

VIRTUAL REALITY FOR TEACHING PRACTICAL SKILLS IN HIGHER EDUCATION: AN INTEGRATIVE REVIEW

Ben Bander Abudawood

Abertay University

b.abudawood2200@abertay.ac.ukDOI: <https://doi.org/10.5281/zenodo.23075866>**Keywords**

virtual reality; higher education; practical skills; psychomotor learning; simulation; immersive learning; integrative review

Article History

Received: 01 January 2026

Accepted: 16 March 2026

Published: 31 March 2026

Copyright @Author

Corresponding Author: *

Ben Bander Abudawood

Abstract

Virtual reality (VR) is increasingly used in higher education to provide repeatable, interactive and comparatively safe environments in which students can rehearse practical tasks. Yet the evidence is distributed across medicine, nursing, engineering, science, laboratory education and other professionally oriented disciplines, making it difficult to determine when VR contributes to genuine skill development rather than short-term engagement. This integrative review synthesizes secondary and empirical literature on the use of VR for teaching practical, procedural and psychomotor skills in higher education. Guided by the integrative-review framework of Whitemore and Knafl, the review draws on peer-reviewed literature identified through targeted searches of scholarly databases and publisher indexes, with emphasis on higher education learners, interactive VR interventions and outcomes related to observable skill performance. Across the literature, VR is generally associated with improved procedural knowledge, task accuracy, confidence and opportunities for deliberate practice. Recent meta-analyses report positive practical-skill effects in science, engineering and nursing, although newer evidence restricted to immersive head-mounted-display simulation found no statistically significant pooled advantage for nursing psychomotor performance. However, effects vary substantially by discipline, degree of immersion, instructional design, duration, pre-training, feedback and the comparison condition. The strongest evidence supports VR as a complement to, rather than an automatic replacement for, supervised physical practice. Important limitations include hardware and content costs, cybersickness, usability problems, unequal access, instructor preparation, weak reporting of intervention design, small samples and a substantially thinner evidence base for long-term retention and transfer than for immediate skill acquisition. The review proposes a pedagogically grounded implementation model in which VR is aligned with explicit skill objectives, preceded by orientation and pre-training, combined with feedback and debriefing, and followed by authentic performance assessment. Future research should prioritize longitudinal transfer, cost-effectiveness, accessibility and standardized reporting of VR interventions.

1. INTRODUCTION

Practical competence is a central outcome of higher education in disciplines in which

students must do more than recall or explain knowledge. Engineers manipulate equipment

and troubleshoot systems; health-professions students perform procedures and respond to rapidly changing situations; science students use instruments and conduct laboratory protocols; and students in technical and vocationally oriented programmes must coordinate perception, judgment and action. These forms of learning are difficult to develop through lectures alone because competence depends on repeated performance, feedback, error correction and exposure to realistic constraints. Conventional practical teaching nevertheless faces persistent limitations. Laboratories and specialist facilities are expensive, equipment may be scarce, class sizes can restrict individual practice time, and some procedures involve safety, ethical or logistical constraints. In clinical education, opportunities to perform a procedure may depend on patient availability and supervision. In engineering and science, high-value equipment, hazardous processes and consumable materials can similarly constrain repetition. These pressures have encouraged universities to explore simulation technologies that allow students to practise before entering higher-risk or resource-intensive environments. Virtual reality is one of the most prominent of these technologies. In this review, VR refers to computer-generated, interactive three-dimensional environments that allow learners to perceive and act within a simulated setting, commonly through head-mounted displays, motion controllers or related interfaces. VR differs from passive video or conventional screen-based multimedia because it can combine spatial representation, interaction, feedback and a sense of presence. These characteristics make it attractive for practical learning: a learner can repeat a procedure, receive immediate information about errors, experience scenarios that are difficult to stage physically and practise without damaging equipment or placing patients, peers or themselves at unnecessary risk. Enthusiasm for VR should not, however, be confused with evidence of educational effectiveness. A visually immersive environment can increase attention and perceived realism while simultaneously imposing extraneous cognitive load. Novelty may improve satisfaction without improving skill. A learner may also become proficient at the interface without

transferring that performance to real equipment or professional practice. Consequently, the relevant question is not simply whether students like VR, but whether VR helps them acquire, retain and transfer practical skills under conditions that justify its cost and complexity.

Recent quantitative syntheses provide an increasingly useful evidence base. Yang et al. [1] analysed 37 empirical studies in science and engineering education and reported a moderate positive overall effect of VR on practical skills (Hedges' $g = 0.477$). In nursing, Kim and Park [2] reported a medium-to-large improvement in clinical skill performance ($g = 0.61$) across randomized controlled trials, while Lin et al. [3] found positive effects on skill, knowledge, confidence and satisfaction. These findings are encouraging, but they also reveal substantial variation in interventions, disciplines, outcome measures and instructional conditions. A cross-disciplinary synthesis is therefore needed to clarify what can reasonably be concluded and how universities should design VR-supported practical teaching.

The purpose of this integrative review is to synthesize current evidence on VR for practical-skill development in higher education, identify pedagogical and technological conditions associated with effective use, examine implementation barriers, and define priorities for research and practice. The review addresses five questions: (1) How is VR being used to teach practical skills in higher education? (2) What evidence exists regarding its effects on skill acquisition? (3) Which instructional and technological features appear to influence outcomes? (4) What barriers affect adoption and scalability? and (5) What gaps remain concerning retention, transfer, accessibility and cost-effectiveness?

