ISSN: 3107-6513

Transfer and Retention of Motor Skills from Virtual Reality Training to Physical Performance: A Critical Evidence Synthesis and Appraisal Framework for Physical Education

[1*]Dhanaji Shivaji Jadhav, [2*]Dr. S.S. Autade
[1*] Phd. Research Scholar, Department of Physical Education, *Saveetha Amravati University, Vijaywada, Andhra Pradesh*
[2*] PhD Guide, Department of Physical Education, *Saveetha Amravati University, Vijaywada, Andhra Pradesh*

Abstract—
Virtual reality (VR) is increasingly proposed for physical education, but its value depends on whether skills practised in a virtual environment improve real movement and remain after training ends. This paper critically synthesises eight primary studies and one systematic review that tested VR training against physical performance, covering baseball batting, table tennis, golf putting, dart throwing, obstacle avoidance and school balance and coordination. Evidence is organised by the type of comparison, the form of transfer measured and the presence of a retention test. Positive transfer was reported against no-training controls and, in some studies, against active physical practice; null group differences and negative or partial transfer were also reported. A power analysis of the reported group sizes shows that the studies, with 14 to 35 participants per group, could reliably detect only standardised differences of 0.69 to 1.10, so small or moderate effects—and equivalence with conventional practice—cannot be inferred from them. An appraisal framework is proposed that separates virtual performance, near and far physical transfer and retention, and that examines physical, perceptual, instructional and transition differences between training and assessment. The framework is intended to guide the design and interpretation of school-based VR studies.

Index Terms—*Virtual reality, motor learning, transfer of training, retention, physical education, evidence synthesis, statistical power, equivalence*

1. Introduction

Motor competence supports participation in physical activity throughout childhood and adolescence [1], and physical education (PE) is the setting in which most students practise fundamental movement skills. VR offers repeatable, adjustable practice with immediate feedback, and it is increasingly marketed to schools. The educational question, however, is not whether students improve in a virtual game. It is whether VR-supported practice improves their physical performance and whether any improvement lasts [2].

A systematic review identified only four articles, from 448 screened, that met its criteria for assessing real-world transfer of VR sports training, and it noted wide variation in technology, duration and outcome measures [3]. Since then further studies have appeared, including school-based work. This paper, derived from a doctoral thesis on VR training and motor skills in PE, contributes:

1) A critical synthesis of primary studies testing VR training against physical performance, organised by comparison type, transfer type and retention.
2) A power analysis showing the smallest effects that each study could reliably detect.
3) An appraisal framework for judging transfer and retention claims, including sources of difference between virtual training and physical assessment.

2. Concepts and Method of Synthesis

2.1 Transfer and Retention

Transfer is the effect of earlier practice on performance in another task or setting. Near transfer concerns a physical version of the practised task; far transfer concerns a larger change, such as an untrained task or a game situation. Transfer may be positive, absent or negative. Schema theory predicts that varied practice builds rules that generalise to new variations of a task [4], which makes the relationship between training and test conditions central. Retention concerns performance after a delay and must report the interval and the activity permitted during it. Fig. 1 shows how these forms of evidence build on each other.

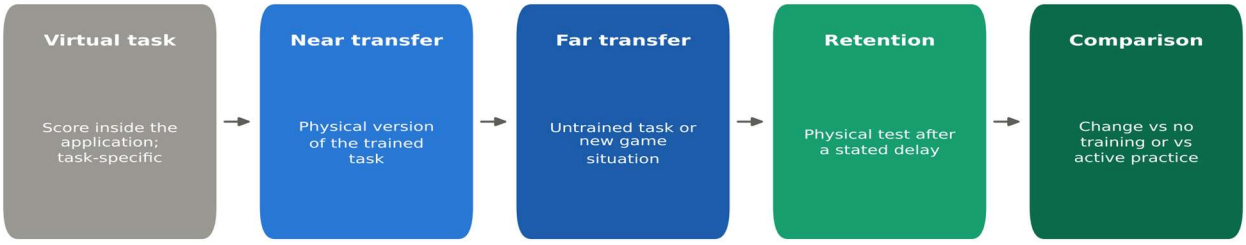


Fig. 1. Forms of evidence about VR motor training. Only physical tests, compared with an appropriate group, support claims about real motor skill development.

