University of Kent

Kent Academic Repository

Hamash, Mahmoud, Tabbaa, Luma, Tiernan, Peter and Young, Gareth W. (2026) ***Co-designing VR chemistry labs: Pre–post learning changes, reshaped classroom dynamics & deployment lessons for real-world rollout in post-primary schools.*** **Computers & Education: X Reality, 9 .**

Downloaded from

https://kar.kent.ac.uk/117148/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.1016/j.cexr.2026.100184

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Computers & Education: X Reality 9 (2026) 100184

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Co-designing VR chemistry labs: Pre–post learning changes, reshaped classroom dynamics & deployment lessons for real-world rollout in post-primary schools

Mahmoud Hamash [a,*], Luma Tabbaa [b], Peter Tiernan [a], Gareth W. Young [c]

[a] DCU Centre for AI and Digital Futures in Education, Institute of Education, Dublin City University, Dublin, Ireland
[b] School of Engineering, Mathematics & Physics, University of Kent, Canterbury, Kent, United Kingdom
[c] School of Computer Science and Statistics, Trinity College Dublin, Dublin, Co. Dublin, Ireland

ARTICLE INFO

Keywords:
Immersive virtual reality
Chemistry education
Post-primary education
Participatory design
Virtual laboratories
Classroom orchestration
Procedural knowledge
Virtual reality in education

ABSTRACT

Experiential learning is central to chemistry education, yet post-primary schools face persistent constraints related to cost, safety, and infrastructure that limit hands-on laboratory access. While immersive VR has demonstrated potential for conceptual and procedural learning, comparatively less is known about how VR-based procedural chemistry learning can be co-designed for adolescents, integrated with teacher orchestration, and implemented under authentic school constraints. This paper investigates how Virtual Reality (VR) can be designed and implemented as a deployable, pedagogically grounded classroom infrastructure for post-primary chemistry education. We report findings from a 25-month, three-phase participatory co-design study involving 137 stakeholders, including students, teachers, pre-service teachers, school leaders, parents, and domain experts. The study culminated in the development and real-world classroom deployment of a VR-based acid–base titration experience. Quantitative analyses showed significant pre–post increases following participation in the VR experience in conceptual knowledge (pre: $M = 7.81$, post: $M = 9.52$, $p < .001$) and curriculum-aligned procedural knowledge (pre: $M = 4.71$, post: $M = 9.14$, $p < .001$). Qualitative findings indicated learner autonomy, confidence, and engagement during VR-supported activities. Beyond individual learning outcomes, the study illustrates how VR can reshape classroom ecology by enabling self-paced procedural practice and shifting teacher roles toward facilitation and orchestration. Findings also highlight the need to balance adolescent learners' autonomy with appropriate scaffolding through adjustable guidance, pacing control, and on-demand support. Finally, the study identifies key sociotechnical conditions for sustainable classroom integration, including alignment with teacher practices, assessment needs, and classroom constraints. By linking learning outcomes with classroom deployment and design rationale, this work contributes a holistic account of how immersive technologies can move beyond novelty toward viable use in post-primary education.

1. Introduction

Chemistry education is inherently practice-oriented: learners are expected to learn not only chemical concepts, but also how to *do* chemistry, handling apparatus, following protocols, conducting experiments, and interpreting experimental outcomes through repeated laboratory work. However, the conditions required to sustain this form of experiential learning are increasingly difficult to guarantee in everyday school realities. Post-primary schools face persistent constraints that restrict opportunities for hands-on experimentation, including limited laboratory infrastructure, rising consumable costs, and stringent safety regulations (Atabek, 2019). This gap is particularly acute in Ireland, where practical laboratory competence is a core requirement of the Leaving Certificate chemistry curriculum, yet national readiness surveys (ASTI Research & Surveys, 2025) highlight substantial barriers to delivering these experiences: around 50% of school leaders report insufficient laboratory space, 44% note that existing labs are not adequately equipped for the delivery of the curriculum, and 47% face ongoing shortages of qualified science teachers who can conduct hands-on experiments. These constraints significantly limit students'

* Corresponding author.
E-mail addresses: mahmoud.hamash2@mail.dcu.ie (M. Hamash), L.A.Tabbaa@kent.ac.uk (L. Tabbaa), peter.d.tiernan@dcu.ie (P. Tiernan), gareth.young@tcd.ie (G.W. Young).

https://doi.org/10.1016/j.cexr.2026.100184
Received 23 April 2026; Received in revised form 20 September 2026; Accepted 29 September 2026
Available online 8 October 2026
2949-6780/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

opportunities to develop crucial conceptual, practical, and Higher-Order Thinking (HOT) skills (e.g., problem-solving, reflection, self-regulation) (Fisser & Thijs, 2015), underscoring the need for alternative, scalable approaches that can supplement laboratory provision while maintaining curricular fidelity and procedural rigour.