2. Conceptual Background

2.1 Practical skills as an educational outcome

Practical skills include procedural, technical and psychomotor capabilities that require coordinated application of knowledge, perception and action. Although the balance differs by discipline, practical competence typically involves knowing the sequence of a task, recognizing relevant cues, executing actions accurately, adapting to variation and evaluating

the result. Skill learning is therefore not reducible to declarative knowledge. A student may correctly describe a procedure while remaining unable to perform it safely or fluently. Experiential and simulation-based approaches are relevant because they create opportunities for active performance and reflection. Kolb's experiential learning model emphasizes cycles of concrete experience, reflective observation, conceptualization and active experimentation [4]. In practical education, this implies that technology is most useful when it enables purposeful action and is integrated with reflection and subsequent application rather than treated as a self-contained experience. Deliberate practice similarly highlights repeated performance of well-defined tasks with feedback and opportunities for correction [5]. VR can support these mechanisms when scenarios are designed around explicit performance objectives.

2.2 Immersion, presence and cognitive processing

VR systems vary from desktop three-dimensional environments to highly immersive head-mounted displays with tracked controllers. Immersion is primarily a property of the technological system, whereas presence describes the user's subjective sense of being located in the mediated environment. Greater immersion can improve spatial interaction and realism, but it does not guarantee learning. Makransky and Petersen's Cognitive Affective Model of Immersive Learning (CAMIL) proposes that immersive features influence learning through factors such as presence, agency, motivation, self-efficacy and cognitive load [6].

This distinction matters for practical skills. Realistic interaction may support embodied rehearsal and spatial understanding, but unnecessary visual complexity, unfamiliar controls or poorly signalled information can consume working-memory resources. Mayer's cognitive theory of multimedia learning predicts that instructional design should manage essential processing and minimize extraneous processing [7]. Accordingly, VR environments for skill training should prioritize task-relevant cues, intuitive interaction, staged complexity

and guidance rather than maximizing sensory stimulation for its own sake.

3. Review Methodology

This article uses a structured integrative-review approach because the research question spans heterogeneous disciplines, technologies and study designs. The analytic logic follows Whittemore and Knafl [8]: problem identification, literature searching, data evaluation, data analysis and presentation. The purpose is interpretive synthesis rather than a new meta-analysis. Accordingly, this article should not be read as a PRISMA systematic review or as an exhaustive census of all VR studies. Its contribution is to integrate higher-level quantitative evidence, systematic reviews and theoretically relevant literature in order to explain when and why VR may support practical-skill development in higher education.

Literature identification was conducted as a structured secondary research search rather than a fully preregistered systematic-review search. Search concepts combined technology terms ("virtual reality", "immersive virtual reality", "head-mounted display", "virtual simulation"), higher-education terms ("higher education", "university, college, undergraduate, student), and skill terms ("practical skills", "clinical skills", "technical skills", psychomotor, procedural, laboratory, performance). PubMed and scholarly publisher/index search interfaces were used for targeted retrieval, supplemented by backward reference checking and citation tracing from relevant reviews. Priority was given to peer-reviewed literature from 2015 onward, while earlier methodological and theoretical sources were retained when foundational. The search was updated through September 2026. Because contemporaneous records of database-specific retrieval yields, duplicate removal and sequential screening counts were not maintained, retrospective PRISMA flow counts are not reported.

Eligibility was determined conceptually against the review questions. Evidence was included when it addressed post-secondary or professional higher education learners; examined interactive VR, immersive VR or closely related virtual simulation; and reported practical, procedural, technical, psychomotor or directly relevant

performance outcomes. Reviews and controlled studies addressing instructional design, adverse effects, accessibility, implementation or transfer were also retained when they materially informed interpretation. School-level studies, entertainment applications and passive 360-degree video without an interactive practical-training component were outside the principal scope. Where multiple publications reported the same underlying study, they were treated as a single evidence source, with the most complete report used for the synthesis. The review was conducted by a single author; therefore, independent duplicate screening and inter-rater agreement were not applicable and this is acknowledged as a limitation.

Evidence evaluation was proportional to study design rather than converted into a single numerical quality score. Meta-analyses and systematic reviews were examined for explicit eligibility criteria, search transparency, sample/study counts, risk-of-bias or quality procedures where reported, heterogeneity and the appropriateness of the authors' conclusions. Controlled primary studies were interpreted with attention to comparator quality, sample size, outcome validity, intervention duration and whether performance was measured inside VR or on an external task. The synthesis was thematic across five dimensions: learning outcomes; instructional design; VR modality and interaction; retention/transfer; and implementation/accessibility. Greater evidential weight was given to replicated quantitative syntheses and controlled evidence, while descriptive studies and theory were used mainly to explain mechanisms or implementation. This design increases transparency but does not provide the exhaustiveness or formal dual-review safeguards of a systematic review.

4. Results and Integrative Synthesis

4.1 Overall pattern and strength of evidence

The evidence base now supports a more differentiated conclusion than the early literature on educational VR. Across higher education, VR is generally associated with positive practical-skill outcomes, but the size and reliability of those gains vary according to discipline, task characteristics, instructional design and the quality of the comparison

condition. The 2024 meta-analysis by Yang et al. [1] is particularly informative because it focused directly on practical skills in science and engineering education. Across 37 empirical studies and 72 effect sizes, the pooled effect was moderate (Hedges' $g = 0.477$). This indicates a meaningful average advantage, but it should not be interpreted as evidence that every VR intervention is superior to conventional instruction. The same meta-analysis detected publication bias and substantial variation across studies, reinforcing the need to interpret the pooled estimate as an average across heterogeneous contexts rather than a universal treatment effect.

