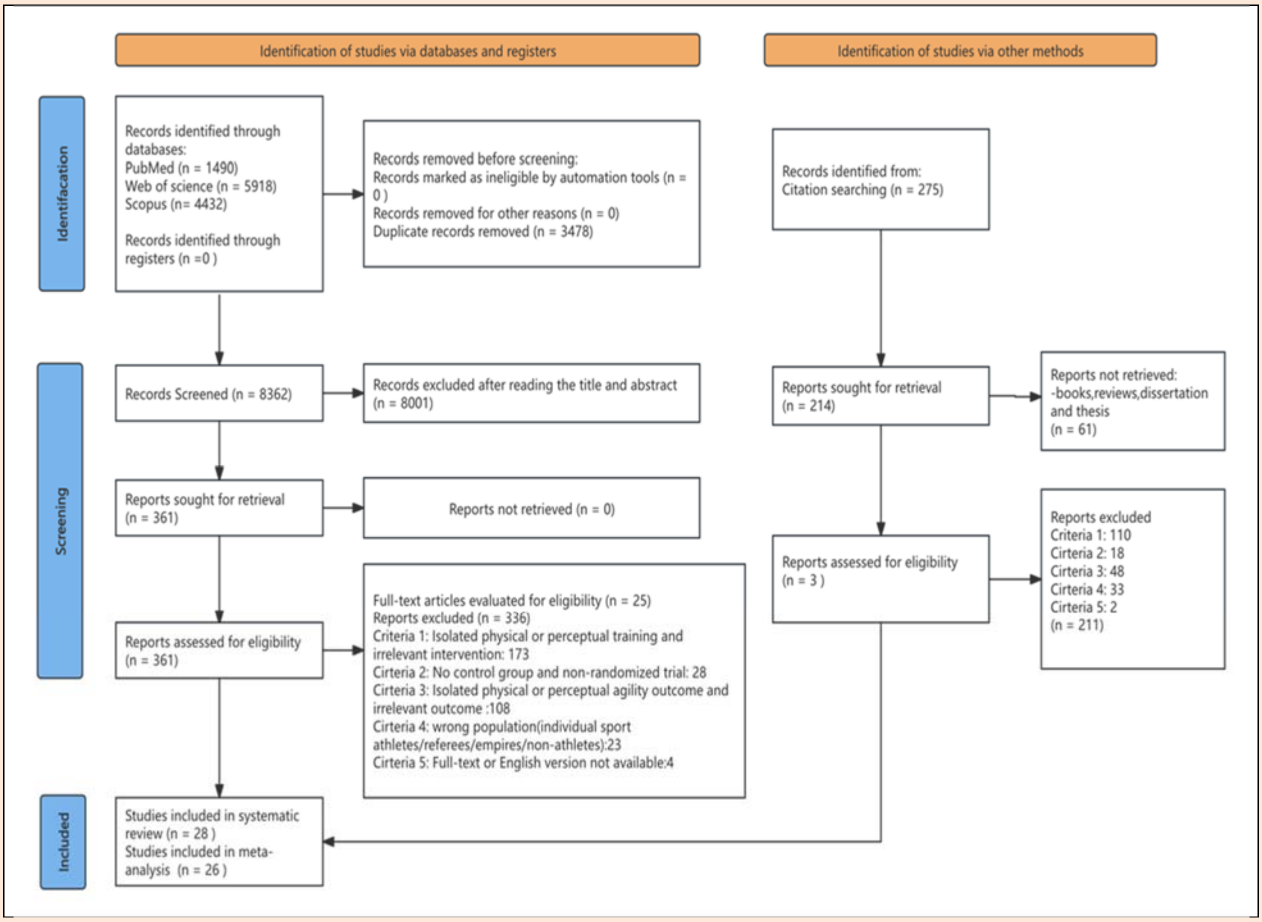

Figure 1. Flow chart of literature identification.

Risk of bias and quality assessment

All included studies were randomized controlled trials and mentioned about randomization (Figure 2). However, only 6 studies (Friebe et al., 2024; Gonzalo-Skok et al., 2023; Lucia et al., 2023; Mancini et al., 2024; Serpell et al., 2011; Trecroci et al., 2016) detailed the method of randomization, which were categorized as “low risk” for selection bias, while others were labeled as “unclear”. One study reported that assessor was blinded for condition assignment, which was rated as “low risk” for detection bias. And one study stating that assessors were not blinded was rated as “high risk”, while the rest were rated as “unclear”. Five studies that experienced attrition of subjects with no significant impact on results were rated as “unclear” for attrition bias, while others were rated as “low risk”. Nine studies clearly reported prespecified outcomes and were therefore rated as “low risk” for reporting bias, while others were rated as “unclear”. Three studies didn’t allow familiarization before data collection, which were rated as “high risk” for familiarization bias. (Born et al., 2016) was coded into “low risk” for reporting participants were well familiarized with test protocols, while the remaining studies were rated as “unclear”. Thirteen studies reporting ICC and CV for reliability were rated as “low risk” for measurement bias, and 3 studies reporting none were rated as “high risk”, while the rest were “unclear”. Six studies reported sample size calculation with appropriate parameters reported and were therefore categorized into “low risk”, while 3 studies with small sample size were categorized as “high risk”. The rest were categorized into “unclear”.

The modified PEDro score of included studies ranged from 16 to 20 (Supplementary Table S5), with an average score of 17.43. Although the overall risk of bias suggests “unclear” to “high risk” for the included studies, the PEDro score indicates that the methodological implementation was relatively complete.

Publication bias and Certainty assessment

Publication bias was evaluated using the multilevel Egger’s regression test along with contour-enhanced funnel plot, which revealed evidence of asymmetry for RAT ($p < 0.01$), while no evidence of asymmetry was detected for PAT and RT ($p > 0.1$). Reflecting on funnel plots (Figure 3), it can be inferred that PAT and RT were slightly asymmetric, while RAT was highly asymmetric. However, unlike typical risk of publication bias where studies with significant or positive results are more likely to be published.

Risk of Bias

Study	D1	D2	D3	D4	D5	D6	D7
Serpell et al. 2011	✓	!	✓	!	!	!	!
Chaouachi et al. 2014	!	!	✓	!	!	✓	!
Young & Rogers 2014	!	!	✓	!	!	!	!
Born et al. 2016	!	!	✓	✓	✓	✓	!
Chaalali et al. 2016	!	!	✓	!	!	✓	!
Engelbrecht et al. 2016	!	!	✓	✓	✗	✓	!
Trecroci et al. 2016	✓	!	!	✓	!	✓	!
Zago et al. 2016	!	!	✓	✓	!	!	✗
Vogt et al. 2018	!	!	✓	!	!	✓	!
Zouhal et al. 2019	!	!	✓	!	!	✓	!
Dunton et al. 2020	!	!	!	✓	!	!	!
Hicheur et al. 2020	!	!	!	!	!	✓	!
Arede et al. 2021	!	!	✓	!	!	!	!
Ehmann et al. 2022	!	✓	!	!	!	!	!
Hassan et al. 2022	!	!	✓	!	!	!	!
Bekris et al. 2023	!	!	✓	!	!	!	✗
Gonzalo-Skok et al. 2023	✓	!	✓	!	!	!	✓
Lucia et al. 2023	✓	!	!	!	!	✓	✓
Zwierko et al. 2023	!	!	✓	✓	!	!	✓
Friebe et al. 2024	✓	✗	✓	!	!	✓	✓
Lee et al. 2024	!	!	✓	!	!	✗	✗
Lucia et al. 2024	!	!	✓	!	✗	✗	✓
Mancini et al. 2024	✓	!	✓	✓	!	✓	!
Steff et al. 2024	!	!	✓	!	✗	✗	!
Zwierko et al., 2024 (a	!	!	✓	✓	!	!	!
Zwierko et al., 2024 (b	!	!	✓	!	!	✓	✓
Klatt et al. 2025	!	!	✓	✓	!	✓	!
Neag et al. 2025	!	!	✓	!	!	!	!

Domins:

D1:Random sequence generation (Selection bias)
D2: Blinding of outcome assessment (Detection bias)
D3: Incomplete outcome data (Attrition bias)
D4: Selective reporting (Reporting bias)
D5: Familiarization of tests (Familiarization bias)
D6: Reporting of reliability (Measurement bias)
D7: Reporting of sample size (Sample size bias)

Jugement	
Low risk	✓
Unclear	!
High risk	✗

Figure 2. Risk of bias assessment for all included studies.

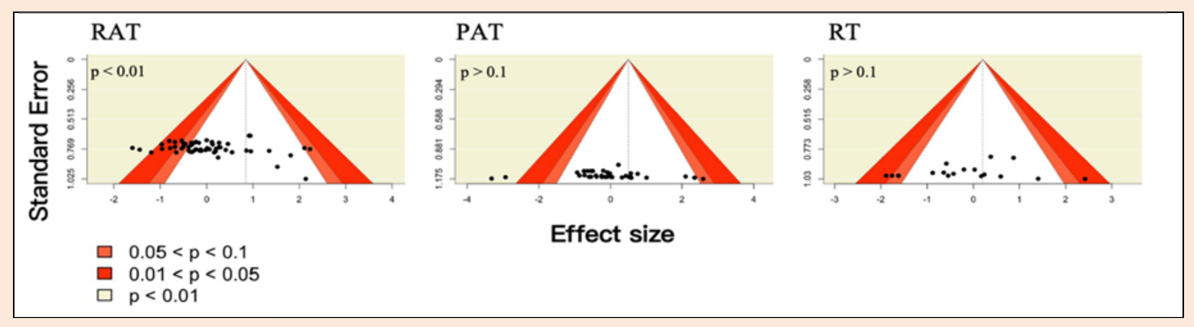

Figure 3. Risk of Publication bias at level 2 and 3. RAT reactive agility test; PAT pre-planned agility test; RT reaction test.

The RAT results included many non-significant or small-effect results. Therefore, the significant publication bias may stem from selective reporting of results, as most studies were unclear regarding reporting bias. The extended version of Egger's test detects cumulative publications bias of each level, making it more sensitive to outcome reporting bias hidden in multiple outcomes than traditional model (Citkowicz and Vevea, 2017). Therefore, highly influential studies would significantly inflate the overall ESs.

Further, sensitivity analyses were conducted using leave-one-out, cook's distance and studentized residual value. The leave-one-out analysis flagged (Hassan et al., 2022) as influential study, changing the overall ES from 0.79 to 0.70. Cook's distance identified Hassan et al., (2022) and Steff et al., (2024) as outliers with absolute cook's distance value exceeding three times the mean value. After removing the outliers, the ES changed from 0.79 to 0.65. Additionally, outlier residual analysis flagged Hassan et al. (2022) with value greater than 4. Therefore, the two studies were excluded from further analysis to ensure stability. After excluding the influential studies, baseline difference tests were conducted with significantly different data being flagged for each outcome. The fitted models were then cross-verified excluding these flagged studies using REML and ML. After excluding the flagged studies, the pooled ES for RAT remained at the same level of 0.65 ($p < 0.01$) without significant variance component change. The pooled ES for RT experienced trivial change from 0.52 ($p > 0.05$) to 0.51 ($p > 0.05$), while the pooled ES for PAT change significantly from 0.55 to 0.70 ($p < 0.01$) after excluding flagged studies. We then further verified the results using change score from PAT data, the pooled results remained as 0.55 ($p < 0.01$). Therefore, after sensitivity analysis only two studies from RAT results were excluded from further analyses. Pre-test baseline differences for PAT and RT were given in Supplementary Tables S6, S7 and S8.

Certainty of evidence (Supplementary Table S12) was moderate for main outcome RAT, mainly owing to the inconsistent measures of RAT such as different angles, possible exit paths, and number of COD during tests. The overall certainty of evidence for secondary outcomes was low, mainly owing to inconsistent measurements and risk of publication bias.

Main effect

Twenty-eight studies provided 136 ESs in total, and the number was reduced to 59 ESs after applying RAT. After excluding highly influential studies, there were 53 ESs and 26 studies in total. Comparison of traditional meta-analysis showed that three-level meta-analysis had lower AIC, BIC, AICc ($p < 0.01$) for RAT and PAT but not for RT (Supplementary Table S9). Therefore, a three-level meta-analysis is necessary and should be utilized for further analyses in RAT and PAT. In addition, the baseline test revealed four pairs of data with significant difference. After conducted sensitivity analysis the overall ES only changed trivially (Supplementary Table S10-11), proving the stability of analysis results.

The main effect of agility training on RAT was large (ES = 0.65, 95% CI: 0.42-0.87, $Q(52) = 180.13$, $p < 0.01$, I^2-level 2 = 61.23%, I^2-level 3 = 0%, total I^2 = 61.23%, GRADE: moderate), indicating a significant improvement on RAT with moderate within-study heterogeneity and negligible between-study heterogeneity (Figure S4). However, it can be a consequence of low sample size (26 studies) and masking effect of high within-study heterogeneity due to inconsistent methodologies and participants' characteristics (Choi and Kang, 2025). Effects on RT (ES = 0.52, $p > 0.05$, GRADE: very low), and PAT (ES = 0.55, $p < 0.01$, GRADE: low) were medium (Figure S1, S2and S3). The results revealed improvements across all outcome dimensions except for RT with p value greater than 0.05.

Subgroup analyses were conducted when effect sizes clustered were greater than 10 according to pre-specified criteria (Figure 4). Although the effects on RAT were large across all age groups, analysis indicated that effects on adolescent (ES = 0.73, CI: 0.49 - 0.97, $p < 0.05$) and junior players (ES = 0.77, CI: 0.44 - 1.09, $p < 0.05$) are larger than senior players (ES = 0.62, CI: 0.27 - 0.97, $p < 0.05$). Although similar improvements were observed in adolescent and junior players, the results should be interpreted with caution due to imbalance sample size across strata. Effects on RAT were larger in studies incorporating both genders (ES = 0.71, CI: 0.25-1.18, $p < 0.05$) than those incorporating only male (ES = 0.64, CI: 0.45 - 0.83, $p < 0.05$). However, it should be noted that female participants were underrepresented in the pooled results. Therefore, results regarding this subgroup are highly preliminary. For training status, effects on national level athletes (ES = 0.80, CI: 0.55 - 1.06, $p < 0.05$) were greater than developmental athletes (ES = 0.58, CI: 0.34 - 0.82, $p < 0.05$). And interventions that used visual stimulus produced larger ES (ES = 0.85, CI: 0.60 - 1.11, $p < 0.05$) than human or visual-auditory stimulus, but it should be noted that pooled results were obtained from a relatively small sample size and imbalance exits between different modalities.