Immersive technologies such as Virtual Reality (VR) have gained increasing attention within Human-Computer Interaction (HCI) research due to their potential to transform how learners engage with complex scientific ideas and practical tasks. VR has been shown to deepen conceptual understanding through embodied interaction (Makransky & Petersen, 2021; Mills & Brown, 2022), make abstract or sub-microscopic processes visible (Chatain et al., 2023, pp. 1–8; Walsh et al., 2025), and provide safe, repeatable environments for rehearsing hazardous or resource-intensive laboratory procedures (Chua et al., 2025; Van Dinther et al., 2023). These affordances position VR as a promising medium for practical science learning, such as chemistry, where conceptual abstraction and procedural precision are central. Despite this promise, most VR research has focused on higher education or adult vocational training (Hamash & Tiernan, 2024, pp. 1–6; Stracke et al., 2025), leaving adolescent learning contexts comparatively underexplored. Only some studies have explored the use and design of VR for adolescents in other contexts, such as biology (Christopoulos et al., 2023), history and culture (Cecotti et al., 2024; Christopoulos et al., 2024), and earthquake escape training (Ye & Fan, 2025, pp. 1–6). As such, it remains unclear how the benefits reported for adult learners in VR can be translated to the context of post-primary chemistry in ways that are developmentally appropriate for adolescents, including their cognitive and representational capacities, attentional and motivational profiles, and physical and visual comfort requirements.

A further limitation in existing research is that VR is often designed and evaluated with limited engagement from all stakeholders who shape whether classroom innovations can be enacted, sustained, and scaled. While adolescents' perspectives are important—particularly in VR, where interaction design, cognitive load, and affective responses can materially influence learning outcomes (Makransky & Petersen, 2021), realistic deployment depends on more than designing for learners alone: teachers, school leaders, laboratory technicians, parents, and subject experts collectively influence feasibility, safety, procurement, and maintenance, and pedagogical legitimacy (Hamash & Tiernan, 2024, pp. 1–6). These tensions underscore the need for research that examines VR not only as a learning tool but also as a practical, realistically deployable pedagogical solution in everyday classroom settings.

To address these gaps, this paper investigates how VR can be meaningfully designed, deployed, and evaluated for practical chemistry learning in post-primary schools. We aim to investigate how a co-designed VR chemistry laboratory experience can be enacted as a meaningful learning tool in post-primary schools. Specifically, we examine pre–post changes in adolescents' conceptual understanding and procedural knowledge following participation in the co-designed VR chemistry laboratory experience, alongside the transferable competencies reported by participants in the VR activity. We then analyse how introducing VR into chemistry teaching reconfigures classroom ecology, particularly teacher–student and student–student dynamics, orchestration practices, and patterns of participation. Finally, we identify what is required for VR to become routinised within everyday post-primary classroom practice, clarifying the sociotechnical constraints and enabling conditions that shape feasible rollout. We contribute to both educational and HCI communities with empirical evidence of learning and wider developmental outcomes from an in-school VR deployment, alongside design and rollout implications that treat classroom ecology and institutional constraints as first-class elements of immersive learning system design.

To guide our investigation, we address the following research questions:

RQ1: To what extent do students' conceptual understanding and procedural knowledge change from pre-to post-participation in a co-designed VR chemistry laboratory experience?
RQ2: How do students and teachers experience the usability, engagement, and pedagogical value of the VR system in classroom settings?
RQ3: How does the introduction of VR reshape classroom ecology, including teacher–student roles, interaction patterns, and orchestration practices?
RQ4: What sociotechnical factors influence the feasibility and sustainability of deploying VR in real-world post-primary classrooms?