Health-professions evidence is positive overall but becomes more nuanced when newer HMD-specific evidence is considered. Kim and Park [2] synthesized 11 randomized controlled trials and reported improved nursing clinical-skill performance ($g = 0.61$), while Lin et al. [3] reported a skill effect of 0.682 across eight randomized trials. Cho and Kim [17], synthesizing 22 studies, reported a larger pooled effect for nursing competency (Hedges' $g = 0.88$, 95% CI 0.47-1.29). In contrast, a 2026 meta-analysis restricted to immersive HMD simulation included 20 randomized trials (2,063 nursing students) and found significant benefits for knowledge and self-efficacy but no statistically significant pooled psychomotor-skill advantage (SMD = 0.49, 95% CI -0.24 to 1.22) [24]. The difference is important: conclusions depend on how VR is defined, which comparators are used and whether the outcome is knowledge, confidence or externally assessed skill. The most defensible conclusion is therefore that VR can support practical learning, but evidence is not uniformly positive across modalities or outcome domains.

This interpretation is consistent with broader immersive-learning reviews. Conrad et al. [19] restricted their 2024 systematic review to immersive HMD-based VR, whereas other reviews aggregate desktop, semi-immersive and fully immersive systems. Stracke et al. [20] similarly emphasize the need to distinguish VR by immersion level, and Santilli et al. [21] found that interactivity may be at least as important as immersiveness in comparative higher-education studies. In this review, evidence is therefore

interpreted using three broad technology categories: non-immersive VR (screen-based interactive 3D environments), semi-immersive/virtual simulation (enhanced visual or simulated environments without full HMD enclosure), and immersive HMD VR. Interaction method and haptic capability are considered separately because two systems with similar visual immersion may offer very different opportunities for motor practice.

4.2 Disciplinary variation and the nature of the target skill

The impact of VR differs across disciplinary contexts because practical skills are not a single class of outcome. Some skills depend primarily on procedural sequencing and spatial judgment; others depend on fine tactile discrimination, interpersonal communication, adaptive decision-making or the integration of several sensory channels. Yang et al. [1] identified disciplinary category as a significant moderator, with different effects across medical, engineering, science and agronomy learners. This finding is theoretically plausible: VR is most likely to add value when the essential structure of a task can be represented through visual-spatial information, interactive choices and repeatable action sequences.

In engineering and laboratory education, VR can represent equipment layouts, assembly sequences, hazardous procedures and spatially complex systems that students might otherwise encounter only briefly. The technology can therefore increase exposure before learners handle costly or dangerous equipment. However, the fidelity required should be defined functionally. For a lockout procedure, for example, the critical educational features may be recognizing hazards, selecting the correct sequence and verifying system state; photorealistic textures may add little. By contrast, a maintenance task requiring precise force, tool resistance or tactile diagnosis may be poorly represented by standard hand controllers. In such cases, VR can prepare cognition and sequence knowledge but cannot fully reproduce the motor demands of the real task.

Health-professions education illustrates the same distinction. VR is well suited to repeated rehearsal of clinical sequences, prioritization and

recognition of deteriorating scenarios, particularly where novice errors would be unsafe in real practice. Yet many clinical procedures also depend on tissue resistance, palpation, interpersonal interaction and subtle contextual cues. The positive meta-analytic effects in nursing [2], [3], [17] therefore support VR as a meaningful training component, not as evidence that virtual practice can replace supervised contact with physical task trainers, standardized patients or clinical placements.

4.3 Procedural, psychomotor and decision-making outcomes

A major strength of VR is its capacity to integrate cognitive and motor components of a task within the same learning episode. Traditional instruction often separates explanation from execution: students first learn what should happen and later attempt to perform it. VR can shorten this distance by requiring learners to make decisions while manipulating objects, moving through a space or completing a procedure. This integration is especially relevant to practical competence because expert performance depends on linking declarative knowledge to action under time and contextual constraints.

Nevertheless, psychomotor learning should not be treated as equivalent to controller proficiency. Performance gains inside a simulation may partly reflect increasing familiarity with the interface. A valid practical-skill assessment should therefore distinguish between actions that are specific to the VR system and actions that correspond to the target professional task. Studies that assess only in-world completion scores risk overestimating educational transfer. Stronger designs use an external performance test, physical simulator, objective structured assessment or real equipment after the VR intervention.

Decision-making is another important component. Virtual scenarios can expose learners to branching consequences, abnormal events and rare conditions that are difficult to schedule in physical teaching. The ability to pause, repeat and vary these situations can support pattern recognition and adaptive reasoning. However, decision quality depends on scenario design. If the system reduces

complex professional judgment to obvious multiple-choice prompts, apparent success may reflect cue recognition rather than authentic reasoning. High-quality practical VR should therefore preserve uncertainty and require learners to integrate task-relevant information rather than merely follow highlighted instructions.

4.4 Repetition, deliberate practice and feedback

The clearest pedagogical advantage of VR is not immersion by itself but the ability to create structured opportunities for deliberate practice. Expertise develops through repeated attempts at appropriately challenging tasks, feedback about errors and opportunities to correct performance [5]. Physical laboratories and placements often limit these opportunities because staff, equipment, consumables or appropriate cases are scarce. Once a simulation has been developed and deployed, VR can make additional repetitions comparatively easier to schedule and can ensure that every learner encounters core scenarios.

Repetition is educationally valuable only when practice quality is controlled. Repeating an incorrect procedure can consolidate errors. VR applications should therefore provide feedback that is timely enough to prevent persistent misconceptions while avoiding excessive guidance that makes learners dependent on prompts. Early attempts may use step cues, demonstrations or corrective feedback; later attempts should progressively remove assistance so that performance becomes independent. This fading of scaffolding provides a stronger test of competence than allowing the learner to complete every attempt with continuous visual guidance.