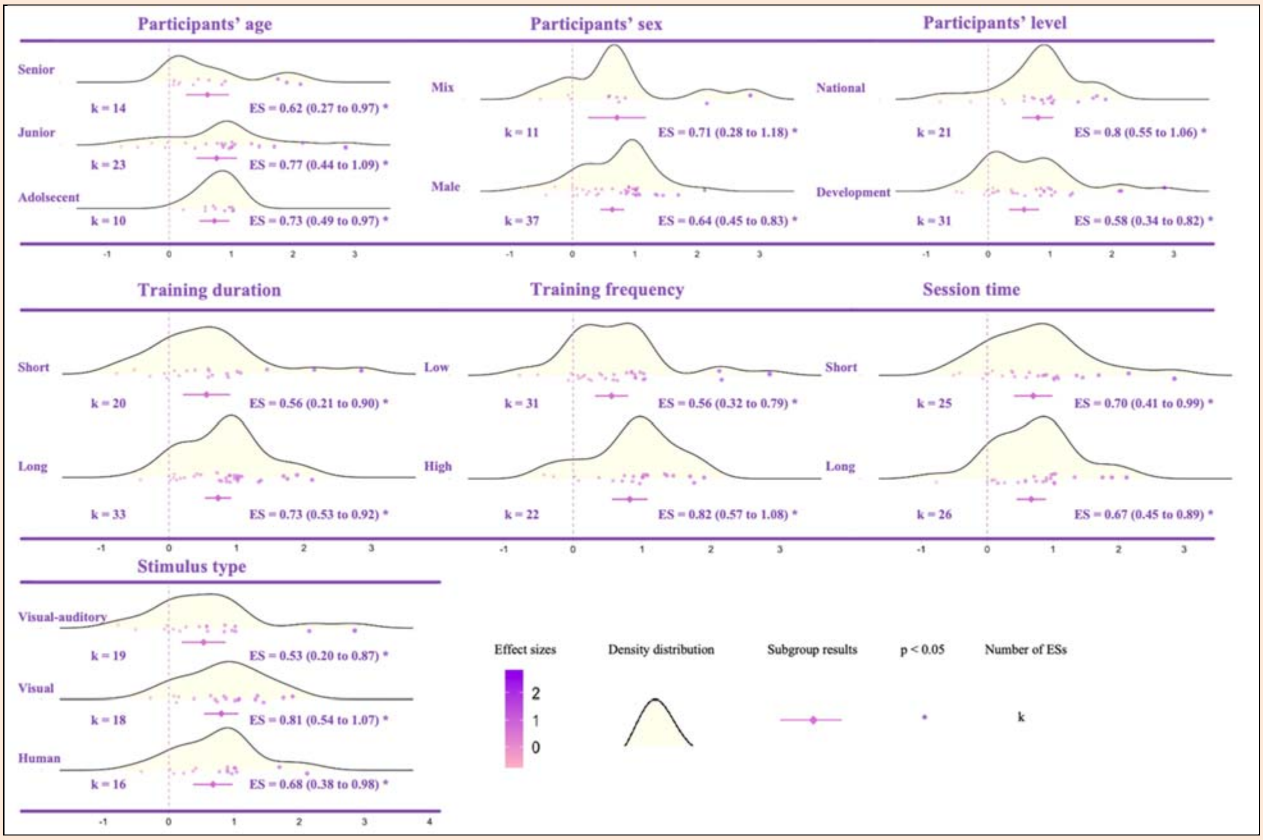

Figure 4. **Subgroup analysis results.**

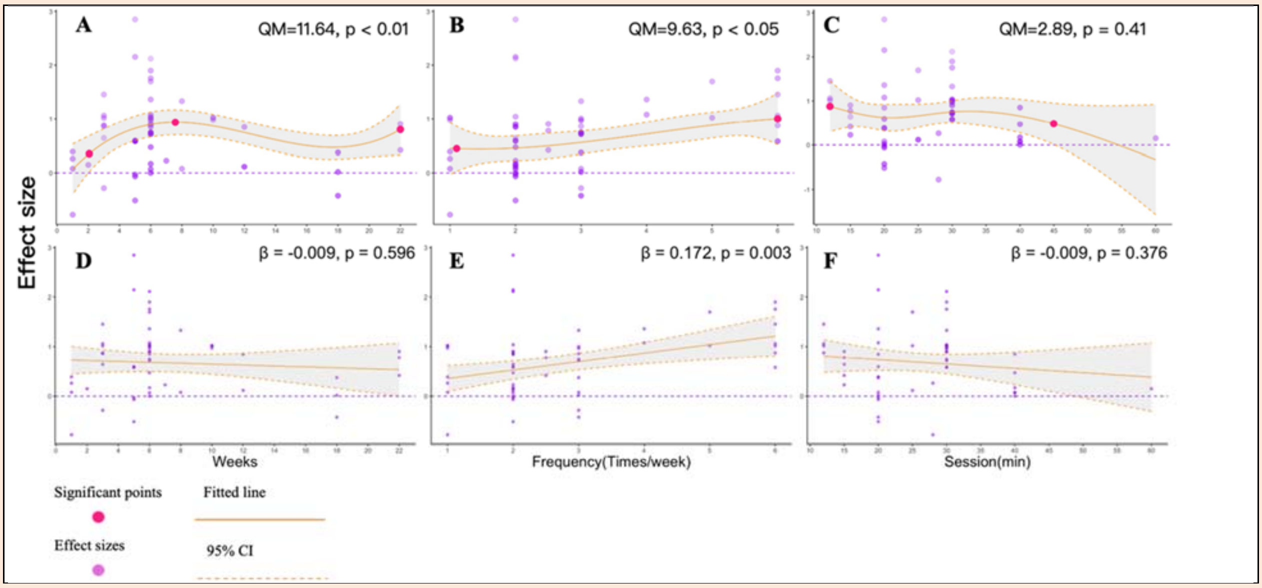

Figure 5. **Regression analyses for RAT. A-C Non-linear regression for training duration (weeks), training frequency (times/week) and session time (min) respectively; D-E Linear regression for training duration (weeks), training frequency (times/week) and session time (min) respectively; CI Confidence Interval.**

Effect sizes ranged from 0.56 (CI: 0.21 - 0.90, $p < 0.05$) to 0.82 (CI: 0.57 - 1.08, $p < 0.05$) across all subgroups. Specifically, interventions that involve visual stimulus with longer intervention (≥ 6 weeks), higher frequency (≥ 3 times/week) and short session time (< 25 minutes/session) had larger ESs.

Linear and non-linear meta-regression

Linear regression showed non-significant relationship between training duration ($\beta = -0.009$, $p = 0.596$, Figure 5), session time ($\beta = -0.009$, $p = 0.376$, Figure 5) and RAT. Analyses revealed significant positive linear relationship between training frequency ($\beta = 0.172$, $p < 0.01$, Figure 5) and RAT. Subsequently, linear regression revealed non-significant relationships between all training modalities and PAT or RT with p value greater than 0.05 (Figure S5). Except for relationship between training frequency and RAT, the linear regression model was implausible due to the goodness of fit.

Therefore, nonlinear regression models were fitted and cross-verified. Based on goodness-of-fit, cubic polynomial model fitted best for relationship between training weeks and RAT, simple linear model fitted training frequency best and restricted cubic spline model fitted session time best. We fitted each training modality with their best-fitted model separately, additionally, we fitted training frequency into cubic polynomial model to see if potential non-linear relationship exists.

Non-linear regression revealed a significant wavy relationship between training duration and RAT (QM = 11.64, $p < 0.01$, Figure 5). Training duration peaked at 7.58 weeks, the minimum significant point occurred at 2.06 weeks and the maximum point occurred at 22 weeks. Restricted cubic spline model yielded non-significant relationship between session time and RAT performance (QM = 3.17, $p = 0.37$, Figure 5). But positive significant ESs were clustered around 20 - 30 minutes per session. The analyses detected a near linear relationship between training frequency and RAT (QM = 9.63, $p < 0.05$, Figure 5). The maximum significant point and the peak point of training frequency overlapped at 6 times per week, while the minimum point occurred at 1 time per week. Therefore, linear model was used for further interpretation.

Discussion

The main aims of this review were to investigate the effects of agility training on agility performance among team-sport athletes, and to identify the optimal training parameters. Our findings indicate that agility training yields large overall effects on RAT performance, with moderate effects on PAT and non-significant effects on RT. Although participant characteristics and training modalities have potential moderating effects on agility training gains, the ecological validity varied among test protocols. Therefore, the overall and moderating effects should be interpreted in sport-specific context. Evidence certainty was moderate for RAT, low for PAT and very low for RT, and the results were characterized by considerable within-study heterogeneity. Non-linear analyses suggested that agility performance improvements play plateau beyond approximately 7-8 weeks of training, with higher frequencies and moderate session durations (20 - 25 minutes) showing favorable trends. These relationships are indicative rather than prescriptive, as the available data preclude precise model-based determination of optimal parameters.

The contribution of perception-action coupling

Previous research has demonstrated the importance of perceptual-cognitive factors in agility performance. While basic cognitive functions like simple reaction time can facilitate agility performance (Pojskic et al., 2018; Sekulic et al., 2019), higher order cognitive (executive) functions, encompassing working-memory, inhibitory control and cognitive flexibility are more crucial (Diamond, 2013). Working memory involves holding, manipulating and refreshing information in dynamic game context, such as monitoring teammates' positions and predicting opponents' actions. Inhibitory control is the ability to suppress automatic or irrelevant responses to focus on task-relevant information, enabling players to filter distractions like noise from spectators. While cognitive flexibility entails the ability to switch between tasks or to adapt to changing conditions, such as decision on a COD running during the transition between attack and defense.

Agility differs from COD ability in that it requires continuous integration of perceptual operations with motor execution (Young et al., 2015). Team-sport players are constantly challenged by simultaneous perceptual tasks (e.g. anticipation and decision-making) and motor execution (e.g. dribbling or running) (Jordet et al., 2020), highlighting the need for perception-action coupling during training (Young et al., 2021). Evidence that knee mechanics and decision processes diverge between reactive and pre-planned tasks further underscores the need for representative practice to maintain perception-action coupling (Brown et al., 2014). The results confirmed that training programs integrating perception-action coupling, such as COD training with external stimulus (Born et al., 2016; Chaalali et al., 2016; Zouhal et al., 2019), SSGs (Neag et al., 2025; Zago et al., 2016), paired drills and mirror games (Engelbrecht et al., 2016), and skill practice with external stimulus (Hicheur et al., 2020; Vogt et al., 2018), yielded greater ESs in RAT than PAT or RT.

To better interpret this, it's necessary to investigate the underlying mechanisms that facilitates this transfer of training. The effectiveness of perception-action coupled training can be understood through several interacting cognitive and neurophysiological mechanisms. A potential pathway might be the dual-process theory, which conceptualized skill acquisition as a gradual shift from controlled, deliberative operations (Type 2 processing) to automatic, intuitive responses (Type 1 processing) (Furley et al., 2015). Repeated exposure to representative perceptual-motor contexts accelerates the transfer from Type 2 to Type 1 processing, thereby, reducing the reliance on working memory and enabling faster, more automated responses (Anderson, 1982). Consequently, the adaptation explains why training tasks that integrate visual-spatial perception and motor execution, such as SSGs and reactive COD training, are more effective in improving RAT performance.

In conclusion, perception-action coupling is fundamental in agility training among team-sport players, as it ensures the seamless integration of perceptual-cognitive factors and motor processes, promoting adaptive performance in complex scenarios.

Moderators and dose-response modelling

Moderators of agility training

This study identified several moderators of agility training including participant characteristics and training modalities. Among these, age (maturation) is a prominent topic in agility training (Horníková et al., 2025; Thieschäfer and Büsch, 2022). Although subgroup analysis revealed that adolescent and junior players had similar training gains, there might be confounding factors underlying. Research indicated that maturation rather than chronological age, influences agility training gains with COD ability culminating before adolescence (around 14 years old) (Thieschäfer and Büsch, 2022). Results from 8 studies with a longitudenal design (Arede et al., 2021; Bekris et al., 2023; Ehmann et al., 2022; Gonzalo-Skok et al., 2023; Klatt et al., 2025; Neag et al., 2025; Trecroci et al., 2016; Vogt et al., 2018; Zago et al., 2016) in the current analysis aligned with previous research showing that younger participants experienced the largest improvement (Thieschäfer and Büsch, 2022). However, only one study reported maturity offset among youth soccer players (Trecroci et al., 2016), showing that agility training introduced pronounced effect (ES = 0.8, $p < 0.05$) on experimental group with small effect size (ES = 0.2) in control group regarding RAT. Rapidly developing physical abilities may contribute to observed large effect size (Di Mascio et al., 2020), which explains the pronounced ES in subgroup junior (11 - 14 years old). However, pronounced agility training gains are also observed after the above-mentioned culminating of COD ability (Pojskic et al., 2018). This may stem from improved perceptual-cognitive factors, which are more correlated with agility performance in middle adolescence than early adolescence (Horníková et al., 2025). Therefore, the moderator of age suggests different emphases in agility training for various maturation groups with physical factors being more important before adolescence and perceptual-cognitive factors being more important after. However, we are not suggesting training components of agility separately, as perception-action coupling is vital in agility performance (Young et al., 2021). Nevertheless, more studies with maturity status are needed to further validate the result.

Sex revealed trivial moderating influence on agility training gains, but only two studies recruited female participants in their experiments (Lee et al., 2024; Mancini et al., 2024), and six studies recruited both genders in their studies. Although subgroup analysis indicated that studies with both genders experienced similar large ES with studies recruiting male participants, the results had limited guidance on real-world practice due to insufficient data.

Previous research demonstrated that untrained team-sport athletes benefit more from various training methods than highly trained ones regarding COD speed (Forster et al., 2022; Nygaard Falch et al., 2019), which illustrated the contribution of physical training to COD speed and the moderating effects of training status. However, the moderating effects on RAT performance remain underexplored. The current analysis found that developmental team-sport players experienced less pronounced responses than national level players regarding RAT performance. Subgroup analysis revealed that overall ES for national level subjects (ES = 0.8, CI: 0.55 - 1.06, $p < 0.05$) was larger than that for developmental participants (ES = 0.58, CI: 0.34 - 0.82, $p < 0.05$). This was cross-verified by independent studies, showing that agility training produced significant group differences among elite participants in RAT performance ($p < 0.05$) (Zouhal et al., 2019), while no significant group effects ($p = 0.26$) were found among developmental participants. Possible mechanisms underlying this may be: physical factors dominate agility performance in developmental athletes, but their influence diminishes with advancing age and training status, while perceptual-cognitive factors become more prominent (Thieschäfer and Büsch, 2022). Therefore, elite athletes who possess greater match experience and better perceptual-cognitive abilities such as decision-making and anticipation can have superior transfer from training to RAT performance.

The study also explored the effects of various stimuli on RAT performance. Subgroup analysis revealed that using visual stimuli can generate very large improvement (ES = 0.81, CI: 0.54 - 1.07, $p < 0.05$) regarding RAT performance, while the second largest ES was observed in human stimuli (ES = 0.68, CI: 0.38 - 0.98, $p < 0.05$). However, the results should be interpreted with caution owing to contradiction to previous research suggesting using human and life-sized video as stimulus in agility training (Paul et al., 2016), possibly reflecting imbalanced samples with more visual-stimulus interventions, which will be discussed in Section 4.3.

Although moderators were explored using subgroup data, the moderate within-study heterogeneity suggest inconsistent internal study design, such as varied number of CODs, angles, entry and exit speed in RATs. Moreover, many studies incorporated less than 10 participants per group (Engelbrecht et al., 2016; Hassan et al., 2022; Hicheur et al., 2020; Lee et al., 2024; Serpell et al., 2011), which may cause unstable within-study and pooled ESs. Previous sensitivity analysis revealed that Hassan et al., (2022) and Steff et al., (2024) as highly influential studies, which changed the pooled estimates from 0.79 to 0.65. Therefore, studies with inconsistent internal study design and small sample size may potentially lead to inflated pooled estimates. Additionally, most samples were gained from soccer and basketball, the results remained an indicative estimate rather than a one-size-fits-all guidance. And interpretation of results should not be extrapolated to team-sport broadly.

Dose-response modelling of agility training

The present meta-regression analysis elucidated dose-response relationships between key training modalities (training duration, frequency and session time) and improvements in agility performance among team-sport athletes. Findings revealed a non-linear pattern between training duration and RAT as well as a linear pattern between training frequency and RAT, highlighting the potential threshold of agility training. For training duration, the wavy non-linear regression indicated a peak ES at 7.58 weeks with a minimum significant ES after 2 weeks, suggesting an initial rapid adaptation phase before 8 weeks followed by a dip and a later rising at 18 weeks.

