

Supplementary Table 1. Database-specific search strategies, search limits, and retrieval results

Database	Exact Executed Search Strategy	Search Fields	Applied Limits	Final Search Date	Records Retrieved
Scopus	TITLE-ABS-KEY(("artificial intelligence" OR "machine learning" OR "computer vision" OR "intelligent feedback" OR "motion tracking" OR "AI- supported") AND ("motor learning" OR "motor skill" OR "skill acquisition" OR "movement learning") AND ("physical education" OR "physical literacy" OR sport* OR "physical activity") AND (learning OR teaching OR education OR pedagogy OR instruction))	TITLE-ABS- KEY	English; Article; 2015– 2026	22 June 2026	80
PubMed	(("artificial intelligence"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "computer vision"[Title/Abstract] OR "intelligent feedback"[Title/Abstract] OR "motion tracking"[Title/Abstract] OR "AI- supported"[Title/Abstract]) AND ("motor learning"[Title/Abstract] OR "motor skill"[Title/Abstract] OR "skill acquisition"[Title/Abstract] OR "movement learning"[Title/Abstract]) AND ("physical education"[Title/Abstract] OR "physical literacy"[Title/Abstract] OR sport*[Title/Abstract] OR "physical activity"[Title/Abstract]) AND (learning[Title/Abstract] OR teaching[Title/Abstract] OR education[Title/Abstract] OR pedagogy[Title/Abstract] OR instruction[Title/Abstract]))	Title/Abstract	English; 2015– 2026	22 June 2026	14

ProQuest	noft("artificial intelligence" OR "machine learning" OR "computer vision" OR "intelligent feedback" OR "motion tracking" OR "AI-supported") AND noft("motor learning" OR "motor skill" OR "skill acquisition" OR "movement learning") AND noft("physical education" OR "physical literacy" OR sport* OR "physical activity") AND noft(learning OR teaching OR education OR pedagogy OR instruction)	Anywhere (sesuai yang benar-benar digunakan)	English; Scholarly Journals; 2015– 2026	22 June 2026	14
----------	---	---	---	-----------------	----

Supplementary Table 2. Classification of included studies according to study design, outcome domain, and synthesis eligibility

No	Study	Design	Primary Outcome Domain	Qualitative Synthesis	Meta-analysis	Reason
1	Ma (2025)	RCT	Skill acquisition	Yes	Yes	Complete n, mean, SD
2	Rakha (2026)	RCT	Skill acquisition	Yes	Yes	Complete statistics
3	Zhang (2025)	Mixed experimental	Skill acquisition	Yes	No	No extractable group statistics
4	Zhao (2025)	Quasi-experimental	Motor proficiency	Yes	Yes	Complete statistics
5	Wu (2025)	RCT crossover	Motor proficiency	Yes	No	Outcome unsuitable for pooled comparison
6	Xiao & Guo (2026)	Quasi-experimental	Motor proficiency	Yes	Yes	Complete statistics
7	Rauter (2019)	Experimental	Movement accuracy	Yes	No	Repeated-measures design
8	Harris (2024)	Mixed experiment	Transfer	Yes	No	Transfer outcome only
9	Yan (2024)	Quasi-experimental	Skill acquisition	Yes	No	Statistics not extractable
10	Qiu (2024)	Experimental	Skill acquisition	Yes	No	Statistics not extractable
11	Tang (2024)	Quasi-experimental	Skill acquisition	Yes	Yes	Complete statistics
12	Takita (2025)	Experimental	Retention	Yes	No	Retention outcome synthesized narratively
13	Hsia (2025)	Quasi-experimental	Skill acquisition	Yes	Yes	Complete statistics
14	Wójcik (2020)	Experimental	Movement accuracy	Yes	No	No comparable control outcome
15	Wójcik & Piekarczyk (2020)	Experimental	Skill transfer	Yes	No	Transfer outcome synthesized narratively

Supplementary Table 3. Mapping of included studies to motor learning outcome domains and outcome selection for quantitative synthesis

No	Study	Outcome Domain	Exact Outcome Selected	Assessment Time Point	Included in Meta-analysis	Rationale for Outcome Selection
1	Ma et al. (2025)	Skill acquisition	Total Baduanjin skill score	Post-intervention	Yes	Primary learner-level motor learning outcome with complete group statistics (n, mean, SD).
2	Rakha (2026)	Skill acquisition	Composite handball skill score (passing, catching, dribbling, shooting, faking, defense, goalkeeping)	Post-intervention	Yes	Composite motor skill score represented the primary intervention outcome and contained complete quantitative data.
3	Zhang (2025)	Skill acquisition	Basketball shooting skill acquisition	Post-intervention	No	Only percentages and descriptive improvements were reported; insufficient extractable statistics for effect-size calculation.
4	Zhao et al. (2025)	Motor proficiency	One-minute jump rope performance	Post-intervention	Yes	Selected as the primary motor learning outcome because it directly reflects acquired motor proficiency and provided complete

						statistical data.
5	Wu (2025)	Motor proficiency	Movement skill proficiency	Post-intervention	No	Improvement percentages were reported without extractable group means and standard deviations.
6	Xiao & Guo (2026)	Motor proficiency	Physical fitness motor proficiency rubric (30-item analytical rubric)	Post-intervention	Yes	Primary learner-level motor proficiency outcome with sufficient quantitative statistics.
7	Rauter et al. (2019)	Movement accuracy	Velocity error / rowing movement accuracy	Post-intervention	No	Outcomes were reported as repeated-measures learning curves rather than independent post-test comparison suitable for meta-analysis.
8	Harris et al. (2024)	Skill transfer	Real-world golf putting transfer performance	Transfer test	No	Primary outcome assessed transfer rather than immediate performance and did not provide compatible statistics for pooling.
9	Yan et al. (2024)	Skill acquisition	Sport-specific motor skill scores (volleyball, basketball, soccer, badminton,	Post-intervention	No	Statistical significance was reported without complete extractable