2.2 Why Transfer May Succeed or Fail

Transfer is expected when the virtual and physical tasks share the information and actions that control performance. VR can present realistic timing, target positions and decisions, and it allows difficulty to be adjusted systematically, which can support learning of the shared components [2]. Adaptive difficulty was central to the strongest positive result in the reviewed evidence [5]. Transfer may fail when the virtual task differs in ways that matter for control: controllers replace real objects, haptic feedback is absent, depth cues and latency differ, and feedback in the application may not be available during physical performance. Under these conditions learners may adapt to the virtual task itself, and the adaptation can fail to carry over or even interfere briefly with physical performance [6], [7].

Retention adds a further question. Practice that makes performance easy during training, such as continuous guidance or constant conditions, can produce good immediate performance but weaker retention, whereas varied practice supports generalisation [4]. A VR study that measures only immediate post-training performance therefore cannot show whether the programme has produced lasting learning.

2.3 Selection and Appraisal

Studies were selected for this critical (non-systematic) synthesis when they trained a motor task in VR and measured performance in a corresponding physical task. For each, the design, group sizes, practice dose, comparison group, outcome and reported finding were extracted from the published article. The smallest standardised difference detectable with 80% power at two-sided $\alpha = .05$ was calculated from the reported group sizes using the noncentral t distribution. This is a property of the design, not of the observed result, and indicates which effects a study could reliably reveal.

3. Evidence Synthesis

3.1 Studies Reporting Positive Transfer

Gray assigned 80 baseball players equally to adaptive VR training, extra VR batting, extra real batting or no extra training, with two 45-minute sessions per week for six weeks [5]. The adaptive VR group improved more on most measures, including an on-field batting test, and had better league statistics in the following season. Michalski *et al.* randomised 57 participants to VR table tennis training (n = 29) or a no-training control (n = 28); the VR group completed about 3.5 hours over seven sessions, and an expert coach rated real table tennis performance before and after training [8]. VR training improved real performance relative to no training.

3.2 Studies Reporting No Group Difference

Markwell *et al.* randomised 68 participants to commercial VR golf putting (n = 33) or real putting practice (n = 35) over two consecutive days with 60 practice putts per day [9]. Both groups improved putting accuracy, with a main effect of test but not of group. In the second experiment of Harris *et al.*, VR and real-world practice produced comparable improvements in putting accuracy among 40 novice golfers [6]. Neither study was designed as an equivalence trial, so the absence of a significant difference does not show that the methods are equivalent [10].

3.3 Studies Reporting Negative or Partial Transfer

In the first experiment of Harris *et al.*, 18 expert golfers showed temporarily disrupted gaze control and putting accuracy when moving from VR to real putting [6]. Drew *et al.* trained 41 adults in a single 25-minute session of 100 dart throws in VR (n = 22) or the real world (n = 19); the VR group threw less accurately afterwards than the real-world group and than its own baseline, and reported more visual symptoms [11]. Weber *et al.* trained 32 young adults to avoid sudden virtual obstacles or walk an obstacle-free virtual corridor (14 per group analysed). The virtual adaptation was fully retained after about one week, but transfer to physical obstacles was only partial, and its retention was partial too [7].

3.4 School-Based Studies

Yanar Tunçel and Avşar studied 39 seventh-grade students (20 VR, 19 control) over eight weeks and found greater improvement in the VR group on Y-Balance scores and most KTK coordination subtests, but not single-leg hopping [12]. Fernández-Vázquez *et al.* compared three PE groups—practice teaching style alone, with gamification, and with gamification plus VR—in 75 students over six weeks of two one-hour sessions; the VR group improved on plate tapping, and the authors concluded that combining gamification with VR enhanced some motor skills [13]. Because the groups differed in several features, the VR contribution cannot be separated from that of gamification. Table I summarises all studies.