Since agility performance requires "perception-action coupling" (Young et al., 2021), where movement and stimulus influence each other, training gains can be amplified through this "coupled" training. Therefore, resulting in primary ascending phase of agility performance. Especially for initial neural adaptation. As illustrated by Klatt et al., (2025), participants experienced significant improvement in reaction time ($p < 0.05$) and RAT ($p < 0.05$) after 6 weeks of agility training. In addition, perceptual-cognitive abilities develop rapidly during adolescence (Legault et al., 2022), making maturation a confounding factor contributing to rapid RAT performance improvement. The later stagnant and diminishing phase may stem from plateau effects. It was suggested that training monotony can result in training plateau for physical ability like explosive lower limb power (Oliveira et al., 2025), which is important in agility performance (Sheppard and Young, 2006). However, this can be solved by using short intense training mesocycle, which lowers the chance of training monotony and plateau of adaptation (Forster et al., 2022). Therefore, diverse training design can mitigate the potential plateau effects after 8 weeks of agility training. And only one study with excessive long training duration appeared in the later returning phase (Zago et al., 2016). The observed agility training gains may stem from progressive loading and diverse drill formats (Zago et al., 2016), and progressive loading was proven effective in COD ability among various population (Thompson et al., 2017; Zhou et al., 2024). Therefore, preventing training monotony and establishing progressive loading can assist agility training gains in excessive long training duration. However, more studies with longer intervention duration are warranted to validate the mechanism.

There was no significant relationship between session time and RAT based on linear and non-linear regression results. But reflecting on the non-linear plot for session time, it can be inferred that largest significant ESs clustered around 20 - 30 minutes per session. And previous subgroup analysis showed that shorter session time (< 25 minutes) yielded larger responses. Therefore, it can be concluded that a session time between 20 - 25 minutes is plausible in agility training. Possible mechanism for smaller ESs in longer session time may stem from mental fatigue. Mental fatigue is a psychobiological state caused by excessive long period of demanding cognitive activity (Marcora et al., 2009), which has been proved harmful in sport-specific skills in Australian Football (Weerakkody et al., 2021), soccer (Smith et al., 2016) and cricket (Veness et al., 2017). While in agility training, mental fatigue is inevitable for exercises are performed in response to a stimulus. Therefore, to avoid detrimental effects caused by prolonged training session, agility training should be conducted before fatigue. However, a session time between 20 - 25 minutes should be interpreted with caution due to relatively high heterogeneity and low certainty of evidence. Moreover, evidence was aggregated mostly from soccer and basketball, the results may not be universally applicable in team-sport.

In contrast, the positive "linear" relationship for training frequency indicated that higher frequencies may have favorable adaptive RAT performance. Although most studies clustered at 2 - 3 times/week, maximum ES was achieved at 6 times/week. The result along with relationship between training duration, session time and RAT performance indicated that repeated exposure to sport-specific stimuli in agility training without excessive fatigue for a suitable duration can appropriately enhance RAT performance in team-sport players. Study also revealed similar results regarding training frequency and COD ability with a positive relation of $r = 0.436$ ($p < 0.05$) (Asadi et al., 2017). However, this study only reflected the physical aspect of agility performance, the perceptual-cognitive factors remain underexplored. It should also be noted that training frequency should be adjusted based on volume and intensity to avoid injury and over-training (Bourdon et al., 2017). Therefore, future studies should investigate the influence of training frequency on RAT performance using both short and long session time while controlling training volume and intensity to validate the results.

In conclusion, 7 - 8 weeks interventions delivered at higher frequency with approximately 20 - 25 minutes session time are associated with favorable adaptations. However, monitoring volume and intensity is critical to avoid injury or over-training. Especially when the data regarding training volume and intensity was absent from many studies, scrutiny monitoring (e.g. total distance covered, hear rate, rate of perceived exertion etc.) should accompany training programs. But the results should be interpreted with caution due to the dominance of soccer and basketball samples.

Revisiting the importance of stimuli

In the current research, results indicated that using visual stimuli in agility training was associated with favorable performance outcome ($ES = 0.81$, $p < 0.05$). However, the results need further discussion due to the relatively small sample size, imbalance across modalities and contradiction to prior theory. Team-sport players are rarely required to react to light or video stimulus and most decisions are made based on teammates or opponents' actions. But most included studies were conducting training and testing using visual stimuli (light and video), except for (Chaalali et al., 2016; Chaouachi et al., 2014; Engelbrecht et al., 2016; Lee et al., 2024; Neag et al., 2025; Young and Rogers, 2014; Zago et al., 2016; Zouhal et al., 2019). Although training with these stimuli may improve RAT performance among team-sport athletes, the transferability on on-field performance remains skeptical. Because training stimuli should match the demands of the sports being trained for (Ramirez-Campillo et al., 2023), and light and video stimulus are rare in game context. The widely used light and video stimulus in agility training and testing may stem from its high level of repeatability. For such programmed stimulus can be generated at the same time on each occasion (Paul et al., 2016), making it replicable. And the protocol was used for both training and testing, therefore, when participants become more familiar with the system, the performance in tests would potentially increases. However, in game context, athletes can anticipate opponents' movements based on their experience (Roca et al., 2012), enabling them to make faster decisions, which is an important factor to agility performance (Sheppard and Young, 2006). While using light stimulus restricts anticipation, the stimulus is training and testing the speed of decision-making independent of game context in team-sport athletes. Thus, training agility with sport-specific context is highly recommended (Young et al., 2021). But variability of using human stimulus is also a major concern. To fix the problem, one study tried a novel agility training method in soccer using human stimulus (Chaalali et al., 2016). By touching the ball with both feet, the coach consistently provides at most four possible COD running. Similarly, one study used chasing runs and "mirror" drills in agility training (Trecroci et al., 2016), controlling the variability of using human stimulus. The methods provided players with live interaction and information for anticipation, greatly improving sport-specific stimulus in agility training. Also, agility training gains in these studies were both significant. Subsequently, human stimuli were also utilized in RAT. Attempt was made using a tester to initiate four possible scenarios based on passing movements in soccer and recorded total time, tester time, response time, decision-making time and movement time respectively (Altmann et al., 2022). The results provided valuable information on player development with good overall validity (ICC = 0.82, $p < 0.01$). And the ability of discriminating high-level and low-level players were addressed by a similar test with four possible scenarios in Australian Football (Veale et al., 2010), which correctly classified 75% of the participants. A systematic review also recommended the use of human stimulus in agility studies (Paul et al., 2016), highlighting repeatability of coaches and testers. Therefore, incorporating human stimuli in agility training and testing can provide better game context and discriminability.

However, we're not discouraging the utilization of visual stimuli in agility training. Since advanced technologies such as BATAK pro (Arede et al., 2021), digital screen (Serpell et al., 2011), and Speedcourt (Born et al., 2016) etc. have made using visual stimuli more systematic and reliable. It's more accessible for coaches and practitioners to utilize visual stimuli in agility training. And live-sized video is preferable when using 2-D visual stimulus (Paul et al., 2016). A recent research suggested that decision-making is equally important as anticipation (Williams et al., 2023), where video-based techniques were used to validate the theory. Therefore, we suggest that coaches can utilize visual stimuli (preferably live-sized video) to training decision-making speed to positively influence agility performance at off-season or preparation stage of training. While human stimuli are encouraged during in-season period for better sport-specific stimulus and anticipation involved training.

Retention of agility training

Research suggested that agility performance plateaus around adolescence and have distinct difference between early and middle adolescence (Horníková et al., 2025). Therefore, implementing agility training before this key stage is of great importance. However, dual-career limitations such as balancing training and education for young players (Pink et al., 2018), injury and off-season period may prevent them from performing their regular habitual exercises, which significantly lower their agility training adherence. Therefore, it's pivotal to examine the retention of agility training-induced effects to better inform training schedule. However, most included studies didn't perform retention test except for two studies (Engelbrecht et al., 2016; Zwierko et al., 2023). Investigation on the retention effects of agility training in rugby players using human stimulus revealed that RAT performance was maintained after a 6-week period without receiving the training stimulus (Engelbrecht et al., 2016). Another study also investigated the retention effects of agility training in volleyball players using stroboscopic glasses (Zwierko et al., 2023). Anecdotally, results showed that RAT performance was maintained in control group rather than experimental group after 4-week period without training. This may suggest that retention of agility training using stroboscopic glasses is trivial. Moreover, there is a large body of evidence examining the retention of physical components of agility such as power and strength (Chaouachi et al., 2019), sprint performance (Padrón-Cabo et al., 2025), and dynamic strength (Ingle et al., 2006). While the retention of perceptual-cognitive components of agility and agility performance is under-represented.

Overall, existing evidence suggest agility training gains among team-sport players may be partially maintained. But the retention effects were inconsistent and under-represented, which limits the information on proper scheduling of training sessions. More research is needed to validate the retention of agility training gains, especially the retention of perceptual-cognitive components of agility using different stimulus.

Limitations and Methodological Considerations

While this systematic review with meta-analysis provides evidence on agility training effects and prescriptions, several limitations warrant consideration. First, training volume and intensity were not coded owing to inconsistent reporting in different agility training modalities such as SSGs, speed agility and quickness (SAQ) training, and stroboscopic training etc. Future studies should seek to explore the standardized way to monitor agility training load (e.g. using rate of perceived exertion or heart rate for internal load and using global positioning systems for external load). Second, subgroup analyses for moderators like sex and training status were constrained by small study numbers and the underrepresentation of female subjects, limiting information on targeted population. Third, although the moderator "age" revealed that adolescent team-sport players benefit more from agility training, maturation was not thoroughly considered due to the insufficient data. Thus, it can only be referred that players around the pre-specified chronological age category may benefit more from agility training. Moreover, most evidence was gained from soccer and basketball, narrowing the scope of team-sport. Therefore, results should be applied cautiously, especially regarding team-sport players other than soccer and basketball or female subjects. Lastly, this meta-analysis only included interventions that incorporated exercises with perceptual-cognitive factors, partially overlooked potential research that train components of agility independently.

For methodological consideration, the registration of this study took place after initial article retrieval but before data extraction. Which is allowed in PROSPERO registration but could introduce potential selective reporting. Ideally, registration should precede literature search. Additionally, although dual screening with a third reviewer resolving disagreements aligns with PRISMA 2020 guidelines, not reporting agreement metrics (e.g. Cohen's kappa coefficient) could obscure the consistency of decisions. The heterogeneity across various RAT modalities narrows the applicability of the pooled results. To the best of our knowledge, there isn't a consensus regarding the gold ecological standard of reactive agility testing. Therefore, we only included reactive agility tests modalities that are recognized as spatial agility or universal agility to improve the consistency of the RAT results. While our primary analysis provides evidence for large improvements in RAT performance (ES = 0.65), a key limitation involves the exclusion of two influential outliers influential (Hassan et al., 2022; Steff et al., 2024). These studies were excluded based on pre-specified criteria, including high Cook's distance, and disproportionate contributions to heterogeneity (I^2 from 61.23% to 77.43%). Methodologically, both studies exhibited risks such as small sample sizes, lack of randomization and blinding, and high test-training specificity using FITLIGHT system, which may have inflated their study ESs (1.11 - 3.56). Therefore, excluding them provide a more conservative but less biased pooled estimates, future studies should examine the large effects with rigorous designs to confirm their validity.

Recommendations for Future Agility Studies

To advance agility research, future studies should address current gaps with targeted designs. Included studies had relative high risk of bias regarding familiarization of tests, reporting of reliability and sample size, which should be addressed by future studies for better practical guidance. Upon improving study design, reporting on maturation is warranted in agility studies for more specific exploration on possible "window of opportunity" for agility training among team-sport players. Subsequently, visual stimuli and human stimuli in agility training should be evaluated using on-field performance to better understand the contribution of different perceptual-cognitive factors. But proper training for tester should be provided to ensure consistent stimulus before implementing agility training. In addition, there is a call for agility research on female participants to better inform the gender difference in agility training gains. Also, future studies should compare the agility training gains between elite and non-elite team-sport players to validate current results. Given the barriers like dual-career limitations, injuries and off-season period can prevent players from training, it's of great value to examine the retention of agility training gains using various modalities to better inform training scheduling among various competitive levels.

Conclusion

The results of the current meta-analysis revealed that the overall effect of agility training on RAT performance among team-sport athletes was large. The findings suggest that interventions of approximately 7-8 weeks, performed more than three times per week with sessions lasting around 20-25 minutes, were most frequently associated with favorable adaptations. However, these values reflect the central tendencies of available studies rather than model-predicted optima due to the high heterogeneity, moderate to low certainty of evidence, the dominance of soccer and basketball samples, underrepresentation of female participants, insufficient training load and retention data. Nevertheless, heart rate, rate of perceived exertion and global positioning systems should be utilized during agility training to monitor internal and external load to prevent injury and over-training. Coaches and practitioner may use visual stimuli for decision-making speed during off-season or preparation stage, which can positively influence the overall agility performance. However, human stimuli are encouraged during in-season period as they provide a more sport-specific stimulus, which is critical for anticipation. Overall, the current study is an attempt in aggregating evidence regarding agility training effects among team-sport players. But the limited and imbalanced data restricting practical guidance on team-sport training. Future studies should explore agility training in more diverse demographics.

Acknowledgements

Protocol available from:
https://www.crd.york.ac.uk/PROSPERO/view/CRD420251060149.
The author reports no actual or potential conflicts of interest. The datasets generated and analyzed in this study are not publicly available, but are available from the corresponding author who organized the study upon reasonable request. All experimental procedures were conducted in compliance with the relevant legal and ethical standards of the country where the study was performed.

References

Altmann, S., Neumann, R., Ringhof, S., Rumpf, M. C. and Woll, A. (2022) Soccer-specific agility: Reliability of a newly developed test and correlates of performance. *The Journal of Strength and Conditioning Research* **36(5)**. https://doi.org/10.1519/JSC.0000000000003635

Anderson, J. R. (1982) Acquisition of cognitive skill. *Psychological Review* **89(4)**, 369. https://doi.org/10.1037/0033-295X.89.4.369

Arede, J., Carvalho, M., Esteves, P., de las Heras, B. and Leite, N. (2021) Exploring the effects of LED lighting training program on motor performance among young athletes. *Creativity Research Journal* **33(1)**, 63-73. https://doi.org/10.1080/10400419.2020.1817693

Asadi, A., Arazi, H., Ramirez-Campillo, R., Moran, J. and Izquierdo, M. (2017) Influence of maturation stage on agility performance gains after plyometric training: A systematic review and meta-analysis. *The Journal of Strength and Conditioning Research* **31(9)**, 2609-2617. https://doi.org/10.1519/JSC.0000000000001994

Assink, M. and Wibbelink, C. J. (2016) Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology* **12(3)**, 154-174. https://doi.org/10.20982/tqmp.12.3.p154

Bekris, E., Gioldasis, A., Souglis, A., Zacharakis, E. and Smirniotou, A. (2023) Enhancing soccer-specific motor skills through visual training: A quasi-experimental study in young soccer players. *Baltic Journal of Sport and Health Sciences* **2(129)**, 28-38. https://doi.org/10.33607/bjshs.v2i129.1381

Born, D.-P., Zinner, C., Düking, P. and Sperlich, B. (2016) Multi-directional sprint training improves change-of-direction speed and reactive agility in young highly trained soccer players. *Journal of Sports Science and Medicine* **15(2)**, 314. https://pubmed.ncbi.nlm.nih.gov/27274670/

Bourdon, P. C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M. C., Gabbett, T. J., Coutts, A. J., Burgess, D. J. and Gregson, W. (2017) Monitoring athlete training loads: Consensus statement. *International Journal of Sports Physiology and Performance* **12(S2)**, 161- 170. https://doi.org/10.1123/IJSPP.2017-0208

Brown, S. R., Brughelli, M. and Hume, P. A. (2014) Knee mechanics during planned and unplanned sidestepping: A systematic review and meta-analysis. *Sports Medicine* **44(11)**, 1573-1588. https://doi.org/10.1007/s40279-014-0225-3

Brughelli, M., Cronin, J., Levin, G. and Chaouachi, A. (2008) Understanding change of direction ability in sport. *Sports Medicine* **38(12)**, 1045-1063. https://doi.org/10.2165/00007256-200838120-00007