Feedback can be intrinsic, automated or instructor-mediated. Intrinsic feedback comes from the simulated consequences of an action. Automated feedback can identify missed steps, timing errors, collisions, contamination events or unsafe choices. Instructor feedback can address reasoning, professional standards and nuances that are difficult to encode in software. The strongest model combines these forms. Automated metrics are useful for immediate correction and repeated practice, while

debriefing helps learners interpret why an action mattered and how virtual performance maps onto authentic practice.

4.5 Immersion, presence and cognitive load

Immersion can strengthen the sense that a learner is acting within a task environment, but presence should be regarded as a mediating experience rather than an educational outcome. Makransky et al. [13] demonstrated an important tension: immersive VR increased presence in a science simulation but was also associated with greater cognitive load and poorer learning on some measures. Parong and Mayer [14] likewise showed that immersive experiences can require additional instructional support to convert engagement into durable learning. These findings challenge the assumption that increasing sensory richness automatically improves practical performance.

For novice learners, cognitive load may arise from three sources simultaneously: understanding the disciplinary content, remembering the procedure and operating an unfamiliar interface. Pre-training can reduce this competition for working-memory resources by teaching basic controls and essential concepts before the main scenario. Yang et al. [1] included pre-training among the moderators considered in their meta-analysis, reflecting its growing importance in VR instructional design. The implication is straightforward: time spent teaching learners how to use the technology should be separated from time intended to measure or develop the target skill.

Design simplicity is therefore a strength rather than a limitation when it directs attention to task-relevant information. Decorative realism, unnecessary movement and excessive interface elements can create extraneous processing. For practical training, fidelity should be selective: high for cues and interactions that influence correct performance, lower for features that do not contribute to the competency.

4.6 Blended instruction and the role of conventional practice

The evidence most strongly supports VR as part of a blended instructional sequence. In Yang et al. [1], practice-oriented approaches combined with traditional methods produced particularly favourable effects. This matters because VR and

physical practice have complementary strengths. VR can provide standardization, unlimited reset, controlled difficulty and safe error; conventional laboratories and placements provide authentic tools, tactile information, social interaction and exposure to the unpredictability of real environments.

A productive sequence is therefore preparation, virtual rehearsal, debriefing and authentic application. Conceptual instruction establishes the rationale for a procedure; orientation prevents interface difficulties from dominating the learning episode; VR provides repeated and progressively independent rehearsal; debriefing makes errors and decision processes explicit; and physical practice tests whether competence transfers beyond the simulation. Such sequencing also makes more efficient use of scarce laboratory or clinical time because students arrive with an established procedural model rather than encountering the task for the first time.

This blended interpretation avoids an unhelpful replacement debate. The relevant institutional question is not whether VR should replace laboratories or placements, but which components of a practical curriculum can be moved into virtual rehearsal without losing essential task information. In many disciplines, the highest value may come from shifting early-stage familiarization and error-prone repetition into VR while preserving advanced authentic practice for tasks that require real materials, haptics, teamwork or professional interaction.

4.7 Skill retention: beyond immediate post-test performance

A recurring weakness in the literature is the dominance of immediate post-intervention assessment. Improvement immediately after a VR session demonstrates acquisition, but higher education ultimately requires retention: the ability to perform after practice has ceased. Without delayed assessment, it is difficult to distinguish durable learning from short-lived familiarity with a scenario or interface.

Adjacent evidence reinforces the need to distinguish immediate acquisition from durable learning. Zhong et al. [18] reviewed 64 studies of VR-supported development of the 4Cs—creativity, critical thinking, collaboration and

communication—and found that only nine assessed retention. Because these outcomes are broader cognitive and interpersonal skills rather than direct procedural or psychomotor tasks, their findings should not be treated as direct evidence about clinical procedures, engineering manipulation or laboratory technique. They are nevertheless informative as adjacent evidence that delayed testing is uncommon in VR-supported skill research more generally. The present synthesis identifies the same methodological need within practical-skill education: immediate post-test improvement should not be assumed to represent durable competence.

For procedural training, retention studies should use delayed intervals that are educationally meaningful and should avoid providing extensive refresher exposure immediately before testing. Researchers should also report whether learners received additional conventional practice during the retention interval, because subsequent experience can obscure the independent contribution of the original VR intervention.

4.8 Transfer from virtual performance to authentic practice

Transfer is the most consequential and least adequately tested outcome in practical education. Zhong et al. [18] found only two transfer studies within their broader 4Cs review; this is adjacent rather than direct evidence for procedural education. It nevertheless illustrates a wider tendency for VR research to stop at immediate or near-term outcomes. Within the practical-skills literature synthesized here, external transfer measures are also less common than in-VR or immediate post-training assessments. Accordingly, this review does not use the 4Cs evidence to quantify a procedural transfer deficit; instead, it treats it as corroborating evidence for the need to test practical performance beyond the training environment.

Transfer should be expected when the simulation preserves the functional structure of the target task: the relevant decisions, spatial relationships, sequence, constraints and consequences. Surface resemblance alone is insufficient. A highly realistic environment may

produce weak transfer if learners use simplified interactions that do not correspond to real action, whereas a less visually elaborate simulation may transfer well if it accurately represents the decisions and motor sequence that matter.