ISSN: **3107-6513**

TABLE I STUDIES TESTING VR TRAINING AGAINST PHYSICAL PERFORMANCE

Study	Task and participants	Dose	Comparison	Reported finding
Gray [5]	Baseball batting; 80 players, 4 groups	12 × 45 min over 6 weeks	Extra VR, extra real practice, none	Adaptive VR improved more; better league statistics
Michalski *et al.* [8]	Table tennis; 29 vs 28	≈3.5 h, 7 sessions	No training	Real performance improved vs control
Markwell *et al.* [9]	Golf putting; 33 vs 35	2 days, 60 putts/day	Real practice	Both improved; no group effect
Harris *et al.* [6]	Putting; 18 experts; 40 novices	Short practice	Real practice	Temporary disruption (experts); comparable gains (novices)
Drew *et al.* [11]	Darts; 22 vs 19	1 session, 100 throws	Real practice	Lower real accuracy after VR
Weber *et al.* [7]	Obstacle avoidance; 14 vs 14	1 session; retest at ≈1 week	Obstacle-free VR walking	Partial transfer; partial retention of transfer
Yanar Tunçel and Avşar [12]	Balance, coordination; 20 vs 19 students	8 weeks	Control group	Greater Y-Balance and most KTK gains
Fernández-Vázquez *et al.* [13]	PE motor skills; 75 students, 3 groups	12 × 60 min over 6 weeks	Practice style ± gamification	VR group improved plate tapping

3.5 Evidence Map

Fig. 2 arranges the studies by comparison condition and direction of finding. Positive transfer is concentrated in comparisons with no training or unspecified control activity, whereas comparisons with active physical practice produced positive, null and negative results. This pattern is consistent with the view that VR adds practice opportunities, but that its advantage over equivalent physical practice is task-dependent and not yet established.



Fig. 2. Evidence map of the reviewed studies by comparison condition and direction of the reported physical-performance finding. *Positive on some motor tests; VR combined with gamification.*

3.6 Implications for Classroom Applications

VR applications for PE differ in display, controllers, tracking, movement space and teacher involvement, so "VR training" must be defined through the equipment and tasks actually used. A game that appears active may allow success with little body movement, and software response scores include the system's own detection delay. Table II links typical classroom applications to the physical evidence needed before improvement can be claimed.

ISSN: 3107-6513

TABLE II CLASSROOM APPLICATIONS AND THE PHYSICAL EVIDENCE REQUIRED

Application	Intended movement demand	Evidence required
Balance tasks	Controlled changes in posture and weight distribution	Physical balance test, e.g. Y-Balance
Target-response tasks	Timely selection and initiation of movement	Validated response timing and accuracy
Coordination tasks	Organisation of several body movements	Independent coordination battery, e.g. KTK
Aiming and object control	Accuracy of projected objects	Physical throwing or striking accuracy
Sport-specific simulations	Task-specific perception and action	Physical sport task; transfer to play

4. Statistical Power of the Evidence

Fig. 3 shows the smallest standardised difference *d* that each design could detect with 80% power. With 14 per group, Weber *et al.* could reliably detect only $d \geq 1.10$; designs with about 20 per group could detect $d \geq 0.90$–0.92; and the largest two-group comparisons, with 29–35 per group, could detect $d \geq 0.69$–0.76. A moderate effect of $d = 0.5$ would require about 64 participants per group, as planned in the thesis trial. Three consequences follow. Positive findings in small studies are likely to overestimate the true effect, because only large observed effects reach significance; null findings are uninformative about moderate effects; and none of the studies was large enough to establish equivalence within a narrow margin [10].