Chaalali, A., Rouissi, M., Chtara, M., Owen, A., Bragazzi, N. L., Moalla, W., Chaouachi, A., Amri, M. and Chamari, K. (2016) Agility training in young elite soccer players: Promising results compared to change of direction drills. *Biology of Sport* **33(4)**, 345-351. https://doi.org/10.5604/20831862.1217924

Chaouachi, A., Ben Othman, A., Makhlouf, I., Young, J. D., Granacher, U. and Behm, D. G. (2019) Global training effects of trained and untrained muscles with youth can be maintained during 4 weeks of detraining. *The Journal of Strength and Conditioning Research* **33(10)**. https://doi.org/10.1519/JSC.0000000000002606

Chaouachi, A., Chtara, M., Hammami, R., Chtara, H., Turki, O. and Castagna, C. (2014) Multidirectional sprints and small-sided games training effect on agility and change of direction abilities in youth soccer. *The Journal of Strength and Conditioning Research* **28(11)**, 3121-3127. https://doi.org/10.1519/JSC.0000000000000505

Cheung, M. W.-L. (2019) A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review* **29(4)**, 387-396. https://doi.org/10.1007/s11065-019-09415-6

Choi, G. J. and Kang, H. (2025) Heterogeneity in meta-analyses: An unavoidable challenge worth exploring. *Korean Journal of Anesthesiology* **78(4)**, 301-314. https://doi.org/10.4097/kja.25001

Citkowicz, M. and Vevea, J. L. (2017) A parsimonious weight function for modeling publication bias. *Psychological Methods* **22(1)**, 28. https://doi.org/10.1037/met0000119

Cobley, S., Baker, J., Wattie, N. and McKenna, J. (2009) Annual age-grouping and athlete development. *Sports Medicine* **39(3)**, 235-256. https://doi.org/10.2165/00007256-200939030-00005

Cohen, J. (2013) *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

Currell, K. and Jeukendrup, A. E. (2008) Validity, reliability and sensitivity of measures of sporting performance. *Sports Medicine* **38(4)**, 297-316. https://doi.org/10.2165/00007256-200838040-00003

Di Mascio, M., Ade, J., Musham, C., Girard, O. and Bradley, P. S. (2020) Soccer-specific reactive repeated-sprint ability in elite youth soccer players: Maturation trends and association with various physical performance tests. *The Journal of Strength and Conditioning Research* **34(12)**, 3538-3545. https://doi.org/10.1519/JSC.0000000000002362

Diamond, A. (2013) Executive functions. *Annual Review of Psychology* **64(1)**, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dunton, A., O'Neill, C. and Coughlan, E. K. (2020) The impact of a spatial occlusion training intervention on pass accuracy across a continuum of representative experimental design in football. *Science and Medicine in Football* **4(4)**, 269-277. https://doi.org/10.1080/24733938.2020.1745263

Ehmann, P., Beavan, A., Spielmann, J., Mayer, J., Ruf, L., Altmann, S., Forcher, L., Klever, N., Rohrmann, S. and Nuß, C. (2022) Perceptual-cognitive performance of youth soccer players in a 360-environment - an investigation of the relationship with soccer-specific performance and the effects of systematic training. *Psychology of Sport and Exercise* **61**, 102220. https://doi.org/10.1016/j.psychsport.2022.102220

Engelbrecht, L., Terblanche, E. and Welman, K. E. (2016) Video-based perceptual training as a method to improve reactive agility performance in rugby union players. *International Journal of Sports Science and Coaching* **11(6)**, 799-809. https://doi.org/10.1177/1747954116676106

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P. and Van den Noortgate, W. (2021) Detecting selection bias in meta-analyses with multiple outcomes: A simulation study. *The Journal of Experimental Education* **89(1)**, 125-144. https://doi.org/10.1080/00220973.2019.1582470

Forster, J. W. D., Uthoff, A. M., Rumpf, M. C. and Cronin, J. B. (2022) Training to improve pro-agility performance: A systematic review. *Journal of Human Kinetics* **85**, 35-51.

Friebe, D., Banzer, W., Giesche, F., Haser, C., Hülsdünker, T., Pfab, F., Rußmann, F., Sieland, J., Spataro, F. and Vogt, L. (2024) Effects of 6-week motor-cognitive agility training on football test performance in adult amateur players - a three-armed randomized controlled trial. *Journal of Sports Science and Medicine* **23(2)**, 276. https://doi.org/10.52082/jssm.2024.276

Furley, P., Schweizer, G. and Bertrams, A. (2015) The two modes of an athlete: Dual-process theories in the field of sport. *International Review of Sport and Exercise Psychology* **8(1)**, 106-124. https://doi.org/10.1080/1750984X.2015.1022203

Gonzalo-Skok, O., Tous-Fajardo, J., Maroto-Izquierdo, S., Raya-González, J. and Sánchez-Sánchez, J. (2023) The inclusion of preplanned and random and unanticipated/unexpected events during strength training improves the ability to repeat high-intensity efforts under uncertainty. *International Journal of Sports Physiology and Performance* **19(3)**, 249-256. https://doi.org/10.1123/ijspp.2023-0107

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P. and Schünemann, H. J. (2008) GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* **336(7650)**, 924-926. https://doi.org/10.1136/bmj.39489.470347.AD

Harrer, M., Cuijpers, P., Furukawa, T. and Ebert, D. (2021) *Doing meta-analysis with R: A hands-on guide*. Chapman and Hall/CRC. https://doi.org/10.1201/9781003107347

Hassan, A. K., Alhumaid, M. M. and Hamad, B. E. (2022) The effect of using reactive agility exercises with the FITLIGHT training system on the speed of visual reaction time and dribbling skill of basketball players. *Sports* **10(11)**. https://doi.org/10.3390/sports10110176

Hedges, L. V. and Olkin, I. (2014) *Statistical methods for meta-analysis*. Academic Press.

Hedges, L. V., Tipton, E. and Johnson, M. C. (2010) Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* **1(1)**, 39-65. https://doi.org/10.1002/jrsm.5

Hicheur, H., Chauvin, A., Cavin, V., Fuchslocher, J., Tschopp, M. and Taube, W. (2020) Augmented-feedback training improves cognitive motor performance of soccer players. *Medicine and Science in Sports and Exercise* **52**, 141. https://doi.org/10.1249/MSS.0000000000002118

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L. and Sterne, J. A. C. (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, d5928. https://doi.org/10.1136/bmj.d5928

Horníková, H., Hadža, R. and Zemková, E. (2025) The contribution of perceptual-cognitive skills to reactive agility in early and middle adolescent soccer players. *The Journal of Strength and Conditioning Research* **39(3)**. https://doi.org/10.1519/JSC.0000000000004997

Ingle, L., Sleap, M. and Tolfrey, K. (2006) The effect of a complex training and detraining programme on selected strength and power variables in early pubertal boys. *Journal of Sports Sciences* **24(9)**, 987-997. https://doi.org/10.1080/02640410500457117

Jordet, G., Aksum, K. M., Pedersen, D. N., Walvekar, A., Trivedi, A., McCall, A., Ivarsson, A. and Priestley, D. (2020) Scanning, contextual factors, and association with performance in English Premier League footballers: An investigation across a season. *Frontiers in Psychology* **11**. https://doi.org/10.3389/fpsyg.2020.553813

Kadlec, D., Sainani, K. L. and Nimphius, S. (2023) With great power comes great responsibility: Common errors in meta-analyses and meta-regressions in strength and conditioning research. *Sports Medicine* **53(2)**, 313-325. https://doi.org/10.1007/s40279-022-01766-0

Klatt, S., Thönnißen, L., Fischer, T. and Sandmann, A. (2025) Alternative training methods in soccer: A perceptual-skill intervention in youth soccer players. *German Journal of Exercise and Sport Research*. https://doi.org/10.1007/s12662-025-01017-0

Lee, Y.-S., Lee, D. and Ahn, N. Y. (2024) SAQ training on sprint, change-of-direction speed, and agility in U-20 female football players. *Plos One* **19(3)**, e0299204. https://doi.org/10.1371/journal.pone.0299204

Legault, I., Sutterlin-Guindon, D. and Faubert, J. (2022) Perceptual cognitive abilities in young athletes: A gender comparison. *Plos One* **17(8)**, e0273607. https://doi.org/10.1371/journal.pone.0273607

Lucia, S., Digno, M., Madinabeitia, I. and Di Russo, F. (2023) Testing a multicomponent training designed to improve sprint, agility and decision-making in elite basketball players. *Brain Sciences* **13(7)**, 984. https://doi.org/10.3390/brainsci13070984

Lucia, S., Mattia, D., Iker, M. and Di Russo, F. (2024) Integration of cognitive-motor dual-task training in physical sessions of highly-skilled basketball players. *Journal of Sports Sciences* **42(18)**, 1695-1705. https://doi.org/10.1080/02640414.2024.2408191

Mancini, N., Di Padova, M., Polito, R., Mancini, S., Dipace, A., Basta, A., Colella, D., Limone, P., Messina, G. and Monda, M. (2024) The impact of perception-action training devices on quickness and reaction time in female volleyball players. *Journal of Functional Morphology and Kinesiology* **9(3)**, 147. https://doi.org/10.3390/jfmk9030147

Marcora, S. M., Staiano, W. and Manning, V. (2009) Mental fatigue impairs physical performance in humans. *Journal of Applied Physiology* **106(3)**, 857-864. https://doi.org/10.1152/japplphysiol.91324.2008

Martin-Niedecken, A. L., Bucher, V., Adcock, M., de Bruin, E. D. and Schättin, A. (2023) Impact of an exergame intervention on cognitive-motor functions and training experience in young team sports athletes: A non-randomized controlled trial. *Frontiers in Sports and Active Living* **5**, 1170783. https://doi.org/10.3389/fspor.2023.1170783

McKay, A. K. A., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J. and Burke, L. M. (2022) Defining training and performance caliber: A participant classification framework. *International Journal of Sports Physiology and Performance* **17(2)**, 317-331. https://doi.org/10.1123/ijspp.2021-0451

Morral-Yepes, M., Moras, G., Bishop, C. and Gonzalo-Skok, O. (2022) Assessing the reliability and validity of agility testing in team sports: A systematic review. *The Journal of Strength and Conditioning Research* **36(7)**. https://doi.org/10.1519/JSC.0000000000003753

Neag, I., Mihaila, I., Fleancu, L. J., Stancu, M., Potop, V., Barbu, D., Paun, L. I. and Mihai, I. (2025) Agility development in youth soccer: The efficacy of fixed-role small-sided games. *Frontiers in Sports and Active Living* **7**, 1593906. https://doi.org/10.3389/fspor.2025.1593906

Nygaard Falch, H., Guldteig Rædergård, H. and van den Tillaar, R. (2019) Effect of different physical training forms on change of direction ability: A systematic review and meta-analysis. *Sports Medicine - Open* **5(1)**, 53. https://doi.org/10.1186/s40798-019-0223-y

Oliveira, J. P., Marinho, D. A., Jacinto, P., Sampaio, T. and Morais, J. E. (2025) Characterization of physical performance and change of direction deficit across age groups in young female volleyball players. *BMC Sports Science, Medicine and Rehabilitation* **17(1)**, 209. https://doi.org/10.1186/s13102-025-01264-6

Padrón-Cabo, A., Lorenzo-Martínez, M., De Dios-Álvarez, V., Rey, E. and Solleiro-Durán, D. (2025) Effects of a short-term detraining period on the physical fitness in elite youth soccer players: A comparison between chronological age groups. *The Journal of Strength and Conditioning Research* **39(2)**. https://doi.org/10.1519/JSC.0000000000004975

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A. and Brennan, S. E. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery* **88**, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Paul, D. J., Gabbett, T. J. and Nassis, G. P. (2016) Agility in team sports: Testing, training and factors affecting performance. *Sports Medicine* **46(3)**, 421-442. https://doi.org/10.1007/s40279-015-0428-2

Pink, M. A., Lonie, B. E. and Saunders, J. E. (2018) The challenges of the semi-professional footballer: A case study of the management of dual career development at a Victorian Football League (VFL) club. *Psychology of Sport and Exercise* **35**, 160-170. https://doi.org/10.1016/j.psychsport.2017.12.005

Pojskic, H., Åslin, E., Krolo, A., Jukic, I., Uljevic, O., Spasic, M. and Sekulic, D. (2018) Importance of reactive agility and change of direction speed in differentiating performance levels in junior soccer players: Reliability and validity of newly developed soccer-specific tests. *Frontiers in Physiology* **9**, 506. https://doi.org/10.3389/fphys.2018.00506

Ramirez-Campillo, R., García-Hermoso, A., Moran, J., Chaabene, H., Negra, Y. and Scanlan, A. T. (2022) The effects of plyometric jump training on physical fitness attributes in basketball players: A meta-analysis. *Journal of Sport and Health Science* **11(6)**, 656-670. https://doi.org/10.1016/j.jshs.2020.12.005

Ramirez-Campillo, R., Sortwell, A., Moran, J., Afonso, J., Clemente, F. M., Lloyd, R. S., Oliver, J. L., Pedley, J. and Granacher, U. (2023) Plyometric-jump training effects on physical fitness and sport-specific performance according to maturity: A systematic review with meta-analysis. *Sports Medicine - Open* **9(1)**, 23. https://doi.org/10.1186/s40798-023-00568-6

Roca, A., Williams, A. M. and Ford, P. R. (2012) Developmental activities and the acquisition of superior anticipation and decision making in soccer players. *Journal of Sports Sciences* **30(15)**, 1643-1652. https://doi.org/10.1080/02640414.2012.701761

Rodgers, M. A. and Pustejovsky, J. E. (2021) Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods* **26(2)**, 141. https://doi.org/10.1037/met0000300

Sekulic, D., Foretic, N., Gilic, B., Esco, M. R., Hammami, R., Uljevic, O., Versic, S. and Spasic, M. (2019) Importance of agility performance in professional futsal players; Reliability and applicability of newly developed testing protocols. *International Journal of Environmental Research and Public Health* **16(18)**, 3246. https://doi.org/10.3390/ijerph16183246

Serpell, B. G., Young, W. B. and Ford, M. (2011) Are the perceptual and decision-making components of agility trainable? A preliminary investigation. *The Journal of Strength and Conditioning Research* **25(5)**, 1240-1248. https://doi.org/10.1519/JSC.0b013e3181d682e6

Sheppard, J. M. and Young, W. B. (2006) Agility literature review: Classifications, training and testing. *Journal of Sports Sciences* **24(9)**, 919-932. https://doi.org/10.1080/02640410500457109

Smith, M. R., Coutts, A. J., Merlini, M., Deprez, D., Lenoir, M. and Marcora, S. M. (2016) Mental fatigue impairs soccer-specific physical and technical performance. *Medicine and Science in Sports and Exercise* **48(2)**, 267-276. https://doi.org/10.1249/MSS.0000000000000762

Steff, N., Badau, D. and Badau, A. (2024) Improving agility and reactive agility in basketball players U14 and U16 by implementing Fitlight technology in the sports training process. *Applied Sciences* **14(9)**, 3597. https://doi.org/10.3390/app14093597

Sun, M., Soh, K. G., Ma, S., Wang, X., Zhang, J. and Yaacob, A. B. (2025) Effects of speed, agility, and quickness (SAQ) training on soccer player performance—A systematic review and meta-analysis. *Plos One* **20(2)**, e0316846. https://doi.org/10.1371/journal.pone.0316846

Thieschäfer, L. and Büsch, D. (2022) Development and trainability of agility in youth: A systematic scoping review. *Frontiers in Sports and Active Living* **4**, 952779. https://doi.org/10.3389/fspor.2022.952779