2. Related work

2.1. Chemistry learning in schools: developmental challenges and educational consequences

Chemistry learning at the secondary school level presents persistent challenges that differ fundamentally from learning chemistry as an adult or expert. Unlike adult learners, adolescents are still developing formal operational thinking, spatial reasoning, and the ability to coordinate multiple representational systems (Piaget & Inhelder, 2013). These developmental constraints contribute to robust misconceptions about core topics such as acid–base reactions, chemical bonding, and equilibrium, which are resistant to correction through lecture-based or text-based instruction alone (Üce & Ceyhan, 2019). When adolescents' conceptual understanding remains fragmented, they may succeed in "memorising" solutions without grasping underlying principles, resulting in fragile learning that quickly decays and reinforces perceptions of chemistry as abstract, difficult, and disconnected from real-world relevance (Batamuliza et al., 2024). Importantly, the consequences extend beyond poor academic performance; the lack of conceptual understanding in chemistry has been linked to reduced self-efficacy, disengagement, and ultimately, declining interest in Science, Technology, Engineering & Mathematics (STEM) pathways (Jegstad, 2024; Orosz et al., 2023). Research shows that nurturing students' interest in STEM, and chemistry in particular, supports scientific literacy, informed environmental choices, and future STEM career pathways (Damyana, 2024).

Inquiry-oriented and experiential chemistry pedagogies that scaffold the development of HOT skills (e.g., exploration, critical thinking, autonomy in learning, self-regulation, confidence) have been shown to enhance students' conceptual learning, support engagement, and offer positive learning experiences (Fisser & Thijs, 2015). These competencies are not only educationally valuable within schooling; they are foundational for progression to university life, employability, and civic participation in technology-enabled societies (J. Liu et al., 2024). However, the implementation of laboratory-based enquiry-oriented experiential activities in schools is constrained by curricular pressure, safety concerns, limited laboratory resources, and a lack of appropriate support for teachers (Orosz et al., 2023). Consequently, there is a growing need for pedagogical approaches that support experiential learning, develop chemistry procedural knowledge, foster transferable HOT skills, generate interest in chemistry, and scaffold learning processes that align with adolescents' cognitive development and school realities.

2.2. HCI and immersive technologies for adolescent chemistry learning

HCI research has developed interactive technologies, such as simulations (Bower et al., 2014), serious games (Fang, 2021), tangible interfaces (Ishii & Ullmer, 1997, pp. 234–241), and immersive environments (Cecotti et al., 2024), to address chemistry's representational and motivational challenges by externalising invisible processes and supporting embodied, exploratory learning. Research found that the use of interactive technologies for chemistry learning can reduce

cognitive load and help learners connect symbolic equations with molecular-level behaviour (Van Dinther et al., 2023). VR is particularly promising because it can visualise sub-microscopic phenomena, safely simulate hazardous experiments, and situate laboratory learning in immersive contexts, with studies reporting gains in understanding of molecular geometry, reaction mechanisms, and laboratory procedures (Qorbani et al., 2024; Rahman et al., 2024). Whilst the use of VR has been widely investigated in education contexts, the vast majority of studies focus on adult learners in higher education (Jin et al., 2022, 2024; Woźniak et al., 2020) or professional training (Sabrie et al., 2025; Yang et al., 2024). Far fewer studies have examined how VR can support adolescent learning in educational contexts, despite critical differences from adult learners; engaging adolescents with VR introduces unique developmental, cognitive, emotional, and safeguarding considerations that shape how VR should be designed and used safely, ethically, and responsibly (Kaimara et al., 2022). For instance, depending on how VR is being designed and instrumented, it can generate sensitive data streams (e.g., interaction logs, behavioural traces, potentially biometric or positional data), which in turn raises questions around privacy, an ethical issue that is of heightened importance when it comes to adolescent users (Kablo et al., 2025). In parallel, well-documented wellbeing risks, including nausea, dizziness, and visual fatigue, can arise during VR use (Sharples et al., 2008). Finally, designing VR for adolescents necessitates developmentally appropriate, age-sensitive content and interaction constraints to avoid exposure to unsuitable material and to mitigate potential impacts on behavioural and emotional regulation—concerns that families have also raised regarding children's VR use (Cao et al., 2024, pp. 1554–1572).