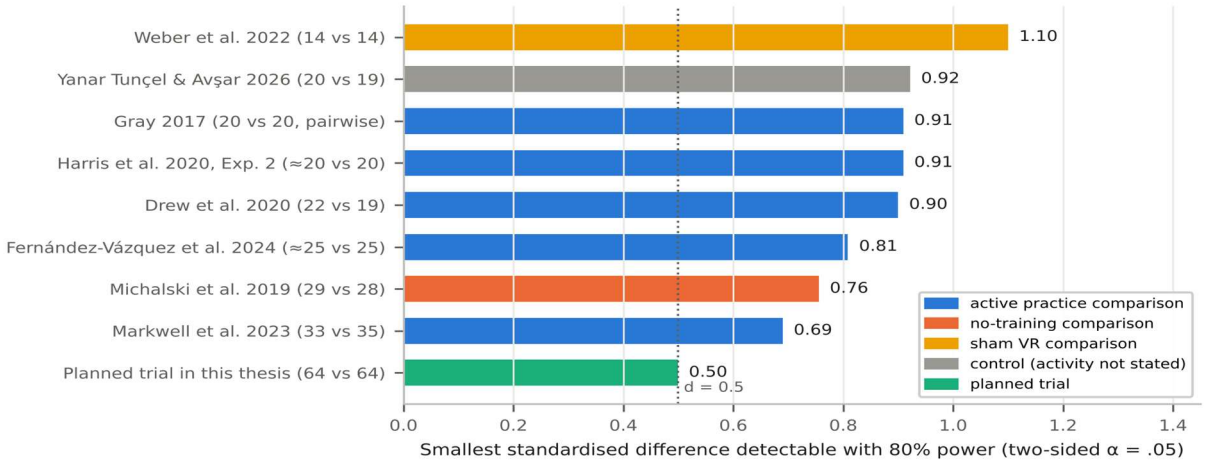


Fig. 3. Smallest standardised difference each study could detect with 80% power, calculated from the reported group sizes (approximate where groups were not reported separately). Colour shows the comparison condition.

5. Appraisal Framework

5.1 Sources of Difference Between Training and Assessment

The contrasting findings suggest that transfer depends on the demands that training and assessment share, not on visual resemblance. Fig. 4 separates five sources of difference. Physical differences concern movement range, object weight, contact and resistance; perceptual differences concern depth cues, visual field, latency and haptic information; instructional differences concern feedback, guidance and practice time. Transition effects concern what happens when learners switch environments, as in the immediate disruption reported among expert golfers [6] and the accuracy decline after VR darts [11]. The assessment itself must be described so that the shared demands are explicit.

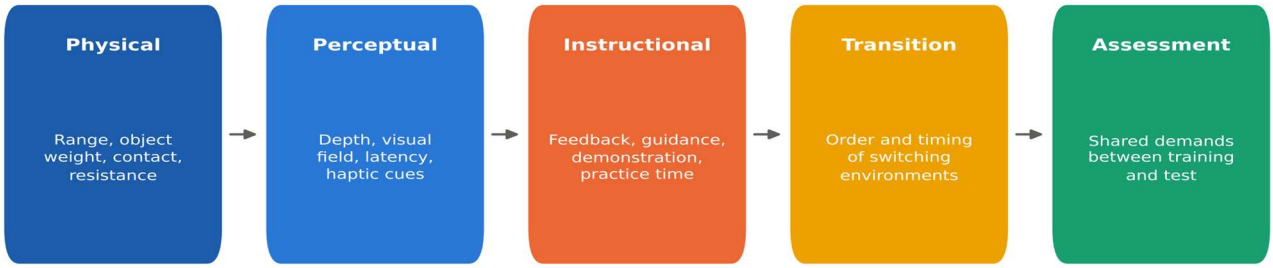


Fig. 4. Sources of difference between VR training and physical assessment that should be described when interpreting transfer.