Thompson, B. J., Stock, M. S., Mota, J. A., Drusch, A. S., DeFranco, R. N., Cook, T. R. and Hamm, M. A. (2017) Adaptations associated with an after-school strength and conditioning program in middle-school-aged boys: A quasi-experimental design. *The Journal of Strength and Conditioning Research* **31(10)**, 2840-2851. https://doi.org/10.1519/JSC.0000000000001712

Trecroci, A., Milanović, Z., Rossi, A., Broggi, M., Formenti, D. and Alberti, G. (2016) Agility profile in sub-elite under-11 soccer players: Is SAQ training adequate to improve sprint, change of direction speed and reactive agility performance? *Research in Sports Medicine* **24(4)**, 331-340. https://doi.org/10.1080/15438627.2016.1228063

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F. and Sánchez-Meca, J. (2013) Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods* **45**, 576-594. https://doi.org/10.3758/s13428-012-0261-6

Veale, J. P., Pearce, A. J. and Carlson, J. S. (2010) Reliability and validity of a reactive agility test for Australian football. *International Journal of Sports Physiology and Performance* **5(2)**, 239-248. https://doi.org/10.1123/ijspp.5.2.239

Veness, D., Patterson, S. D., Jeffries, O. and Waldron, M. (2017) The effects of mental fatigue on cricket-relevant performance among elite players. *Journal of Sports Sciences* **35(24)**, 2461-2467. https://doi.org/10.1080/02640414.2016.1273540

Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36(1)**, 1-48. https://doi.org/10.18637/jss.v036.i03

Vogt, T., Gassen, S., Wrede, S., Spielmann, J., Jedrusiak-Jung, M., Härtel, S. and Mayer, J. (2018) Football practice with youth players in the "Footbonaut". *German Journal of Exercise and Sport Research* **48(3)**, 341-348. https://doi.org/10.1007/s12662-018-0514-3

Weerakkody, N., Taylor, C., Bulmer, C., Hamilton, D., Gloury, J., O'Brien, N., Saunders, J., Harvey, S. and Patterson, T. (2021) The effect of mental fatigue on the performance of Australian football specific skills amongst amateur athletes. *Journal of Science and Medicine in Sport* **24(6)**, 592-596. https://doi.org/10.1016/j.jsams.2020.12.003

Williams, A. M., Thomas, J. L., Jordet, G. and Ford, P. R. (2023) Anticipation and decision-making. In *Science and Soccer* (pp. 124-140). Routledge. https://doi.org/10.4324/9781003148418-10

Xu, K., Blazevich, A. J., Boullosa, D., Ramirez-Campillo, R., Yin, M., Zhong, Y., Tian, Y., Finlay, M., Byrne, P. J. and Cuenca-Fernández, F. (2025) Optimizing post-activation performance enhancement in athletic tasks: A systematic review with meta-analysis for prescription variables and research methods. *Sports Medicine*, 1-32. https://doi.org/10.1007/s40279-024-02170-6

Young, W., Rayner, R. and Talpey, S. (2021) It's time to change direction on agility research: A call to action. *Sports Medicine - Open* **7(1)**, 12. https://doi.org/10.1186/s40798-021-00304-y

Young, W. and Rogers, N. (2014) Effects of small-sided game and change-of-direction training on reactive agility and change-of-direction speed. *Journal of Sports Sciences* **32(4)**, 307-314. https://doi.org/10.1080/02640414.2013.823230

Young, W. B., Dawson, B. and Henry, G. J. (2015) Agility and change-of-direction speed are independent skills: Implications for training for agility in invasion sports. *International Journal of Sports Science and Coaching* **10(1)**, 159-169. https://doi.org/10.1260/1747-9541.10.1.159

Zago, M., Giuriola, M. and Sforza, C. (2016) Effects of a combined technique and agility program on youth soccer players' skills. *International Journal of Sports Science and Coaching* **11(5)**, 710-720. https://doi.org/10.1177/1747954116667109

Zhang, M., Li, F., Jiao, J., Liang, W., Gomez, M.-A. and Scanlan, A. T. (2025) Effects of different training methods on open-skill and closed-skill agility in basketball players: A systematic review and meta-analysis. *Sports Medicine - Open* **11(1)**, 50. https://doi.org/10.1186/s40798-025-00842-9

Zhou, Z., Xin, C., Zhao, Y. and Wu, H. (2024) The effect of multi-directional sprint training on change-of-direction speed and reactive agility of collegiate tennis players. *PeerJ* **12**, e18263. https://doi.org/10.7717/peerj.18263

Zouhal, H., Abderrahman, A. B., Dupont, G., Truptin, P., Le Bris, R., Le Postec, E., Sghaeir, Z., Brughelli, M., Granacher, U. and Bideau, B. (2019) Effects of neuromuscular training on agility performance in elite soccer players. *Frontiers in Physiology* **10**, 947. https://doi.org/10.3389/fphys.2019.00947

Zwierko, M., Jedziniak, W., Popowczak, M. and Rokita, A. (2023) Effects of in-situ stroboscopic training on visual, visuomotor and reactive agility in youth volleyball players. *PeerJ* **11**, e15213. https://doi.org/10.7717/peerj.15213

Zwierko, M., Jedziniak, W., Popowczak, M. and Rokita, A. (2024) Effects of six-week stroboscopic training program on visuomotor reaction speed in goal-directed movements in young volleyball players: A study focusing on agility performance. *BMC Sports Science, Medicine and Rehabilitation* **16(1)**, 59. https://doi.org/10.1186/s13102-024-00848-y

Zwierko, T., Tapia, V., Vera, J., Redondo, B., Morenas-Aguilar, M. D. and García-Ramos, A. (2024) Enhancing reactive agility in soccer: The impact of stroboscopic eyewear during warm-up across fatigued and non-fatigued conditions. *European Journal of Sport Science* **24(12)**, 1798-1808. https://doi.org/10.1002/ejsc.12224

Key points

- Agility training yields large improvements in RAT and medium effects on PAT among team-sport players, with no significant gains in reaction test (RT). Training effects are larger in junior and national-level athletes.
- Perception-action coupling is crucial in agility training. While practitioners can use visual stimulus for controlled agility training gain, using human stimulus is encouraged to provide more sport-specific stimulus.
- 7 - 8 weeks agility training at a frequency higher than 3 times per week and a session duration of 20-25 minutes is associated with favorable adaptations in team-sport players.

AUTHOR BIOGRAPHY

Zhiwei ZHAO
Employment
School of Physical Education and Sports, Central China Normal University, Wuhan, China
Degree
MEd
Research interests
Agility training; Soccer; Strength and Conditioning
E-mail: zhou4210elliot@gmail.com

Yuhang LIU
Employment
School of Physical Education, Wuhan Sports University, Wuhan 430079, China
Degree
PhD
Research interests
Public health; Strength and conditioning
E-mail: liuyuhangcoach@163.com

Kai XU
Employment
School of Athletic Performance, Shanghai University of Sport, Shanghai, China.
Degree
PhD
Research interests
High-intensity Interval Training; Energy Contribution; Strength and conditioning
E-mail: 2221152066@sus.edu.cn

✉ **Kai Xu**
School of Medical and Health Sciences, Centre for Human Performance, Edith Cowan University, Joondalup, 6027, Australia

Supplementary Materials

Table S1. Search strategy.

Databases		Search Strategy
WOS	#1	ALL = (agility OR quickness OR change-of-direction OR reactive OR plyometric OR neuromuscular OR stimulu* OR SSG OR small-sided-game OR dual-task OR cognitive OR perceptual) AND (training OR intervention OR program*)
	#2	ALL = (athlete OR player OR recreational OR elite OR professional OR trained) AND (soccer OR football OR rugby OR basketball OR volleyball OR team-sport)
	#3	#1 AND #2
PUBMED	#1	(agility[Title/Abstract] OR quickness[Title/Abstract] OR change-of-direction[Title/Abstract] OR reactive[Title/Abstract] OR plyometric[Title/Abstract] OR neuromuscular[Title/Abstract] OR stimulu*[Title/Abstract] OR SSG[Title/Abstract] OR small-sided-game[Title/Abstract] OR dual-task[Title/Abstract] OR cognitive[Title/Abstract] OR perceptual[Title/Abstract]) AND (training[Title/Abstract] OR intervention[Title/Abstract] OR program*[Title/Abstract])
	#2	(athlete[Title/Abstract] OR player[Title/Abstract] OR recreational[Title/Abstract] OR elite[Title/Abstract] OR professional[Title/Abstract] OR trained[Title/Abstract]) AND (soccer[Title/Abstract] OR football[Title/Abstract] OR rugby[Title/Abstract] OR basketball[Title/Abstract] OR volleyball[Title/Abstract] OR team-sport[Title/Abstract])
	#3	#1 AND #2
Scopus	#1	TITLE-ABS-KEY ((agility OR quickness OR change-of-direction OR reactive OR plyometric OR neuromuscular OR stimulu* OR ssg OR small-sided-game OR dual-task OR cognitive OR perceptual) AND (training OR intervention OR program*))
	#2	TITLE-ABS-KEY ((athlete OR player OR recreational OR elite OR professional OR trained) AND (soccer OR football OR rugby OR basketball OR volleyball OR team-sport))
	#3	#1 AND #2

Table S2. Inclusion and exclusion criteria.

Complete eligibility criteria for systematic review
Inclusion criteria
P (Population) Healthy team-sport athlete, with no age or sex restrictions.
I (Intervention) Interventions integrate perceptual stimuli and motor movement (e.g. sided-games, COD drill with perceptual stimulus and etc.).
C (Control) Active or passive control.
O (Outcome) These tests were evaluation metrics that showed at least one reactive agility and/or open-skill performance (tests with direction or movement change in response to stimulus and/or open-skill tests score).
S (Study design) Randomized controlled, parallel and crossover trials.
Exclusion criteria
P (Population) Studies with individual/closed-skill sports (tennis, track, etc.), special populations (e.g. referees, umpires, military officers), or injured population.
I (Intervention) Studies with interventions that are purely physical (plyometrics/COD without reacting to external stimuli) or purely computer-based/perceptual (no motor movement involved or restricted to finger movement).
C (Control) Studies without control.
O (Outcome) Studies fail to report at least one reactive agility or open-skill performance indicator. Or studies reporting physical and perceptual aspects of agility separately.
S (Study design) Study design other than randomized control trails.

Table S2-1. Exclusion details.

PICOS	Reasons for exclusion	Number of studies	Total exclusion
P (Population)	Individual sport athlete	12	23
	Referees/empires	5	
	Non-athlete	6	
I (Intervention)	Isolated physical training	104	173
	Isolated perceptual training	43	
	Irrelevant training modality	26	
C (Control)	No control group	11	11
O (Outcome)	Physical outcomes only	70	108
	Perceptual outcomes only	38	
S (Study design)	Non-randomized trail	17	17
Others	Full-text or English version not available	4	4

Table S3. Characteristics of included studies.

Study	No. of participants	Sex	Age (years)	Sport	Training status	Method	W/F/T	Stimuli	Outcomes
Serpell et al. 2011	EG = 7 CG = 7	Male	18-20	Rugby	International level	RCOD training program using sport-specific video based stimuli	3/2/15	visual	RCOD
Chaouachi et al. 2014	EG = 12 CG = 12	Male	14.2 ± 0.9	Soccer	International level	SSG using formats like 1 vs. 1, 2 vs. 2, and 3 vs. 3.	6/3/30	human	RCOD; RCOD with ball
Young & Rogers 2014	EG = 13 CG = 12	Male	17.4 ± 0.7	AFL	International level	SSG using 4 v 4 or 2 v 2 games on 20 × 23 m or 15 × 15 m fields	7/2/15	human	RCOD
Born et al. 2016	EG = 10 CG = 9	Male	14 ± 0.6	Soccer	National level	COD drills performed in response to a visual stimulus on the Speedcourt.	3/6/12	visual	RCOD
Chaalali et al. 2016	EG = 11 CG = 10	Male	14.5±0.9	Soccer	International level	COD training with human stimuli	6/5/25	human	RCOD; RCOD with ball
Engelbrecht et al. 2016	EG = 10; 9 CG = 7; 7	Male	19-23	Rugby	Developmental	Coach command drills, pairs working together, mirror drills, evasion drills, and small-sided games.	6/2/30	human	RCOD
Trecroci et al. 2016	EG = 20 CG = 15	NA	10.57 ± 0.26	Soccer	Developmental	SAQ training program	12/2/25	visual	RCOD
Zago et al. 2016	EG = 10 CG = 10	Male	11.5 ±0.27	Soccer	Developmental	Structured SSGs on artificial constraints using tapes (14x12 meter matrix)	22/2.5/15	human	LSPT
Vogt et al. 2018	EG = 11 CG = 11	Male	13.5 ± 1.0	Soccer	National level	A football-specific course involving dribbling, passing, jumping,and sprinting in Footbonaut	1/1/28	visual-auditory	RT; Passing Accuracy
Zouhal et al. 2019	EG = 10 CG = 10	Male	17.25 ± 0.72	Soccer	International level	RCOD training program using human stimuli	6/2/30	human	RCOD
Dunton et al. 2020	EG = 10 CG = 10	NA	19.3 ± 2.3	Soccer	Developmental	Soccer passing task using spatial occlusion goggles	2/2/60	visual	Passing error; reaction time
Hicheur et al. 2020	EG = 9 CG = 9	Male	14.5 ± 0.52	Soccer	International level	Soccer passing training using visually moving target visuoauditory feedback after each pass, provided only to the AF group	3/3/NA	visual; visual-auditory	Passing time
Arede et al. 2021	EG = 11 CG = 11	Mix	13.8 ± 1.7	Soccer/ Basketball/Handball	Developmental	BATAK PRO training program using visual cues	12/2/20	visual	RCOD;PAT
Ehmann et al. 2022	EG = 14 CG = 13	Mix	12.3 ± 0.7	Soccer	Developmental	Passing training in Footbonaut with illuminated ball machine and target field	5/2/20	visual-auditory	Passing score; Passing accuracy; RT
Hassan et al. 2022	EG = 10 CG = 10	NA	14.7 ± 0.73	Basketball	Developmental	RCOD training using FITLIGHT system	8/4/45	visual	RCOD; RCOD with ball; RT
Bekris et al. 2023	EG = 10 CG = 10	Male	11.55 ± 0.49	Soccer	Developmental	Dribbling exercises with COD under coaches' visual cues	8/3/30	visual	RCOD with ball
Gonzalo-Skok et al. 2023	EG = 11 CG = 7	Male	15.5 ± 0.8	Soccer	Developmental	Unilateral flywheel training with random, unanticipated, and unexpected executions of multidirectional movements	10/1/30	visual-auditory	RCOD; PAT

W: Weeks; *F:* Frequency; *T:* Session time; *EG:* Experimental Group; *CG:* Control Group; *RCOD:* Reactive Change-of-direction; *COD:* Change-of-direction; *SSG:* Small-sided Game; *SAQ:* Speed Agility and Quickness; *AFL:* Australian Football League; *LSPT:* Loughborough Soccer Passing Test; *PAT:* Pre-planned Agility Test; *RT:* Reaction Test