Furthermore, the pedagogical realities of designing and delivering curricula for adolescents have distinct implications for the design and deployment of VR in this context. Unlike adult education, which often assumes learner independence and minimal scaffolding, adolescent learning takes place in tightly structured classroom contexts where pacing, teacher facilitation, and Special Educational Needs (SEN) inclusivity are critical to learners' success (Southgate, 2022). Where the design of VR (Christopoulos et al., 2023; Ye & Fan, 2025, pp. 1–6) or AR (Chen et al., 2024; Moon et al., 2025, pp. 1–8; Woodward et al., 2022; Woźniak et al., 2020; Wu et al., 2023, pp. 1–18) has been explored specifically for adolescent learning, there has been limited co-design and engagement with stakeholders, including students, post-primary teachers, and school leaders, whose perspectives are critical to its deployment, integration, and adoption into school curricula. As a result, perspectives on the pedagogical value, practical feasibility, and mitigating risks of VR in adolescent learning remain largely absent. Accordingly, we still lack evidence on how to design and deploy VR that is both developmentally appropriate for adolescents and practically adoptable in schools. We therefore co-design a VR chemistry lab with key stakeholders to support conceptual understanding and procedural knowledge acquisition, examine emerging transferable competencies and shifts in classroom ecology and student–teacher dynamics, and derive rollout lessons grounded in authentic deployment constraints and ethical, privacy, and wellbeing considerations.

2.3. Research gaps and positioning of this work

Despite substantial prior work on virtual laboratories, molecular visualisation tools, and VR-based experimental training in chemistry education, important gaps remain. First, immersive VR for post-primary (secondary-level) chemistry education remains underexplored. A recent review of the literature by Hamash, Ghreir, et al. (2025a), conducted according to PRISMA guidelines, evaluated research on Immersive Virtual Reality (IVR) in post-primary education between 2013 and 2024. From a search across major databases, including ScienceDirect, ERIC, JSTOR, Taylor & Francis, and Scopus, only 31 studies were focused on VR for post-primary education. While the findings indicated a strong overall focus on STEM and noted that VR-enhanced experiential learning significantly boosted student motivation and conceptual understanding, only one of these 31 studies specifically addressed chemistry. This scarcity of chemistry-focused research is further supported by another systematic literature review conducted by Van Dinther et al. (2023). In a broader systematic review of 3406 publications regarding IVR in post-primary education, only 30 studies met the inclusion criteria. Of these, only seven were related to chemistry, and the rigor of even that small sample is notable: two of the seven were limited to a poster presentation and a conference series, respectively.

Second, although procedural learning is increasingly represented in VR research, chemistry studies have more commonly emphasised conceptual understanding, molecular visualisation, or learning outcomes than the design and classroom enactment of procedural laboratory training for adolescent learners (Hamash & Tiernan, 2024, pp. 1–6; Hamerská et al., 2024; Makransky & Petersen, 2021; Van Dinther et al., 2023; M. Wiedbusch et al., 2025; M. D. Wiedbusch et al., 2026). Where procedural training is considered, it is often evaluated in controlled or higher-education settings rather than within real classroom environments. The broader evidence base increasingly supports the potential of VR for procedural and practical learning. Systematic review evidence indicates that immersive VR can support active, practice-oriented learning (Conrad et al., 2024), while a recent meta-analysis reported a moderate positive effect of VR on practical skills in science and engineering education (Yang et al., 2024). The unresolved question is therefore not simply whether VR can support procedural learning, but how VR-based procedural chemistry learning can be designed for adolescent learners, integrated with teacher practice, and implemented under authentic school constraints.

Third, existing studies frequently treat VR as an isolated instructional intervention, focusing on individual learning outcomes while abstracting away from the realities of classroom implementation, including teacher orchestration, institutional constraints, and long-term integration. In addition, stakeholder engagement is often limited, with relatively little involvement from the broader ecosystem of actors, such as school leaders, parents, and practitioners, who shape adoption and sustainability (Docter et al., 2024; Hamash et al., 2024; Hamash & Mohamed, 2023; Han et al., 2025). These implementation questions are reinforced by recent K–12 STEM evidence identifying learner discomfort, classroom-control challenges, changing teacher roles, limited teacher technical competence, infrastructure constraints, and cost as persistent barriers to VR/AR adoption in schools (Jiang et al., 2025).

Fourthly, while immersive technologies have been studied across a range of age groups, relatively less attention has been given to the specific needs of post-primary (adolescent) learners. This group occupies a transitional stage between highly scaffolded primary education and the more autonomous learning environments typical of higher education. As a result, designing for this population requires careful calibration between guidance and independence, as learners are developing both conceptual understanding and self-regulated learning skills. This study contributes to this gap by examining how VR can support this balance in real classroom contexts.