			table tennis, aerobics)			summary statistics.
10	Qiu (2024)	Skill acquisition	Composite badminton motor skill score	Post-intervention	No	Post-test means or adjusted comparisons were reported, but the standard deviations required to calculate a compatible standardized mean difference were unavailable.
11	Tang (2024)	Skill acquisition	Motor training performance score	Post-intervention	Yes	The VR and traditional-instruction groups each included 40 participants, and complete post-test means and standard deviations for the motor skill training score were reported.
12	Takita et al. (2025)	Skill retention	Retention performance score	Retention test	No	Focused primarily on delayed retention rather than immediate motor performance; synthesized narratively.
13	Hsia et al. (2025)	Skill acquisition	Yoga performance score	Post-intervention	Yes	Primary learner-level motor learning outcome with complete

						quantitative statistics.
14	Wójcik (2020)	Movement accuracy	Hand-foot synchronization accuracy	Post-intervention	No	Comparative design did not provide extractable intervention-versus-control statistics compatible with meta-analysis.
15	Wójcik & Piekarczyk (2020)	Skill transfer	Transfer of motor learning (RMSE-based movement quality)	Transfer assessment	No	Primary outcome evaluated transfer performance and lacked compatible statistics for pooled quantitative synthesis.

Supplementary Table 4. Leave-one-out sensitivity analysis of the meta-analysis for short-term motor skill performance

Analysis	Studies (n)	Participants (AI / Comparator)	SMD (95% CI)	τ^2	χ^2 (df=4)	I ²	Z	p-value
Full model	6	344 / 342	3.15 (1.86–4.44)	2.38	210.35	98%	4.78	<0.00001
Excluding Hsia et al. (2025)	5	264 / 266	3.95 (2.18–5.71)	3.76	185.03	98%	4.39	<0.0001
Excluding Ma et al. (2025)	5	316 / 310	3.80 (2.23–5.37)	2.94	206.80	98%	4.74	<0.00001
Excluding Rakha (2026)	5	316 / 314	2.70 (1.43–3.97)	1.91	160.70	98%	4.16	<0.0001
Excluding Tang (2024)	5	304 / 302	1.68 (0.78–2.59)	0.97	86.62	95%	3.66	0.0002
Excluding Xiao & Guo (2026)	5	224 / 222	3.90 (1.96–5.84)	4.62	209.12	98%	3.93	<0.0001
Excluding Zhao et al. (2025)	5	296 / 296	3.42 (1.91–4.93)	2.71	190.32	98%	4.44	<0.00001

Leave-one-out sensitivity analysis was performed by sequentially omitting one study at a time from the random-effects model. Although the pooled standardized mean difference varied across iterations, all analyses continued to favor AI-supported interventions. The greatest reduction in the pooled effect size occurred after exclusion of Tang (2024), indicating that this study exerted the largest influence on the magnitude of the overall effect, whereas exclusion of the remaining studies resulted in only modest changes to the pooled estimate. Between-study heterogeneity remained high in most iterations ($I^2 = 95\text{--}98\%$).

Supplementary Table 5a. Domain-level risk-of-bias assessment of randomized controlled trials using RoB 2

Study	Randomization Process (D1)	Deviations from Intended Interventions (D2)	Missing Outcome Data (D3)	Measurement of the Outcome (D4)	Selection of the Reported Result (D5)	Overall Judgment
Ma et al. (2025)	Low	Some concerns	Low	Low	Some concerns	Some concerns
Rakha (2026)	Low	Low	Low	Some concerns	Some concerns	Some concerns
Wu (2025)	Some concerns	Some concerns	Some concerns	Some concerns	High	High
Takita et al. (2025)	Some concerns	Some concerns	Low	Some concerns	Some concerns	Some concerns
Harris et al. (2024)	Low	Some concerns	Low	Low	Some concerns	Some concerns
Wójcik & Piekarczyk (2020)	Some concerns	Some concerns	Low	Low	Some concerns	Some concerns

Domain Legend (RoB 2):

D1: Bias arising from the randomization process

D2: Bias due to deviations from intended interventions

D3: Bias due to missing outcome data

D4: Bias in measurement of the outcome

D5: Bias in selection of the reported result

Supplementary Table 5b. Domain-level risk-of-bias assessment of non-randomized studies using ROBINS-I

Study	Confounding (D1)	Selection of Participants (D2)	Classification of Interventions (D3)	Deviations from Intended Interventions (D4)	Missing Data (D5)	Measurement of Outcomes (D6)	Selection of the Reported Result (D7)	Overall Judgment
Zhang (2025)	Serious	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Serious
Zhao et al. (2025)	Moderate	Low	Low	Moderate	Low	Low	Moderate	Moderate
Xiao & Guo (2026)	Moderate	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Moderate
Rauter et al. (2019)	Moderate	Low	Low	Moderate	Low	Low	Moderate	Moderate
Yan et al. (2024)	Serious	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Serious
Qiu (2024)	Moderate	Moderate	Low	Moderate	Low	Moderate	Moderate	Moderate
Tang (2024)	Serious	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Serious
Hsia et al. (2025)	Moderate	Moderate	Low	Moderate	Low	Low	Moderate	Moderate
Wójcik (2020)	Moderate	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Moderate

Domain Legend (ROBINS-I):

D1: Bias due to confounding

D2: Bias in selection of participants into the study

D3: Bias in classification of interventions

D4: Bias due to deviations from intended interventions

D5: Bias due to missing data

D6: Bias in measurement of outcomes

D7: Bias in selection of the reported result