The assessment of transfer should therefore be hierarchical. Near transfer can be tested using a different virtual scenario or physical simulator that requires the same underlying procedure. Farther transfer requires performance with authentic equipment, real cases or workplace-like conditions. Studies should report these outcomes separately from in-VR scores. Until such evidence becomes more common, claims that VR produces real-world competence should remain more cautious than claims about immediate skill acquisition.

4.9 Motivation, confidence and learner variability

Affective outcomes can support practical learning but should not be conflated with competence. Lin et al. [3] found positive effects on confidence and satisfaction alongside skill gains, and other reviews report high acceptance of immersive learning. Confidence may encourage students to attempt difficult tasks and persist through repeated practice. However, confidence without accurate performance feedback can become miscalibrated. Objective assessment is therefore necessary to determine whether increased self-efficacy reflects genuine competence.

Learner characteristics can also moderate outcomes. Prior gaming or VR experience may reduce the initial burden of navigation and controller use, while students with little exposure may require more orientation. Prior disciplinary knowledge can influence how effectively learners interpret simulated cues. Physical comfort, visual characteristics and susceptibility to cybersickness can determine whether a learner can participate for the intended duration. These differences mean that equal access to a headset does not necessarily produce equal access to the learning experience. Novelty effects should also be considered. Initial enthusiasm may increase attention and satisfaction during short studies, but the effect may decline once VR becomes routine. Longer

interventions and repeated-use studies are therefore more informative for institutional decision-making than single-session demonstrations.

4.10 Safety, accessibility and adverse effects

VR can provide a lower-consequence environment for practising hazardous or high-stakes tasks, but headset use introduces its own tolerability issues. A 2024 systematic review of healthcare-student VR studies reported generally low levels of cybersickness alongside high presence, while emphasizing that robust evidence remains limited [22]. A newer 2026 meta-analysis of immersive nursing simulation likewise reported predominantly mild visual and vestibular adverse effects [24]. Cybersickness, visual discomfort, disorientation and fatigue therefore should be measured and reported rather than assumed to be negligible, particularly when sessions are long or repeated.

Accessibility should be treated as an empirical and design issue rather than assumed from the flexibility of digital delivery. Dudley et al. [23] reviewed accessibility efforts across VR and AR and documented continuing barriers affecting users with different sensory, motor and other access needs. In higher education specifically, Santilli et al. [21] found that none of the 71 comparative studies in their review involved students with disabilities or specific learning disorders, highlighting a striking evidence gap. Practical accommodations may include alternative input methods, captions or multimodal cues, adjustable visual presentation, seated modes and an equivalent non-headset pathway, but the suitability of any accommodation should be tested with the learners it is intended to support rather than inferred from average-user studies.

Adverse effects and accessibility also matter methodologically. If participants who experience discomfort withdraw, cannot use the interface, or are excluded before enrolment, reported outcomes may overrepresent learners who tolerate the technology well. Future studies should therefore report adverse events, attrition, accessibility exclusions and reasons for non-completion alongside educational outcomes. Inclusive evaluation is especially important because current comparative higher-education

evidence contains very little direct participation by disabled students [21], [23].

4.11 Cost, scalability and institutional sustainability

Lower-cost standalone headsets have reduced the initial hardware barrier, but implementation costs extend beyond device purchase. Recent higher-education reviews identify acquisition and maintenance costs, specialized software, technical requirements, usability and infrastructure as recurring barriers to adoption and scale [20], [25]. Institutions may additionally incur local costs for device management, storage, charging, hygiene, technical support, instructor preparation and curriculum integration. Because these components vary by institution and application, the present review treats them as implementation considerations rather than claiming a universal cost advantage for VR.

Scalability is similarly context dependent. A simulation demonstrated with a small research cohort may require different scheduling, supervision and support when used across a large programme. Conversely, independent virtual rehearsal may expand access where laboratory time, consumables or specialist facilities are constrained. Current higher-education reviews continue to identify cost, infrastructure, usability and equity as barriers [20], [25]. Cost-effectiveness should therefore be evaluated locally using both educational outcomes and operational data rather than inferred from headset price or short-term efficacy alone.

Institutional evaluation should combine educational and operational indicators. Relevant measures include performance improvement, transfer, learner completion, technical failure rates, staff time, device utilization, laboratory hours saved and the marginal cost of additional practice. The literature currently provides much stronger evidence about learning efficacy than about long-term economic sustainability, leaving an important gap for university decision-makers.

4.12 Quality of the evidence and sources of uncertainty

Although the direction of evidence is generally favourable, several methodological issues limit certainty. Many individual studies use small

samples, short interventions and locally developed outcome measures. Blinding is often impossible, and participants may know that they are receiving a novel technology. Intervention descriptions are sometimes insufficient to reproduce the experience, particularly with respect to interaction design, feedback, pre-training and the amount of practice. These weaknesses make it difficult to identify which component of an intervention caused an observed effect.

Publication bias is another concern. Yang et al. [1] reported a significant Egger test, although trim-and-fill analysis did not change the combined effect estimate. This suggests that the positive overall result is reasonably robust while still requiring caution about the distribution of published findings. The field would benefit from preregistered trials, clearer reporting of null results and greater use of common performance measures.

Conceptual heterogeneity in the term 'virtual reality' is itself a source of uncertainty. Studies may label desktop 3D simulations, semi-immersive systems, HMD-based immersive environments, virtual patients and interactive panoramic experiences as VR. Conrad et al. [19] addressed this problem by restricting their synthesis to HMD-based immersive VR, while Stracke et al. [20] explicitly distinguish non-immersive, semi-immersive and immersive forms. The present review therefore avoids treating immersion as a binary exposure and interprets findings in relation to display modality, interactivity and haptic capability. This classification does not eliminate heterogeneity, but it makes the boundaries of broad claims about 'VR effectiveness' more explicit.