Table S3. Continue…

Study	No. of participants	Sex	Age (years)	Sport	Training status	Method	W/F/T	Stimuli	Outcomes
Gonzalo-Skok et al. 2023	EG = 11 CG = 7	Male	15.5 ± 0.8	Soccer	Developmental	Unilateral flywheel training with random, unanticipated, and unexpected executions of multidirectional movements	10/1/30	visual-auditory	RCOD; PAT
Lucia et al. 2023	EG = 15 CG = 15	Mix	15-17	Basketball	National level	Cognitive-Motor Dual-Task Training (CMDT) integrated into Multicomponent Training (MCT)	5/2/30	visual-auditory	RCOD
Zwierko et al. 2023	EG = 25 CG = 25	Mix	16.5 ± 0.6	Volleyball	National level	Stroboscopic eyewear during volleyball-specific training tasks	6/3/30	visual	RCOD; RT
Friebe et al. 2024	EG = 15 CG = 14	Male	27±6	Soccer	Developmental	Agility with integrated multiple-object tracking training: Using SKILLCOURT system	6/2/40	visual-auditory	RCOD; RCOD with ball; LSPT
Lee et al. 2024	EG = 9 CG = 10	Female	18.89 ± 0.80	Soccer	Developmental	Technical and tactical (dribbling and SSG) training using human stimuli	8/3/40	human	RCOD; PAT
Lucia et al. 2024	EG = 15 CG = 15	Mix	16.1 ± 1.1	Basketabll	National level	Dual-task training (CMDT) integrated into 30-minute physical training sessions.	5/6/30	visual;visual-auditory	RCOD
Mancini et al. 2024	EG = 12 CG = 12	Female	20.35 ± 1.03	Volleyball	National level	Choosing signals with ReactionX technological system with up to eight wireless LED lights.	6/6/30	visual	Tapping with stimuli; RT
Steff et al. 2024	EG = 18;17 CG = 18;17	Male	14-16	Basketball	National level	Using FITLIGHT system during technical and tactical training (including positioning and COD)	18/3/60	visual	RCOD; RCOD with ball; PAT; PAT with ball
Zwierko et al., 2024 (a	EG = 25 CG = 25	Mix	16.5 ± 0.6	Volleyball	National level	Use of stroboscopic glasses (Senaptec Strobe) during volleyball-specific training exercises.	6/3/30	visual	RCOD; PAT; RT
Zwierko et al., 2024 (b	EG = 11 CG = 11	Male	23.6 ± 4.4	Soccer	Developmental	Stroboscopic eyewear during the ball-specific phase of soccer warm-up	1/1/20	visual	RCOD; RCOD with ball
Klatt et al. 2025	EG = 11 CG = 11	Male	10.98 ± 0.26	Soccer	Developmental	Soccer-specific training sessions using peripheral color cues to develop and practice perceptual-cognitive skills.	6/4/20	visual	RCOD; RCOD with ball
Neag et al. 2025	EG = 16 CG = 15	Male	10.76 ± 0.42	Soccer	Developmental	Fixed-Role Small-Sided Games (FRSSGs)	18/3/20	human	RCOD; RCOD with ball; PAT; PAT with ball; RT

W: Weeks; *F:* Frequency; *T:* Session time; *EG:* Experimental Group; *CG:* Control Group; *RCOD:* Reactive Change-of-direction; *COD:* Change-of-direction; *SSG:* Small-sided Game; *SAQ:* Speed Agility and Quickness; *AFL:* Australian Football League; *LSPT:* Loughborough Soccer Passing Test; *PAT:* Pre-planned Agility Test; *RT:* Reaction Test

Table S4. **Training load data.**

Study	Sport	Sets/ Bouts	Reps	Intensity	Description
Serpell et al. 2011	Rugby	NA	10	NA	Each reactive agility training session involved participants completing 10 perception action guided discovery reactive agility drills per session
Chaouachi et al. 2014	Soccer	COD:2	2-4	Session-RPE	Detailed in Table 1 and Table 2of the article
Zwierko et al. 2023	EG = 25 CG = 25	SSG:1-2	30s-2min	Volleyball	National level
Young & Rogers 2014	AFL	COD:4	36-48	NA	COD: 4 activities per session to provide variety. The number of changes of direction per session progressed from 36 at the start of the programme to 48 by the end.
Lee et al. 2024	EG = 9 CG = 10	SSG:3-4	30-45s	Soccer	SSG:All games were 30–45 s bouts with four sessions in first and second games and three sessions in last game.
Born et al. 2016	Soccer	4	5	NA	4 sets of 5 sprints, which were separated by 5 min of active rest.
Chaalali et al. 2016	Soccer	2-3	1	NA	Detailed in Table 1 of the article
Engelbrecht et al. 2016	Rugby	3-4	NA	NA	Detailed in Table 3 of the article without exact number of repetitions. Intensity was only suggested qualitatively (e.g. low-moderate intensity)
Trecroci et al. 2016	Soccer	4	120s	NA	Detailed in Table 1 and Table 2 of the article
Zago et al. 2016	Soccer	8-20	4-15s	NA	Detailed in Table 2 with qualitatively suggested intensity (maximal or almost maximal)
Vogt et al. 2018	Soccer	4	4min	90-95%HR	4 intervals of 4min in the course with 3min of active recov- ery between intervals. Heart rates (HR) were consistently monitored and recorded after each in terval. Monitoring HR served as a control to target 90 to 95% of the maximum HR (HRmax) in accordance with previ ously measured HR dur- ing the Yo-Yo intermittent recovery test
Zouhal et al. 2019	Soccer	NA	NA	NA	Training sessions varied greatly within study, intensity was only qualitatively stated (light plyometric intensity initially; progressed gradually; exercises performed at maxi- mal effort during COD drills)
Dunton et al. 2020	Soccer	4	15 passes	NA	The 60 passes were divided in to four blocks of 15 passes.
Hicheur et al. 2020	Soccer	1	32 passes	NA	During each training session, players of the AF group and the NF group also performed 32 passes. Intensity was mentioned qualitatively ("pass the ball as accurately and as quickly as possible")
Arede et al. 2021	Soccer/ Basketball/ Handball	2-3	4-5	NA	The training protocol consisted of 20 minutes sessions, 2–3 sets of 4–5 exercises using touch-board and LED light- ing equipment
Ehmann et al. 2022	Soccer	3	6 min	NA	Each training session was divided into three 6-min training blocks with a 1-min break between each block. Only “adap- tive speed adjustment” was suggested for intensity
Hassan et al. 2022	Basketball	2-3	2-8	NA	The researchers used load intensity (less than maximum in- tensity–maximum intensity). The training method used over the period was high intensity. The number of duplicates was 2–8. The number of groups was 2–3.
Bekris et al. 2023	Soccer	6	5 min	NA	Detailed in Table 1 of the article
Gonzalo-Skok et al. 2023	soccer	1	4	NA	Both UTG and PTG consisted of 1 set of 4 unilateral exercises: front step, backward lunges, defensive-like shuf- fling steps, and lateral crossover cutting
Lucia et al. 2023	Basketball	4	1	NA	Two 30-min sessions per week, consisting of four exercise circuits repeated once each. Intensity was manipulated by task difficulty.
Zwierko et al. 2023	Volleyball	NA	NA	NA	Volume was only provided in minutes per session. Intensity was implied (stroboscopic protocols were restricted to 2.5 min with a 2.5-min break interval)
Friebe et al. 2024	Soccer	NA	NA	Borg scale	Only training hours per week is provided
Lee et al. 2024	Soccer	NA	NA	NA	Only session duration is provided. Intensity was monitored but not recorded ("However, we did not document these records as they were not within the scope of the study.")

Reps: Repetitions performed in a training session; ***COD:*** Change-of-direction; ***RPE:*** Rating of Perceived Exertion; ***AFL:*** Australian rule football; ***NA:*** Not applicable; ***HR:*** Heart rate; ***SSG:*** Small-sided games; ***EG:*** Experimental group; ***CG:*** Control group

Table S4. Continue…

Study	Sport	Sets/ Bouts	Reps	Intensity	Description
Lucia et al. 2024	Basketball	4	1	NA	Four circuits repeated once with two exercises, one on skills and the other on stability
Mancini et al. 2024	Volleyball	NA	NA	NA	Only duration was provided in Table 2 of the article. Intensity was qualitatively stated (",efforts were made to ensure that there was no overall difference in the training volume between the EG and CG")
Steff et al. 2024	Basketball	NA	NA	NA	Training sessions varied greatly within study, intensity was only qualitatively stated ("players performed all exercises at maximal speed and with rapid reaction to light cues.")
Zwierko et al., 2024 (a	Volleyball	NA	NA	NA	Volume was only provided in minutes per session. Intensity manipulated only by visual difficulty (strobe flicker rate and duty cycle), not by physiological exertion.
Zwierko et al., 2024 (b	Soccer	NA	NA	17.0 ± 2.0 (RPE)	Training sessions varied greatly within study.
Klatt et al. 2025	Soccer	NA	NA	NA	Only session duration is provided. Intensity was manipulated through task difficulty ("Visual training can lead to mental fatigue as higher perceptual loads affect information processing and may hinder effectiveness")
Neag et al. 2025	Soccer	1-2	4-6	NA	Intensity was stated qualitatively (e.g. dribbling at maximum speed)

Reps: Repetitions performed in a training session; ***COD:*** Change-of-direction; ***RPE:*** Rating of Perceived Exertion; ***AFL:*** Australian rule football; ***NA:*** Not applicable; ***HR:*** Heart rate; ***SSG:*** Small-sided games; ***EG:*** Experimental group; ***CG:*** Control group

Table S5. PEDro scale.

Study	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Total
Serpell et al. 2011	0	2	2	2	2	2	2	1	2	2	17
Chaouachi et al. 2014	1	1	2	1	2	2	2	1	2	2	16
Young & Rogers 2014	2	1	2	2	1	2	2	1	2	2	17
Born et al. 2016	1	1	2	2	2	2	2	1	2	2	17
Chaalali et al. 2016	1	1	2	1	2	2	2	1	2	2	16
Engelbrecht et al. 2016	1	1	2	2	2	2	2	1	2	2	17
Trecroci et al. 2016	1	2	2	2	2	2	2	2	2	2	19
Zago et al. 2016	1	1	2	2	2	2	2	2	2	2	18
Vogt et al. 2018	1	1	2	2	2	2	2	0	2	2	16
Zouhal et al. 2019	1	1	2	2	2	2	2	1	2	2	17
Dunton et al. 2020	1	1	2	1	2	2	2	1	2	2	16
Hicheur et al. 2020	1	1	2	2	2	2	2	1	2	2	17
Arede et al. 2021	1	1	2	2	2	2	2	2	2	2	18
Ehmann et al. 2022	1	1	2	2	2	2	2	1	2	2	17
Hassan et al. 2022	0	1	2	2	2	2	2	2	2	2	17
Bekris et al. 2023	1	1	2	1	2	2	2	1	2	2	16
Gonzalo-Skok et al. 2023	2	2	2	2	2	2	2	2	2	2	20
Lucia et al. 2023	2	2	2	2	2	2	2	1	2	2	19
Zwierko et al. 2023	2	1	2	2	2	2	2	2	2	2	19
Friebe et al. 2024	2	2	2	2	2	2	2	2	2	2	20
Lee et al. 2024	1	1	2	2	2	2	2	1	2	2	17
Lucia et al. 2024	2	1	2	2	2	2	2	1	2	2	18
Mancini et al. 2024	2	2	2	2	2	2	2	1	2	2	19
Steff et al. 2024	1	1	2	1	2	2	2	2	2	2	17
Zwierko et al., 2024 (a	1	1	2	2	2	2	2	1	2	2	17
Zwierko et al., 2024 (b	2	1	2	1	2	2	2	1	2	2	17
Klatt et al. 2025	1	1	2	2	2	2	2	1	2	2	17
Neag et al. 2025	2	1	2	0	2	2	2	2	2	2	17

D1 = Inclusion criteria stated
D2 = Subjects assigned appropriately (randomized, ability level)
D3 = Intervention described (protocols equated)
D4 = Dependent variables defined (reliable outcome measures)
D5 = Assessments practical (easy to implement)
D6 = Training duration practical
D7 = Appropriate statistics (normality, significant differences)
D8 = Results detailed (mean, standard deviation, percent change, effect size)
D9 = Conclusions concise (clear, concise, future directions)
D10 = point measures of variability.

Table S6. Pre-test baseline difference.

Study	N1	Mean1	SD1	N2	Mean2	SD2	t	p	df
Serpell et al. 2011	7	-1.89	0.16	7	-1.92	0.17	0.34	0.74	11.956
Chaouachi et al. 2014	12	-2.27	0.08	12	-2.29	0.09	0.575	0.571	21.702
Chaouachi et al. 2014	12	-2.65	0.09	12	-2.68	0.08	0.863	0.398	21.702
Young & Rogers 2014	13	-2.64	0.13	12	-2.56	0.1	-1.732	0.097	22.313
Born et al. 2016	10	-15.7	1.5	9	-15.8	1.2	0.161	0.874	16.796
Born et al. 2016	10	-16.2	1.3	9	-15.9	0.8	-0.612	0.549	15.15
Born et al. 2016	10	-15.4	0.9	9	-15.8	1.1	0.862	0.402	15.532
Born et al. 2016	10	-15.5	1.2	9	-15.5	0.8	0	1	15.76
Chaalali et al. 2016	11	-2.15	0.09	10	-2.12	0.08	-0.809	0.429	18.994
Chaalali et al. 2016	11	-2.56	0.07	10	-2.57	0.16	0.182	0.858	12.076
Engelbrecht et al. 2016	10	-2.79	0.18	7	-2.75	0.16	-0.482	0.637	14.01
Engelbrecht et al. 2016	9	-2.88	0.13	7	-2.75	0.16	-1.747	0.107	11.474
Trecroci et al. 2016	20	-2.85	0.15	15	-2.78	0.19	-1.178	0.25	25.967
Zago et al. 2016	10	-51.35	3.91	10	-50.22	4.11	-0.63	0.537	17.955
Zago et al. 2016	10	-17.2	8	10	-20.4	5.3	1.054	0.308	15.624
Zago et al. 2016	10	-68.67	10.99	10	-60.72	9.18	-1.756	0.097	17.447
Vogt et al. 2018	11	-0.92	0.08	11	-0.96	0.08	1.173	0.255	20
Vogt et al. 2018	11	-4.25	0.24	11	-4.2	0.21	-0.52	0.609	19.654
Zouhal et al. 2019	10	-0.385	0.03	10	-0.381	0.025	-0.324	0.75	17.433
Zouhal et al. 2019	10	-1.451	0.08	10	-1.468	0.07	0.506	0.619	17.688
Dunton et al. 2020	*10*	*-18.5*	*5.8*	*10*	*-11*	*9.4*	*-2.147*	*0.049*	*14.985*
Hicheur et al. 2020	9	-0.943	0.68	9	-0.888	0.131	-0.238	0.817	8.593
Hicheur et al. 2020	9	-0.902	0.105	9	-0.888	0.131	-0.25	0.806	15.276
Arede et al. 2021	11	-1.92	0.16	11	-2.03	0.9	0.399	0.698	10.631
Ehmann et al. 2022	14	47.38	8.59	13	47.69	9.85	-0.087	0.931	23.921
Ehmann et al. 2022	14	52.14	14.9	13	50.31	17.1	0.296	0.77	23.913
Ehmann et al. 2022	14	-3.22	0.29	13	-3.35	0.25	1.25	0.223	24.875
Ehmann et al. 2022	14	47.38	8.59	13	50	10.46	-0.708	0.486	23.301
Ehmann et al. 2022	14	52.14	14.9	13	59.77	13.98	-1.373	0.182	24.996
Ehmann et al. 2022	14	-3.22	0.29	13	-3.31	0.29	0.806	0.428	24.852
Bekris et al. 2023	10	-31.1	2.1	10	-30.4	1	-0.952	0.359	12.882
Gonzalo-Skok et al. 2023	11	-3.14	0.08	7	-3.06	0.09	-1.918	0.08	11.766
Gonzalo-Skok et al. 2023	*11*	*-3.27*	*0.06*	*7*	*-3.18*	*0.08*	*-2.554*	*0.028*	*10.274*
Gonzalo-Skok et al. 2023	*11*	*-3.38*	*0.07*	*7*	*-3.29*	*0.09*	*-2.248*	*0.047*	*10.569*
Lucia et al. 2023	15	-5.92	0.34	15	-5.9	0.3	-0.171	0.866	27.573
Zwierko et al. 2023	25	-18.18	1.23	25	-18.61	1.09	1.308	0.197	47.316
Friebe et al. 2024	15	-11.4	0.81	14	-11.4	0.95	0	1	25.658
Friebe et al. 2024	15	-41.7	2.53	14	-40.7	3.12	-0.944	0.354	25.081
Friebe et al. 2024	*15*	*-25.1*	*1.44*	*14*	*-23.7*	*1.39*	*-2.664*	*0.013*	*26.965*
Friebe et al. 2024	15	-5.1	0.36	14	-5.2	0.52	0.598	0.556	22.965
Friebe et al. 2024	15	-16.8	0.72	14	-16.7	0.95	-0.318	0.753	24.213
Lee et al. 2024	9	-2.28	0.86	10	-2.28	0.09	0	1	8.158
Lucia et al. 2024	15	-5.64	0.35	15	-5.63	0.34	-0.079	0.937	27.977
Mancini et al. 2024	12	-5.86	0.244	12	-5.801	0.266	-0.566	0.577	21.838
Mancini et al. 2024	12	-9.153	0.489	12	-9.103	0.444	-0.262	0.796	21.798
Zwierko et al. 2024 (a	25	-18.18	1.23	25	-18.61	1.09	1.308	0.197	47.316
Zwierko et al., 2024 (b	11	-49.9	4.2	11	-50.7	4	0.457	0.652	19.953
Zwierko et al., 2024 (b	11	-59	2.9	11	-61	3.9	1.365	0.189	18.469
Klatt et al. 2025	11	-3.5	0.35	11	-3.47	0.25	-0.231	0.82	18.096
Klatt et al. 2025	11	-4.55	0.52	11	-4.55	0.41	0	1	18.968
Neag et al. 2025	16	-2.68	0.14	15	-2.6	0.13	-1.65	0.11	28.998
Neag et al. 2025	16	-3.16	0.3	15	-3.08	0.21	-0.864	0.395	26.903
Neag et al. 2025	16	-9.74	1.01	15	-9.33	0.83	-1.238	0.226	28.531