In contrast, this work conceptualises VR as part of a sociotechnical classroom system rather than a standalone tool. Through a 25-month participatory co-design process involving students, teachers, pre-service teachers, school leaders, parents, and domain experts, and a real-world classroom deployment, we examine not only pre–post learning changes but also how procedural chemistry learning is co-designed for adolescents, enacted within classroom ecology, and shaped by the conditions required for sustainable rollout.

3. Methodology overview

3.1. Methodology approach & overview

This study adopts a real-world, deployment-oriented evaluation

approach rather than a controlled experimental design. While controlled comparisons can isolate causal effects, we aim to understand how VR can be meaningfully integrated into everyday classroom practice, which requires studying the system in authentic educational contexts. As such, the evaluation focuses on a combination of learning outcomes, user experience, and classroom integration, rather than isolating VR as a single independent variable. Pre–post measures are used to examine learning gains, while qualitative data provide insight into how these outcomes emerge through interaction, orchestration, and classroom dynamics. This approach prioritises ecological validity and deployment realism over strict causal attribution.

In this research, we partnered with seven post-primary schools in Ireland to co-design, develop, test, and implement VR in their day-to-day teaching contexts. The research takes place in Ireland, where Post-primary chemistry education faces significant laboratory-infrastructure constraints and concomitant reductions in opportunities for practical work at the post-primary level, a barrier reiterated by our stakeholders (ASTI Research & Surveys, 2025). The methodological framework draws on Participatory Design, a well-established HCI approach that emphasises co-creation in authentic contexts and sustained dialogue with intended users (DiSalvo, 2016, pp. 4459–4463), Participatory Design has been widely applied in the design of technologies for and with children and other diverse user groups (Benton & Johnson, 2014; Hendriks et al., 2013; Kim, 2024; Maun et al., 2021).

Over the period of 25 months, key stakeholders engaged in a three-phase study to design and develop a realistically deployable VR solution for chemistry education (See Fig. 1). In Phase One, we collaborated with key stakeholders (n = 23), including chemistry teachers, school management staff, subject experts, and VR practitioners, to investigate the pedagogical and practical design and deployment needs for integrating VR into post-primary schools. Insights from Phase One, coupled with a series of ideation sessions (8 sessions with 9 co-designers), shaped the design of the initial prototype, which was then evaluated by key stakeholders (n = 83), including students, chemistry teachers, subject experts, and practitioners. Finally, in Phase Three, an advanced prototype was developed, shaped by feedback from key stakeholders in Phase Two, and subsequently evaluated by students (n = 42), teachers (n = 6), and parents (n = 2).

3.2. Ethical considerations

This study received ethical approval from the Research Ethics Committee (REC) of the DCU Research Ethics Committee (REC). (Approval No. DCUREC/2025/010) in Ireland. Furthermore, written institutional permission was obtained from every participating school. Written informed consent was obtained from all adult participants prior to participation (i.e., parents/guardians, subject experts, teachers, researchers, and developers). All stakeholder groups were provided with a participant information sheet tailored to their role in the study. The procedures for the selection and participation of children were conducted in accordance with ethical guidelines and are fully described in Subsection 3.3.

3.3. Selection and participation of children

Designing with and for children requires careful ethical consideration to protect children's rights and respectfully include their perspectives. Researchers with direct contact with students were vetted by An Garda Síochána, and all institutional and school-level child protection policies were followed. Ethical approval for the study was obtained from the researchers' institutional research ethics committee prior to any engagement with schools or students. Written informed consent was obtained from parents or legal guardians, and written assent was obtained from all participating students. Age-appropriate participant information sheets explained the purpose of the study, what participation involved, the voluntary nature of participation, the right to withdraw, and how data would be used. All activities took place in school settings under standard safeguarding procedures, with teachers present or nearby as appropriate. Students were treated as valued contributors, with care taken to avoid coercion or academic pressure. Data were anonymised at collection, and information regarding confidentiality, storage, and future use of anonymised data was clearly communicated to students and parents/guardians. In this study, post-primary students participated as stakeholders in Phases Two and Three. Their involvement was intentionally timed after substantial consultation with teachers, subject experts, and VR practitioners (Phase One and early Phase Two design enquiries), ensuring that students engaged only once the prototype had reached a level of maturity appropriate for safeguarding, safety, and wellbeing in VR use. Phase Three participation (where the majority of student participation took place) was conducted during the National Science Week activities, during which all 5th- and 6th-year students in each participating school were invited to participate as part of planned enrichment activities, ensuring equitable access and no disruption to regular curricular teaching. Student participation was facilitated by schools following institutional permission and approval.