4.13 Integrative interpretation of the evidence

Taken together, the literature supports a conditional model of effectiveness. VR contributes most clearly when it expands meaningful practice opportunities: learners can perform a relevant task, make consequential decisions, receive informative feedback and repeat the task under controlled variation. The technology contributes less when it functions primarily as an immersive presentation medium

or when essential physical properties of the target skill cannot be represented.

Three propositions emerge from the synthesis. First, practical-skill gains depend more on task alignment and instructional design than on immersion alone. Second, VR is most defensible as a complement to authentic practice, particularly during early and intermediate stages of skill development. Third, the evidence is considerably stronger for acquisition than for retention and real-world transfer. These propositions reconcile the generally positive meta-analytic findings with the continued caution expressed in systematic reviews.

The resulting implication for higher education is that adoption should be competency-led rather than technology-led. A university should first identify the performance problem, determine which elements of the skill can be represented faithfully in VR, design feedback and progression around those elements, and then verify learning outside the virtual environment. When these conditions are met, VR can extend access to deliberate practice and improve preparedness for authentic tasks. When they are absent, immersion risks becoming an expensive substitute for more appropriate instructional methods.

5. Discussion

The integrated evidence supports a conditional rather than universal account of VR effectiveness. Meta-analyses in science, engineering and nursing report positive practical or competency effects [1]-[3], [17], yet the newer HMD-only nursing synthesis found no statistically significant pooled psychomotor advantage despite benefits for knowledge and self-efficacy [24]. These findings are not necessarily contradictory: they arise from different definitions of VR, outcome domains, comparators and intervention designs. The central implication is that practical-skill effectiveness cannot be inferred from immersion alone.

The most coherent interpretation is one of instructional complementarity. VR is particularly useful when it expands safe, repeatable and feedback-rich rehearsal before access to scarce or high-stakes authentic practice. Physical laboratories, placements and task

trainers remain important where tactile information, material resistance, interpersonal interaction or unpredictable real-world constraints are integral to competence. This position is consistent with Yang et al. [1] and with comparative higher-education evidence showing that interactivity and active pedagogy matter alongside technical immersion [21].

A second implication concerns evidence standards. Performance inside VR should be treated as a proximal outcome, not automatically as proof of occupational or clinical competence. Delayed retention and external transfer tests are needed, and adjacent evidence from broader VR skill research indicates that both remain comparatively uncommon [18]. For practical disciplines, the strongest studies will therefore separate in-world performance, physical-simulator performance and authentic-task performance.

A third implication is that implementation research must include learners who may experience barriers. Cybersickness is usually mild in the limited healthcare evidence but remains relevant [22], [24], while comparative higher-education studies have rarely included disabled students [21]. Evaluation should therefore incorporate tolerability, accessibility, attrition and alternative access pathways alongside learning outcomes.

Taken together, the evidence shifts the decision question from whether VR 'works' to whether a particular VR design adds educational value for a defined competency, learner population and institutional context. This framing also provides the basis for the conceptual framework proposed below.

6. Proposed Framework for VR-Supported Practical Skills

The six-stage framework in Table 1 is an author-derived conceptual framework generated from the integrative synthesis; it has not been independently validated as a predictive model. Its stages map to recurring evidence and theory: defining competence reflects the need for task-aligned outcomes and external performance measures [1], [19], [21]; pre-training is supported by cognitive-load/CAMIL principles [6], [7], [13], [14]; guided practice and deliberate repetition draw on deliberate-practice theory [5];

debriefing and reflection are supported by simulation literature showing the educational importance of structured feedback and debriefing [26], [27]; and authentic transfer reflects the evidence gap identified in this review

and adjacent transfer literature [18]. The framework is therefore offered as a transparent synthesis-based design heuristic for future testing, not as an empirically validated standard.

Table 1. Proposed six-stage framework for VR-supported practical-skills education.

Stage	Purpose	Recommended practice	Evidence/assessment
1. Define competence	Specify the real-world skill.	Break the skill into observable decisions, actions and safety criteria.	Task alignment and external performance measures [1], [19], [21].
2. Pre-train	Reduce interface-related cognitive load.	Orient learners to controls, navigation, safety and core concepts before complex scenarios.	CAMIL, multimedia learning and cognitive-load evidence [6], [7], [13], [14].
3. Guided VR practice	Build an accurate procedural model.	Use cues, demonstration, scaffolding and immediate feedback for early attempts.	Deliberate practice and guided feedback [5].
4. Deliberate repetition	Develop fluency and adaptability.	Repeat with fading guidance and controlled variation in scenario conditions.	Deliberate practice with progressively independent performance [5].
5. Debrief and reflect	Connect action to disciplinary reasoning.	Review decisions, errors and alternative strategies with an instructor or structured prompts.	Simulation debriefing and structured reflection [26], [27].
6. Authentic transfer	Verify competence beyond VR.	Move to physical equipment, simulation lab, placement or authentic performance task.	External transfer and retention gap [18], [24].

7. Implications for Higher Education Practice

Institutions should begin with a defined performance problem rather than a technology purchase. VR is most plausible where conventional practice is constrained by risk, rarity, cost, spatial complexity or limited access, but the relevant task features must be representable in the chosen system. Comparative and immersive-VR reviews emphasize that interactivity, pedagogical integration and research design influence

outcomes at least as much as the label 'VR' itself [19]-[21].