RAT: Reactive agility test; ***N1:*** Participants in experimental group; ***Mean1:*** Pre-test mean value in experimental group; ***SD1:*** Pre-test standardised deviation in experimental group; ***N2:*** Participants in control group; ***Mean2:*** Pre-test mean value in control group; ***SD2:*** Pre-test standardised deviation in control group; ***t:*** unpaired t-test results; ***p:*** significance value of the t-test; ***df:*** degree of freedom

Table S7. **Pre-test baseline difference for PAT.**

Study	N1	Mean1	SD1	N2	Mean2	SD2	t	p	df
Serpell et al. 2011	7	-1.61	0.12	7	-1.64	0.15	0.41	0.69	11.45
Chaouachi et al. 2014	12	-3.62	0.19	12	-3.68	0.36	0.51	0.62	16.69
Chaouachi et al. 2014	12	-4.82	0.26	12	-4.92	0.43	0.69	0.50	18.09
Chaouachi et al. 2014	12	-10.67	0.67	12	-10.79	0.67	0.44	0.67	22.00
Chaouachi et al. 2014	12	-7.70	0.28	12	-7.74	0.32	0.33	0.75	21.62
Young & Rogers 2014	13	-8.67	0.29	12	-8.65	0.45	-0.13	0.90	18.55
Born et al. 2016	10	-17.80	0.30	9	-18.20	0.90	1.27	0.23	9.59
Chaalali et al. 2016	11	-3.52	0.12	10	-3.48	0.06	-0.98	0.34	15.00
Chaalali et al. 2016	11	-4.45	0.27	10	-4.55	0.21	0.95	0.35	18.59
Chaalali et al. 2016	11	-2.49	0.17	10	-2.52	0.14	0.44	0.66	18.84
Engelbrecht et al. 2016	10	-1.97	0.08	7	-2.04	0.14	1.19	0.26	8.75
Engelbrecht et al. 2016	9	-2.02	0.14	7	-2.04	0.14	0.28	0.78	13.04
Trecroci et al. 2016	20	-1.53	0.08	15	-1.49	0.07	-1.57	0.13	32.14
Trecroci et al. 2016	***20***	***-4.27***	***0.24***	***15***	***-4.11***	***0.22***	***-2.05***	***0.05***	***31.60***
Trecroci et al. 2016	20	-13.35	0.66	15	-13.10	0.69	-1.08	0.29	29.55
Zago et al. 2016	10	-8.44	0.26	10	-8.26	0.40	-1.19	0.25	15.45
Zago et al. 2016	10	-10.13	0.48	10	-10.27	0.44	0.68	0.51	17.87
Zago et al. 2016	10	-24.17	1.87	10	-25.17	1.24	1.41	0.18	15.63
Arede et al. 2021	11	-2.61	0.19	11	-2.66	0.09	0.79	0.44	14.27
Gonzalo-Skok et al. 2023	11	-9.26	0.30	7	-9.31	0.15	0.47	0.65	15.43
Friebe et al. 2024	15	-10.60	0.54	14	-10.70	0.87	0.37	0.72	21.45
Lee et al. 2024	***9***	***-9.41***	***0.24***	***10***	***-9.09***	***0.29***	***-2.63***	***0.02***	***16.90***
Lee et al. 2024	9	-9.32	0.27	10	-9.12	0.21	-1.79	0.09	15.10
Lee et al. 2024	9	-10.96	0.28	10	-11.14	0.58	0.87	0.40	13.26
Lee et al. 2024	9	-10.99	0.42	10	-10.68	0.44	-1.57	0.13	16.93
Steff et al. 2024	***18***	***-12.35***	***1.22***	***18***	***-13.32***	***0.98***	***2.63***	***0.01***	***32.49***
Steff et al. 2024	***18***	***-17.13***	***2.12***	***18***	***-18.61***	***2.03***	***2.14***	***0.04***	***33.94***
Steff et al. 2024	18	-17.83	1.52	18	-18.68	1.18	1.87	0.07	32.03
Steff et al. 2024	18	-20.41	2.45	18	-21.09	1.44	1.02	0.32	27.49
Steff et al. 2024	***17***	***-12.41***	***1.35***	***17***	***-10.98***	***0.36***	***-4.22***	***0.00***	***18.26***
Steff et al. 2024	***17***	***-14.02***	***2.23***	***17***	***-12.03***	***0.67***	***-3.52***	***0.00***	***18.87***
Steff et al. 2024	***17***	***-19.66***	***1.34***	***17***	***-16.22***	***0.95***	***-8.63***	***0.00***	***28.84***
Steff et al. 2024	***17***	***-20.82***	***1.61***	***17***	***-16.84***	***0.97***	***-8.73***	***0.00***	***26.26***
Zwierko et al. 2024 (a	25	-14.54	0.96	25	-15.01	0.92	1.77	0.08	47.91
Neag et al. 2025	16	-2.92	0.11	15	-2.96	0.19	0.71	0.48	22.14
Neag et al. 2025	***16***	***-6.22***	***0.26***	***15***	***-5.91***	***0.37***	***-2.68***	***0.01***	***24.97***
Neag et al. 2025	16	-8.92	0.62	15	-8.51	0.62	-1.84	0.08	28.87

PAT: Pre-planned agility test; ***N1:*** Participants in experimental group; ***Mean1:*** Pre-test mean value in experimental group; ***SD1:*** Pre-test standardised deviation in experimental group; ***N2:*** Participants in control group; ***Mean2:*** Pre-test mean value in control group; ***SD2:*** Pre-test standardised deviation in control group; ***t:*** unpaired t-test results; ***p:*** significance value of the t-test; ***df:*** degree of freedom.

Table S8. **Pre-test baseline difference for RT.**

Study	N1	Mean1	SD1	N2	Mean2	SD2	t	p	df
Serpell et al. 2011	7	-0.34	0.20	7	-0.33	0.33	-0.07	0.95	9.88
Young & Rogers 2014	***13***	***-0.35***	***0.06***	***12***	***-0.28***	***0.07***	***-2.67***	***0.01***	***21.79***
Young & Rogers 2014	13	-0.97	0.08	12	-0.96	0.10	-0.27	0.79	21.09
Born et al. 2016	10	-0.26	0.10	9	-0.24	0.06	-0.71	0.49	15.19
Born et al. 2016	10	-0.27	0.09	9	-0.23	0.04	-1.41	0.18	12.77
Born et al. 2016	10	-0.26	0.10	9	-0.24	0.06	-0.52	0.61	14.54
Born et al. 2016	10	-0.26	0.08	9	-0.24	0.06	-0.48	0.64	16.55
Dunton et al. 2020	10	-0.98	0.13	10	-0.97	0.12	-0.20	0.85	17.97
Ehmann et al. 2022	14	-384.41	50.39	13	-384.62	51.15	0.01	0.99	24.79
Ehmann et al. 2022	14	-384.41	50.39	13	-365.75	48.53	-0.98	0.34	24.96
Hassan et al. 2022	10	-0.40	0.10	10	-0.41	0.09	0.21	0.83	17.80
Hassan et al. 2022	10	-0.48	0.09	10	-0.51	0.07	0.94	0.36	16.97
Zwierko et al. 2023	25	-270.96	44.51	25	-273.72	52.66	0.20	0.84	46.70
Mancini et al. 2024	12	-0.48	0.01	12	-0.48	0.01	0.51	0.62	20.09
Mancini et al. 2024	12	-0.61	0.03	12	-0.62	0.03	0.64	0.53	22.00
Zwierko et al. 2024 (a	25	-3.64	0.80	25	-3.60	0.80	-0.18	0.86	48.00
Neag et al. 2025	16	-0.50	0.03	15	-0.51	0.04	0.78	0.44	25.93
Neag et al. 2025	16	-0.51	0.04	15	-0.51	0.06	0.00	1.00	24.18

RT: Reaction test; ***N1:*** Participants in experimental group; ***Mean1:*** Pre-test mean value in experimental group; ***SD1:*** Pre-test standardised deviation in experimental group; ***N2:*** Participants in control group; ***Mean2:*** Pre-test mean value in control group; ***SD2:*** Pre-test standardised deviation in control group; ***t:*** unpaired t-test results; ***p:*** significance value of the t-test; ***df:*** degree of freedom.

Table S9. Comparison of model fit superiority.

Model		df	AIC	BIC	AICc	LogLik	p-value
RAT	Three-level meta-analysis	3	165.0635	171.2448	165.5079	-79.5317	NA
	Traditional meta-analysis	2	208.1994	212.3203	208.4176	-102.0997	**<0.001**
PAT	Three-level meta-analysis	3	121.4455	126.1961	122.1955	-57.7228	NA
	Traditional meta-analysis	2	253.1902	256.3572	253.5538	-124.5951	**<0.001**
RT	Three-level meta-analysis	3	44.4058	46.9054	46.2520	-19.2029	NA
	Traditional meta-analysis	2	42.4058	44.0722	43.2629	-19.2029	1

AIC: Akaike Information Criterion; ***BIC:*** Bayesian Information Criterion; ***AICc:*** Corrected AIC; ***LogLik:*** Log-Likelihood; ***RAT:*** Reactive agility test; ***PAT:*** Pre-planned agility test; ***RT:*** Reaction test; ***NA:*** not applicable

Table S10. Sensitivity analysis.

	Effect size				Test of heterogeneity				
Stage	**Model**	**ES**	**95%-CI**	**p-value**	**I^2-Level 2**	**I^2-Level 3**	**df**	**Q**	**p-value**
Pooled estimate for RAT at each stage									
B-Outliers	**REML**	0.79	0.53-1.06	< **0.001**	62.45%	14.98%	58	222.64	< **0.001**
	ML	0.79	0.53-1.05	< **0.001**	64.75%	11.91%	58	222.64	< **0.001**
Full	**REML**	0.65	0.42-0.87	< **0.001**	61.23%	0%	52	180.13	< **0.001**
	ML	0.65	0.42-0.87	< **0.001**	60.73%	0%	52	180.13	< **0.001**
Reduced	**REML**	0.66	0.42-0.89	< **0.001**	63.79%	0%	48	178.28	< **0.001**
	ML	0.66	0.42-0.89	< **0.001**	63.26%	0%	48	178.28	< **0.001**
Pooled estimate for PAT at each stage									
B-Outliers	**REML**	0.57	0.17-0.96	< **0.01**	87.32%	0%	36	249.65	< **0.001**
	ML	0.57	0.17-0.96	< **0.01**	86.95%	0%	36	249.65	< **0.001**
Full	**REML**	0.56	0.17-0.96	< **0.01**	87.32%	0%	36	249.65	< **0.001**
	ML	0.56	0.17-0.96	< **0.01**	86.95%	0%	36	249.65	< **0.001**
Reduced	**REML**	0.70	0.29-1.11	< **0.01**	13.65%	60.20%	27	94.41	< **0.001**
	ML	0.70	0.30-1.10	< **0.01**	14.36%	57.87%	27	94.41	< **0.001**
Change	**REML**	0.55	0.22-0.88	< **0.01**	63.23%	0%	36	103.06	< **0.001**
	ML	0.55	0.23-0.87	< **0.01**	61.19%	0%	36	103.06	< **0.001**
Pooled estimate for RT at each stage									
B-Outliers	**REML**	0.54	-0.12-1.21	> 0.05	0%	83.30%	17	82.24	< **0.001**
	ML	0.54	-0.09-1.17	> 0.05	0%	81.48%	17	82.24	< **0.001**
Full	**REML**	0.54	-0.12-1.21	> 0.05	0%	83.30%	17	82.24	< **0.001**
	ML	0.54	-0.09-1.17	> 0.05	0%	81.48%	17	82.24	< **0.001**
Reduced	**REML**	0.51	-0.17-1.19	> 0.05	0%	83.74%	16	81.00	< **0.001**
	ML	0.51	-0.14-1.15	> 0.05	0%	81.97%	16	81.00	< **0.001**

RAT: Reactive agility test; ***PAT:*** Pre-planned agility test; ***RT:*** Reaction test; ***B-Outliers:*** Before outliers were excluded; ***Full:*** Model fitted using data after influential studies excluded; ***Reduced:*** Model fitted after significant different baseline data was excluded; ***df:*** Degree of freedom; ***REML:*** Restricted maximum-likelihood; ***ML:*** maximum-likelihood; ***Change:*** Calculating effect size using change score; ***Level 2:*** Heterogeneity within the study; ***Level 3:*** Heterogeneity between the study

Table S11. Cluster-robust variance estimation.

Outcomes	r	ES	95%-CI	df	p-value
RAT	0.4	0.65	0.44-0.86	52	<0.001
	0.6	0.65	0.42-0.87	52	<0.001
	0.8	0.64	0.41-0.88	52	<0.001
	1	0.68	0.47-0.88	52	<0.001
PATP	0.4	0.55	0.11-0.99	36	< 0.05
	0.6	0.55	0.09-1.01	36	< 0.05
	0.8	0.54	0.07-1.02	36	< 0.05
	1	0.57	0.17-0.96	36	< 0.05
RT	0.4	0.53	-0.10-1.16	17	> 0.05
	0.6	0.52	-0.10-1.14	17	> 0.05
	0.8	0.51	-0.10-1.12	17	> 0.05
	1	0.54	-0.12-1.21	17	> 0.05

RAT: Reactive agility test; *PAT:* Pre-planned agility test; *RT:* Reaction test; ***r:*** Correlation coefficient; *ES:* Effect size

Table S12. Certainty of evidence.

Certainty Assessment						No. of participants			Certainty
N	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	EG	CG	ES(95%CI)	
RAT (follow-up: range 1 weeks to 22 weeks)									
26	serious	serious	not serious	not serious	not serious	351	332	ES 0.68 (0.47 to 0.88)	□□□◯ Moderate
PAT (follow-up: range 3 weeks to 22 weeks)									
15	serious	serious	not serious	not serious	not serious	224	206	ES 0.57 (0.17 to 0.96)	□□◯◯ Low
RT (follow-up: range 2 weeks to 18 weeks)									
10	serious	serious	serious	not serious	serious	142	138	ES 0.54 (0.12 to 1.21)	□◯◯◯ Very low

N: Number of studies; *EG:* Experimental group; *CG:* Control group; *RAT:* Reactive agility test; *PAT:* Pre-planned agility test; *RT:* Reaction test; *ES:* Effect size; *CI:* confidence interval

Table S13. Deviation from registered protocol.