5.2 Appraisal Questions

Table III lists questions for appraising a transfer or retention claim. They distinguish what a study can show from what it cannot, and they apply equally to positive and negative findings.

TABLE III QUESTIONS FOR APPRAISING EVIDENCE ON VR TRANSFER AND RETENTION

Question	Why it matters	Example from the evidence
Is the outcome a physical test?	Virtual scores show task-specific learning	All included studies used physical tests
What is the comparison?	No-training and active comparisons answer different questions	Table tennis vs no training [8]; putting vs real practice [9]
Was the assessor unaware of allocation?	Reduces biased scoring of performance	Coach rating in table tennis study [8]
Is transfer near or far?	Near transfer does not show general competence	League batting outcomes as far transfer [5]
Is retention measured, and when?	Immediate gains may not last	One-week retest in obstacle avoidance [7]
Could the design detect the effect claimed?	Small samples miss moderate effects	Detectable d of 0.69–1.10 (Fig. 3)
Is equivalence claimed from a null result?	Requires prespecified margin and power	No-group-difference studies [9], [6]

6. Discussion

6.1 What the Evidence Supports

The evidence supports a limited claim: VR training of a specific task can improve the corresponding physical skill, at least relative to no extra training, and in one adaptive system relative to additional physical practice [5]. It does not support the general claim that VR games improve motor skills, nor that VR is equivalent to conventional practice. The negative and partial findings show that transfer can fail, especially immediately after switching environments and when perceptual information differs between VR and the real task [11], [7].

6.2 Implications for School Research

School studies should use physical outcomes, active comparisons with matched time, assessors unaware of allocation, a retention test with a stated interval and samples large enough for the effects of interest. Interventions should be described fully so that their dose, feedback and progression can be reproduced [14], and trials should be reported according to CONSORT [15]. Where VR is combined with linked physical practice, as proposed in the thesis programme, a favourable result will concern the complete teaching approach rather than isolating the contribution of the virtual component.

Table IV translates the appraisal framework into design recommendations for school-based studies, each linked to a weakness observed in the reviewed evidence.

TABLE IV DESIGN RECOMMENDATIONS FOR SCHOOL-BASED VR STUDIES

Recommendation	Weakness addressed	Practical step
Time-matched active comparison	No-training controls inflate apparent benefit	Conventional practice with the same objectives and minutes
Physical primary outcome	Virtual scores mistaken for skill	Pre-specified physical test, e.g. Y-Balance
Untrained transfer task	Near transfer over-interpreted	Task not practised in either group
Delayed retention test	Immediate gains assumed to last	Follow-up after a stated interval without training
Adequate sample size	Only large effects detectable	Power for a meaningful effect in original units
Separate VR from game features	Gamification confounded with VR	Match rewards, praise and feedback across groups

6.3 Limitations

This is a critical synthesis rather than a systematic review: the study selection is not exhaustive, risk of bias was not formally scored, and effects were not pooled because tasks, doses and outcomes differ greatly. Group sizes for two studies were approximated from reported totals.

6.4 Research Agenda

Three priorities follow from the synthesis. First, adequately powered trials in schools should compare VR-supported teaching with time-matched conventional instruction on physical outcomes. Second, studies should include delayed retention tests and at least one untrained transfer task. Third, studies should report the physical, perceptual and instructional features of the VR task so that differences in findings can be related to differences in design, and should state whether a null result was intended as evidence of equivalence, with a margin specified in advance [10].

7. Conclusion and Future Work

VR training can transfer to physical performance in specific tasks, but transfer is neither automatic nor uniform, and

evidence on retention is scarce. The available studies were small, able to detect only large effects ($d \geq 0.69$ in the largest two-group comparison), and mostly short. The proposed framework separates virtual performance, near and far transfer and retention, and it directs attention to the physical, perceptual, instructional and transition differences between training and testing. Future work should test VR-supported PE against time-matched conventional instruction in adequately powered trials with delayed retention tests.