Curriculum integration should specify what students do before, during and after VR. Pre-training can reduce interface-related load [6], [7], [13], [14]; repeated practice should include feedback rather than unguided repetition [5]; and structured debriefing can connect simulated action to disciplinary reasoning [26], [27]. Where the intended outcome is practical competence, assessment should ultimately

include performance outside the training interface.

Accessibility and inclusion require explicit planning and evidence. Reviews of immersive accessibility document unresolved barriers across sensory and motor access needs [23], while a 2025 comparative higher-education review found no included experiments involving students with disabilities or specific learning disorders [21]. Institutions should therefore provide reasonable alternative access pathways and evaluate accommodations with affected learners rather than assuming that a standard headset-based design is universally accessible.

Operational evaluation should accompany educational evaluation. Higher-education reviews identify cost, infrastructure, maintenance, technical requirements and equity as recurring adoption barriers [20], [25]. Institutions should therefore monitor local measures such as staff support, device utilization, technical failure, learner completion and cost per learner alongside competence outcomes; this is an implementation recommendation rather than evidence that VR is inherently cheaper than conventional teaching.

8. Limitations of the Review

This review has important limitations. It is a structured integrative review rather than a preregistered systematic review: database-by-database retrieval counts, duplicate-removal logs and dual-review screening records were not maintained, so exhaustive coverage and formal PRISMA reproducibility cannot be claimed. The search and synthesis were conducted by a single author, increasing the possibility of selection and interpretation bias. The underlying literature is also heterogeneous in its definitions of VR, level of immersion, interaction methods, disciplines and performance measures. Practical skills are assessed differently across fields, limiting direct comparison, and this review does not calculate a new pooled effect. Priority was given to peer-reviewed English-language evidence discoverable through PubMed, publisher/index searches and citation tracing; relevant studies in other languages or less visible databases may have been missed. Finally, rapid changes in hardware and software mean that implementation findings may age faster than

pedagogical principles. These limitations are the reason the conclusions are framed as an integrative interpretation rather than an exhaustive estimate of effect.

9. Future Research Directions

Future research should place greater emphasis on delayed retention and authentic transfer. Studies should assess whether performance gains persist weeks or months after training and whether learners can execute the skill on real equipment, in laboratories, with patients or in workplace-like settings. Transfer measures should be defined before data collection and reported separately from in-VR performance.

More rigorous comparative studies are also needed. Rather than comparing VR only with no intervention, researchers should compare it with credible alternatives such as physical simulation, video demonstration, desktop simulation and blended practice. Reporting should describe hardware, software, field of view, interaction method, session duration, feedback, pre-training and debriefing sufficiently for replication.

Cost-effectiveness is another major gap. Educational gains should be considered alongside development costs, hardware life cycles, staffing, throughput and any reduction in physical consumables or laboratory demand. Multi-institution studies would improve generalizability and help determine whether locally successful systems remain effective at scale.

Finally, accessibility, cybersickness and differential effects across learner groups require direct study. Existing healthcare evidence suggests that cybersickness is often mild but the evidence base remains limited [22], [24], while comparative higher-education research has rarely included disabled students [21]. Future trials should report adverse effects and attrition, recruit more diverse learners, describe accommodations, and test whether inclusive design changes preserve or improve practical-skill outcomes.

10. Conclusion

Virtual reality has a credible but heterogeneous evidence base for supporting practical learning in higher education. Positive meta-analytic

findings in science, engineering and nursing indicate that VR can improve selected practical or competency outcomes [1]-[3], [17], while newer HMD-specific evidence shows that psychomotor advantages are not consistently statistically significant [24]. The appropriate conclusion is therefore not that VR is superior by default, but that it can add value when technology, task and pedagogy are well aligned. The most defensible role for VR is as a structured practice environment within a broader curriculum. It can increase opportunities for safe repetition and controlled variation, while authentic laboratories, placements and physical task environments remain necessary where tactile, social and contextual information is essential. The author-derived six-stage framework translates this synthesis into a testable design heuristic rather than a validated model.

Future evidence should prioritize external performance, delayed retention, real-world transfer, accessibility and implementation sustainability. Clearer differentiation among non-immersive, semi-immersive and HMD-based VR will also improve interpretation. Progress in these areas would allow universities to judge VR by durable competence and equitable educational value rather than novelty or presence alone.

11. Declarations

11.1 Author Contributions

The sole author was responsible for the conceptualization, literature identification and synthesis, manuscript preparation, revision, and final approval of the manuscript.

11.2 Acknowledgment

Not applicable.

11.3 Funding

No external funding was received for this review.

11.4 Conflict of Interest

The author declares no conflict of interest.

11.5 Institutional Review Board Statement

Not applicable. This study is an integrative review of previously published literature and did not involve human participants or animal subjects.

11.6 Informed Consent Statement

Not applicable.

11.7 Data Availability Statement

No new primary dataset was generated. The evidence synthesized in this review is available in the published sources cited in the reference list.