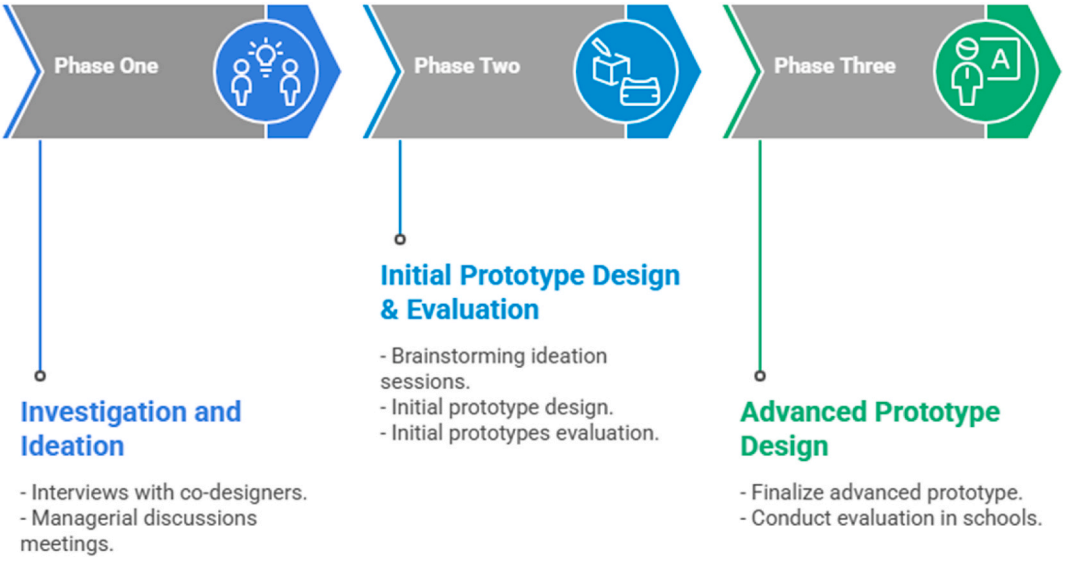


Fig. 1. The three phases of the research.

4. Phase one: investigating the practical & pedagogical requirements for VR deployment in post-primary schools

4.1. Participants

We conducted semi-structured interviews and meetings with a total of 23 participants, including post-primary chemistry teachers (n = 5, minimum of 7 years of teaching experience), principal of a post-primary school (n = 1), post-primary vice principals (n = 3), a post-primary school managerial staff member (n = 1), an expert in chemistry (n = 1), lecturers in chemistry education (n = 4), an expert in computational chemistry (n = 1), a researcher in educational technology (n = 1), a researcher in STEM education (n = 1), experts in computer science (n = 2), a researcher in HCI (n = 1), and VR developers (n = 2).

4.2. Data collection & analysis

Interviews with teachers and subject experts were designed to elicit in-depth accounts of the pedagogical and practical considerations under which VR could be meaningfully integrated into post-primary chemistry teaching, and to identify the core design features regarded as essential for an effective, educationally valuable, and engaging VR laboratory experience. Meetings with senior leaders and managerial staff from five post-primary schools captured organisational perspectives on VR adoption, focusing on school-level deployment realities: classroom layout and setup, technical infrastructure and support capacity, and institutional requirements for ethics, health, and safety. All interviews (~40–60 min) were audio-recorded and transcribed verbatim by one researcher, and the transcripts were subsequently checked by a second researcher to ensure transcriptional accuracy. Meeting discussions were documented in detailed minutes prepared by one researcher.

4.3. Phase One findings: the Practical & Pedagogical Requirements for VR deployment in post-primary schools

Taking the views of a total of 23 stakeholders, we collated the corpus of data, then clustered the requirements using affinity mapping (Fig. 2), a synthesis technique used to distil qualitative data into coherent clusters (Lucero, 2015). We also included quote excerpts to illustrate each finding from each affinity cluster.