References

[1] D. F. Stodden *et al.*, "A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship," *Quest*, vol. 60, no. 2, pp. 290–306, 2008, doi: 10.1080/00336297.2008.10483582.

[2] D. E. Levac, M. E. Huber, and D. Sternad, "Learning and transfer of complex motor skills in virtual reality: A perspective review," *J. NeuroEng. Rehabil.*, vol. 16, Art. no. 121, 2019, doi: 10.1186/s12984-019-0587-8.

[3] S. C. Michalski, A. Szpak, and T. Loetscher, "Using virtual environments to improve real-world motor skills in sports: A systematic review," *Front. Psychol.*, vol. 10, Art. no. 2159, 2019, doi: 10.3389/fpsyg.2019.02159.

[4] R. A. Schmidt, "A schema theory of discrete motor skill learning," *Psychol. Rev.*, vol. 82, no. 4, pp. 225–260, 1975, doi: 10.1037/h0076770.

[5] R. Gray, "Transfer of training from virtual to real baseball batting," *Front. Psychol.*, vol. 8, Art. no. 2183, 2017, doi: 10.3389/fpsyg.2017.02183.

[6] D. J. Harris *et al.*, "The effect of a virtual reality environment on gaze behaviour and motor skill learning," *Psychol. Sport Exerc.*, vol. 50, Art. no. 101721, 2020, doi: 10.1016/j.psychsport.2020.101721.

[7] A. Weber *et al.*, "Limited transfer and retention of locomotor adaptations from virtual reality obstacle avoidance to the physical world," *Sci. Rep.*, vol. 12, Art. no. 19655, 2022, doi: 10.1038/s41598-022-24085-w.

[8] S. C. Michalski, A. Szpak, D. Saredakis, T. J. Ross, M. Billinghurst, and T. Loetscher, "Getting your game on: Using virtual reality to improve real table tennis skills," *PLOS ONE*, vol. 14, no. 9, Art. no. e0222351, 2019, doi: 10.1371/journal.pone.0222351.

[9] L. T. Markwell, K. Cochran, and J. M. Porter, "Off the shelf: Investigating transfer of learning using commercially available virtual reality equipment," *PLOS ONE*, vol. 18, no. 10, Art. no. e0279856, 2023, doi: 10.1371/journal.pone.0279856.

[10] D. Lakens, A. M. Scheel, and P. M. Isager, "Equivalence testing for psychological research: A tutorial," *Adv. Methods Pract. Psychol. Sci.*, vol. 1, no. 2, pp. 259–269, 2018, doi: 10.1177/2515245918770963.

[11] S. A. Drew *et al.*, "The trade-off of virtual reality training for dart throwing: A facilitation of perceptual-motor learning with a detriment to performance," *Front. Sports Act. Living*, vol. 2, Art. no. 59, 2020, doi: 10.3389/fspor.2020.00059.

[12] N. Yanar Tunçel and Z. Avşar, "Virtual reality experience in physical education and sports lessons: An investigation into its effect on the balance and coordination skills of middle school students," *PLOS ONE*, vol. 21, no. 8, Art. no. e0354258, 2026, doi: 10.1371/journal.pone.0354258.

[13] D. Fernández-Vázquez *et al.*, "Influence of virtual reality and gamification combined with practice teaching style in physical education on motor skills and students' perceived effort: A mixed-method intervention study," *Sustainability*, vol. 16, no. 4, Art. no. 1584, 2024, doi: 10.3390/su16041584.

[14] T. C. Hoffmann *et al.*, "Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide," *BMJ*, vol. 348, Art. no. g1687, 2014, doi: 10.1136/bmj.g1687.

[15] S. Hopewell *et al.*, "CONSORT 2025 statement: Updated guideline for reporting randomised trials," *BMJ*, vol. 389, Art. no. e081123, 2025, doi: 10.1136/bmj-2024-081123.