Registered Plan		Deviation from Plan	
Section	**Description**	**Yes/No**	**Detailed Deviation**
Aims in protocol	1) to integrate reactive agility training methods	No	/
	2)to examine the effect of these training methods on reactive agility performance among team-sport athletes;	No	/
	3) to establish the optimal training modality in reactive agility training.	No	/
Effect size calculation	The individual effect size of the included studies will be estimated through Hedge's g formula.	No	/
Heterogeneity analysis	Heterogeneity will be assessed using the I^2 test and the heterogeneity will be reported with percentile reference values: (<25%) low, (25–75%) medium, and (>75%) high.	No	/
Publication bias detection	In order to check the publication bias of the included studies funnel plot asymmetry, and Egger's test will be applied.	Yes	An extended version of Egger's test was employed to address the nested data.
Sensitivity analysis	In addition, sensitivity analyses will be performed to check extreme values and to test the suitability for models.	Yes	Detailed outlier check methods were not stated in registration.

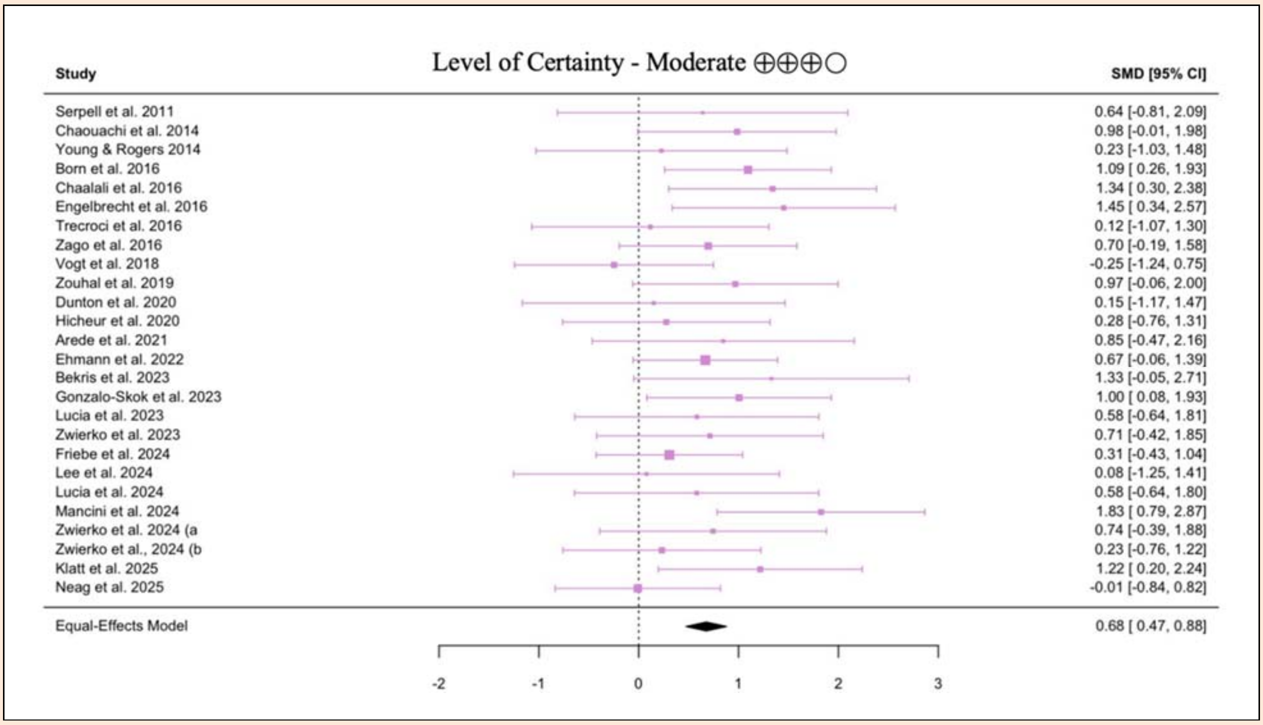

Figure S1. Main effects for RAT ($p < 0.01$). SMD: Standardized Mean Differences; CI: Confidence Interval, GRADE: Certainty of evidence.

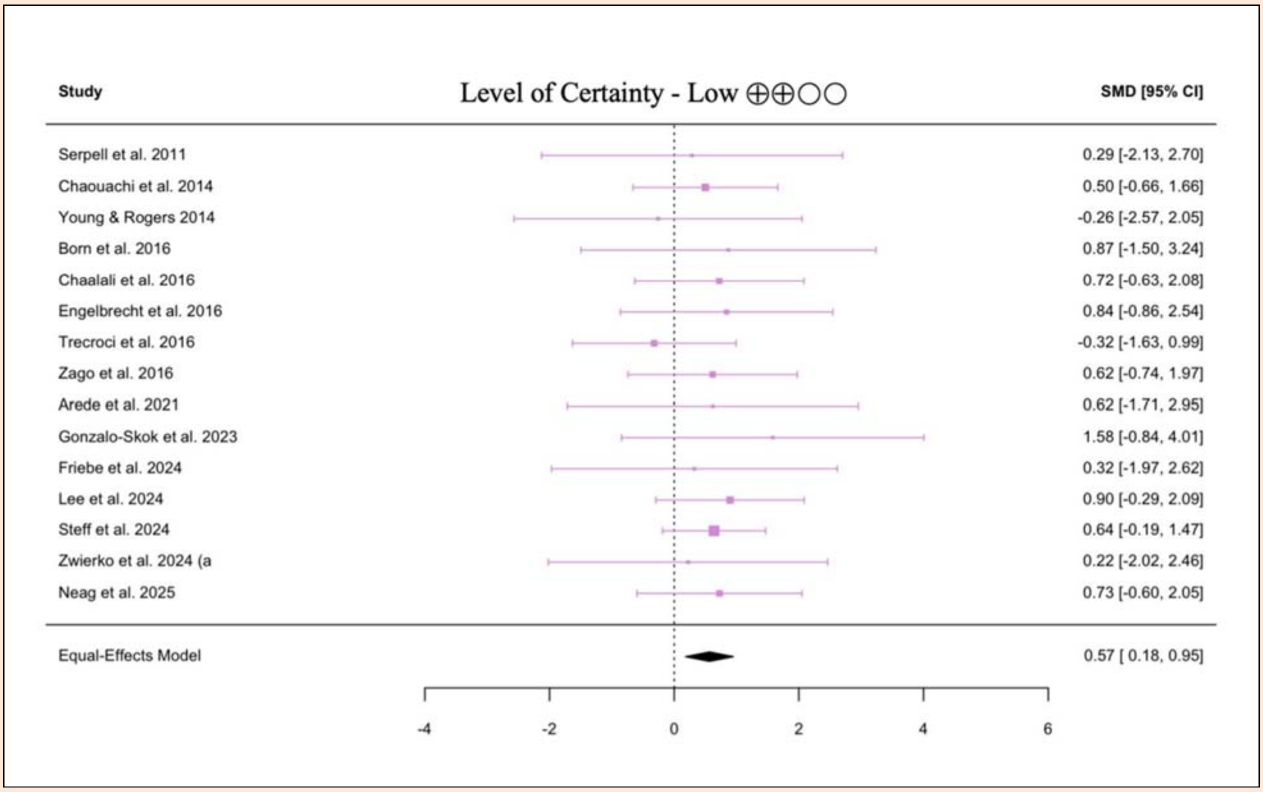

Figure S2. Main effects for PAT ($p < 0.01$). SMD: Standardized Mean Differences; CI: Confidence Interval, GRADE: Certainty of evidence.

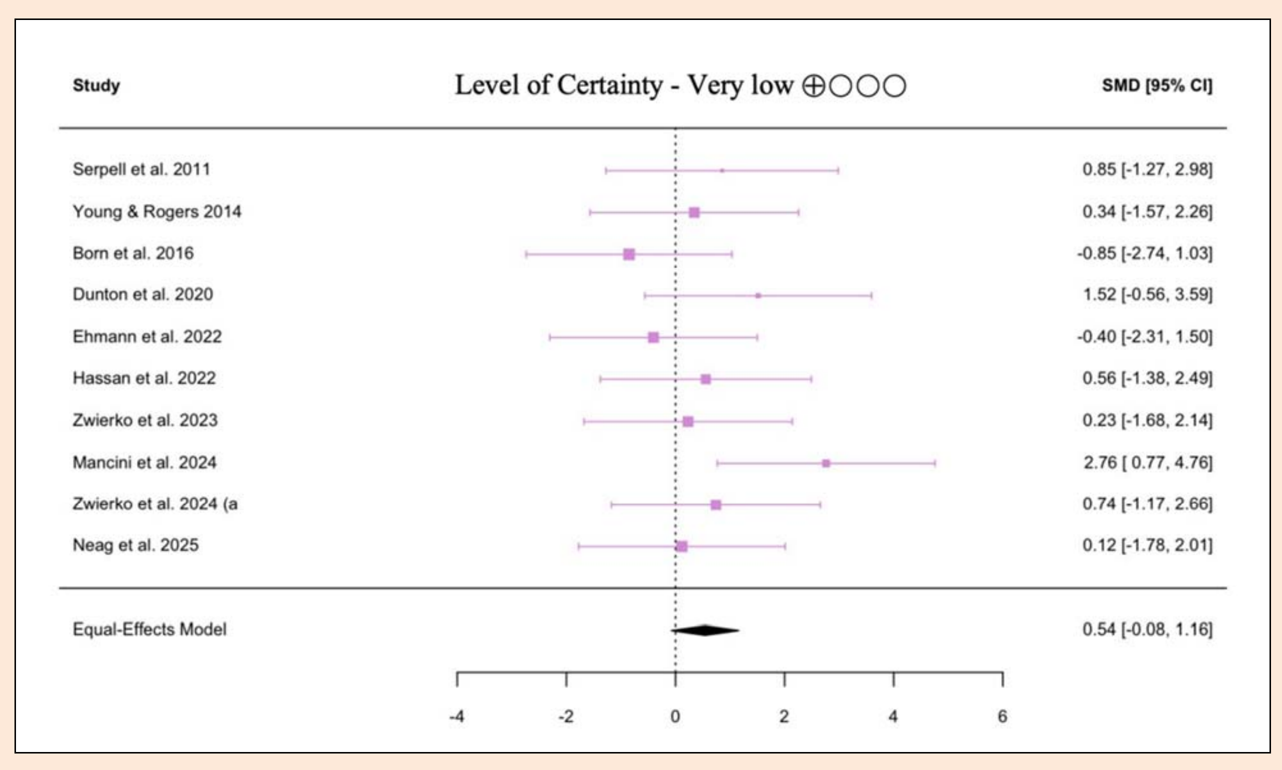

Figure S3. Main effects for RT (p > 0.05). SMD: Standardized Mean Differences; CI: Confidence Interval, GRADE: Certainty of evidence

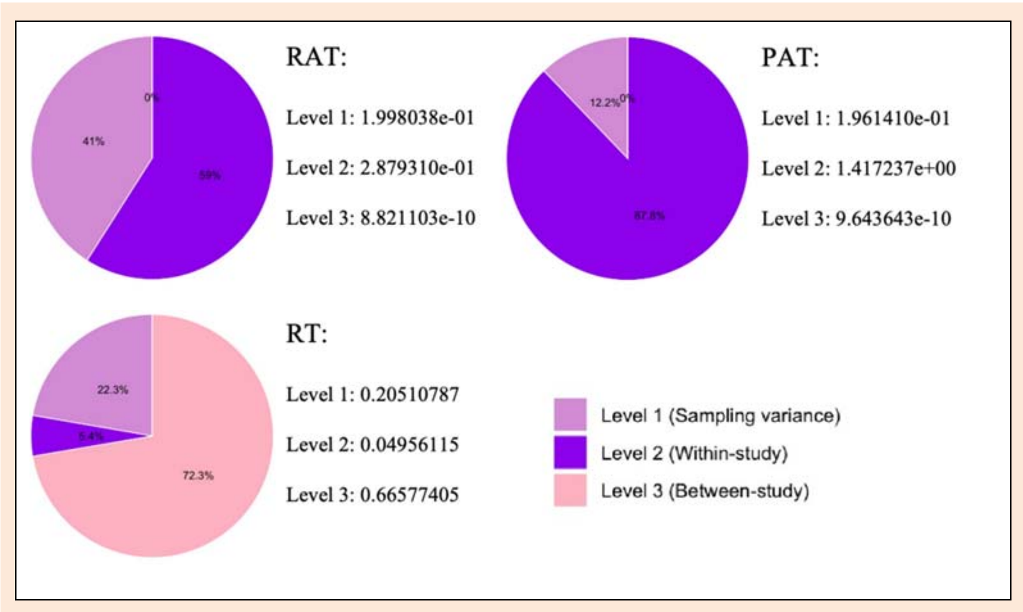

Figure S4. Variance-component pie chart of each level. *Level 1:* sampling variance extracted from vi; *Level 2:* within-study variance extracted from σ²; *Level 3*: between study variance extracted from σ².

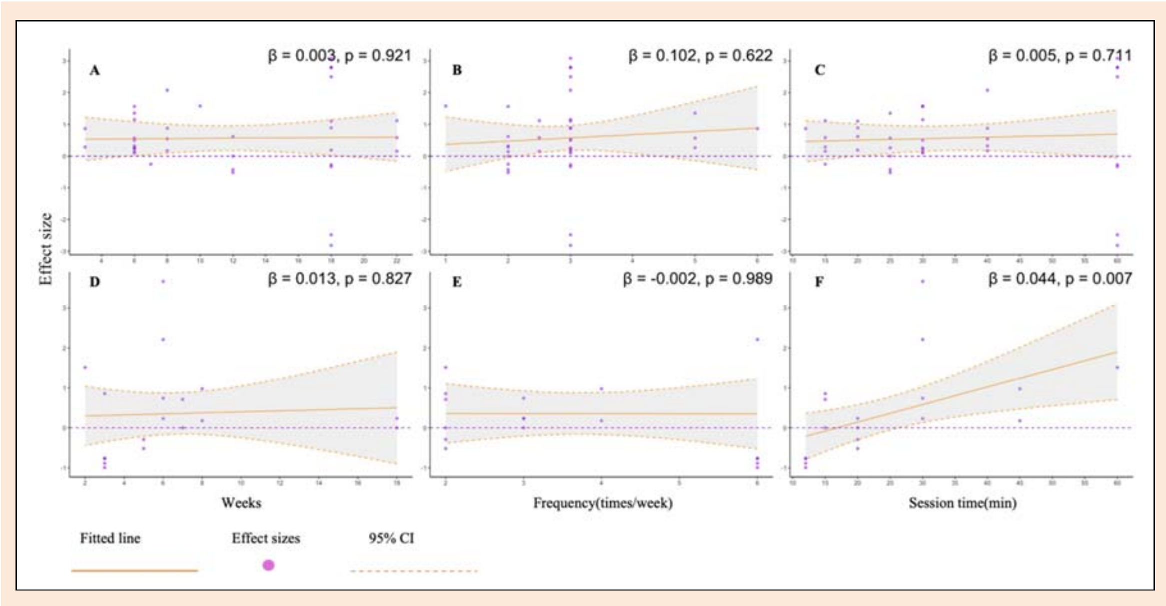

Figure S5. **Linear regression for secondary results. A, B, C results for linear regression of training duration, training frequency and session time for Pre-planned agility test; D, E, F results for linear regression of training duration, training frequency and session time for Reaction test.**