12. REFERENCES

- [1] C. Yang, J. Zhang, Y. Hu, X. Yang, M. Chen, M. Shan, and L. Li, "The impact of virtual reality on practical skills for students in science and engineering education: a meta-analysis," *International Journal of STEM Education*, vol. 11, art. 28, 2024. doi: 10.1186/s40594-024-00487-2.
- [2] Y. Kim and H.-Y. Park, "Effects of virtual reality training on clinical skill performance in nursing students: A systematic review, meta-analysis and meta-regression," *International Journal of Nursing Practice*, vol. 30, no. 6, e13284, 2024. doi: 10.1111/ijn.13284.
- [3] M.-Y. Lin, M.-Z. Huang, and P.-C. Lai, "Effect of virtual reality training on clinical skills of nursing students: A systematic review and meta-analysis of randomized controlled trials," *Nurse Education in Practice*, vol. 81, 104182, 2024. doi: 10.1016/j.nepr.2024.104182.
- [4] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1984.
- [5] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The role of deliberate practice in the acquisition of expert performance," *Psychological Review*, vol. 100, no. 3, pp. 363-406, 1993. doi: 10.1037/0033-295X.100.3.363.
- [6] G. Makransky and G. B. Petersen, "The Cognitive Affective Model of Immersive Learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality," *Educational Psychology Review*, vol. 33, pp. 937-958, 2021. doi: 10.1007/s10648-020-09586-2.
- [7] R. E. Mayer, *Multimedia Learning*, 3rd ed. Cambridge, UK: Cambridge University Press, 2021.

- [8] R. Whittemore and K. Knafl, "The integrative review: updated methodology," *Journal of Advanced Nursing*, vol. 52, no. 5, pp. 546-553, 2005. doi: 10.1111/j.1365-2648.2005.03621.x.
- [9] J. Ma, "Research on the influence of virtual reality on the learning effect of technical skills of science and engineering college students: Meta-analysis based on 32 empirical studies," *International Journal of Digital Multimedia Broadcasting*, vol. 2022, art. 6202370, 2022. doi: 10.1155/2022/6202370.
- [10] L. Jensen and F. Konradsen, "A review of the use of virtual reality head-mounted displays in education and training," *Education and Information Technologies*, vol. 23, pp. 1515-1529, 2018. doi: 10.1007/s10639-017-9676-0.
- [11] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & Education*, vol. 147, 103778, 2020. doi: 10.1016/j.compedu.2019.103778.
- [12] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, and T. J. Davis, "Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis," *Computers & Education*, vol. 70, pp. 29-40, 2014. doi: 10.1016/j.compedu.2013.07.033.
- [13] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," *Learning and Instruction*, vol. 60, pp. 225-236, 2019. doi: 10.1016/j.learninstruc.2017.12.007.
- [14] J. Parong and R. E. Mayer, "Learning science in immersive virtual reality," *Journal of Educational Psychology*, vol. 110, no. 6, pp. 785-797, 2018. doi: 10.1037/edu0000241.
- [15] M. Kyaw et al., "Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration," *Journal of Medical Internet Research*, vol. 21, no. 1, e12959, 2019. doi: 10.2196/12959.
- [16] R. Shorey and E. D. Ng, "The use of virtual reality simulation among nursing students and registered nurses: A systematic review," *Nurse Education Today*, vol. 98, 104662, 2021. doi: 10.1016/j.nedt.2020.104662.
- [17] M. K. Cho and M. Y. Kim, "Enhancing nursing competency through virtual reality simulation among nursing students: a systematic review and meta-analysis," *Frontiers in Medicine*, vol. 11, 1351300, 2024. doi: 10.3389/fmed.2024.1351300.
- [18] Y. Zhong, K. Guo, L. Xu, and L. K. Fryer, "From virtual to reality: A systematic review of the impact of virtual reality on skill acquisition, retention, and transfer," *Educational Research Review*, vol. 50, art. 100767, 2026. doi: 10.1016/j.edurev.2026.100767.
- [19] M. Conrad, D. Kablitz, and S. Schumann, "Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings," *Computers & Education: X Reality*, art. 100053, 2024. doi: 10.1016/j.cexr.2024.100053.
- [20] C. M. Stracke et al., "Immersive virtual reality in higher education: a systematic review of the scientific literature," *Virtual Reality*, vol. 29, art. 64, 2025. doi: 10.1007/s10055-025-01136-x.
- [21] T. Santilli, S. Ceccacci, M. Mengoni, and C. Giaconi, "Virtual vs. traditional learning in higher education: A systematic review of comparative studies," *Computers & Education*, vol. 227, art. 105214, 2025. doi: 10.1016/j.compedu.2024.105214.
- [22] E. Akyüz and Ş. Ergöl, "The Effects of Virtual Reality Used in Healthcare Education on Cybersickness and Sense of Presence: A Systematic Review," *PRESENCE: Virtual and Augmented Reality*, vol. 33, pp. 161-178, 2024. doi: 10.1162/pres_a_00414.

- [23] J. Dudley, L. Yin, V. Garaj, and P. O. Kristensson, "Inclusive Immersion: a review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse," *Virtual Reality*, vol. 27, pp. 2989–3020, 2023. doi: 10.1007/s10055-023-00850-8.
- [24] X. Li et al., "Effects of Immersive Virtual Reality Simulation on Nursing Students' Learning Outcomes: A Systematic Review and Meta-Analysis," *Journal of Nursing Management*, 2026, art. 4686851. doi: 10.1155/jonm/4686851.
- [25] D. Llanos-Ruiz, V. Abella-García, and V. Ausín-Villaverde, "Virtual Reality in Higher Education: A Systematic Review Aligned with the Sustainable Development Goals," *Societies*, vol. 15, no. 9, art. 251, 2025. doi: 10.3390/soc15090251.
- [26] J. P. Duff et al., "Debriefing Methods for Simulation in Healthcare: A Systematic Review," *Simulation in Healthcare*, vol. 19, no. 1S, pp. S112–S121, 2024. doi: 10.1097/SIH.0000000000000765.
- [27] A. Cheng et al., "Debriefing for technology-enhanced simulation: a systematic review and meta-analysis," *Medical Education*, vol. 48, no. 7, pp. 657–666, 2014. doi: 10.1111/medu.12432.