4.3.1. Establishing pedagogical legitimacy of VR

VR as a Distinctive and Credible Pedagogical Value Proposition: Co-designers emphasised that VR should be integrated into post-primary education only when it serves a clear pedagogical purpose by addressing challenges that conventional approaches cannot readily meet; in chemistry, this value lies in visualising invisible processes, safely simulating hazardous scenarios, and extending opportunities for procedural practice beyond what physical labs can reliably provide *"If there's certain equipment, they don't have … then run it in this space [VR]" [T07, INT]*.[1] Furthermore, stakeholders stressed that pedagogical accuracy is foundational to VR's credibility in education, framing it as a prerequisite for students to treat it as a credible educational resource. Co-designers described disciplinary correctness and representational precision as *"non-negotiable" [E3, INT]*, warning that misinformation in immersive experiences risks being internalised and retained: *"students are very critical … if you spread some wrong information … they may remember it for their whole life" [E6, INT]*. Here, *"accuracy"* extends beyond content quality; it is an issue of epistemic trust and pedagogical credibility, underpinning whether VR can be accepted as a reliable mediator of scientific knowledge and practice in the classroom. Beyond these content-specific benefits, participants also constructed legitimacy in affective terms, positioning VR as a means of reducing student anxiety and building confidence ahead of real laboratory participation: *"It can be great to build confidence before they go into the real lab" [T09, INT]*.

The Positioning of VR Within the Classroom Ecology: Many participants highlighted concerns about how VR would be positioned within the broader learning journey and within institutional expectations of post-primary schooling. Many participants emphasised that VR requires an explicitly defined pedagogical role, such as pre-lab preparation, revision, formative assessment, or scenario-based training. Most described VR as *"a good first step" [E05, INT]*, followed by a clarification that *"It must not replace the real thing. Just* support *it" [T05, INT]*. Importantly, this *"supplement rather than replace"* framing was also justified through a broader view of what laboratory learning entails: managerial stakeholders cautioned that *"educational technology should not displace peer learning" [M4, MM]*, emphasising that post-primary science education also develops interpersonal competencies in ways that VR may not be able to replace. In parallel, legitimacy was tied to curricular accountability; stakeholders repeatedly stressed that VR activities must map directly onto curriculum standards and learning outcomes in order to be adoptable in practice *"We must align this [VR experience] with curriculum outcomes" [T1, INT]*; *"link[ing] simulations to learning goals is very important" [E13, INT]*. In this sense, VR becomes legitimate not only through what it can do, but through how it is made compatible with the social, curricular, and organisational logics of post-primary schooling.

4.3.2. Designing meaningful learning VR experiences for adolescent learners

Accountable Engagement Through Pedagogically Driven Design: Stakeholders framed VR as having strong motivational affordances that could bring less engaged students into chemistry learning *"This might actually get some of the less interested students engaged" [T1, INT]*. However, this motivational potential was simultaneously treated as a risk: without explicit instructional intent, *"VR could be trivialised and experienced as purposeless entertainment rather than purposeful learning" [M1, INT]*. As a result, stakeholders did not simply call for *"engaging"* VR, but for accountable engagement, experiences in which novelty is deliberately channelled into epistemically meaningful action. This requirement surfaced most clearly in debates over activity structure. While some advocated tightly sequenced simulations with progression gating that *"require correct step sequence before progression" [T12, INT]*, others preferred more exploratory interaction. Importantly, both positions point to the same underlying design problem: VR must provide sufficient structure to support procedural knowledge acquisition while retaining enough autonomy to sustain curiosity and inquiry. In this framing, the key question is not whether VR is structured or open-ended, but how its interaction design ensures that immersion produces learning rather than mere completion.

The VR-Teacher Distributed Efforts in Scaffolding and Guidance: Stakeholders described effective VR learning as an orchestrated activity, where co-designers envisaged VR as providing guidance *"from inside"* the simulation through step-by-step scaffolding and progress verification, particularly to prevent superficial interaction (students *"click through everything to get it over with" [T5, INT]*). In parallel, teachers positioned their role as external orchestration: framing the purpose of the activity *"The teacher acts as a facilitator" [T7, INT]*, prompting reflection *"the students need to be provoked to think" [E7, INT]*, monitoring engagement, and leading debriefing to connect VR actions to conceptual understanding and laboratory practice. This points to a coordinated division of labour in which the system scaffolds procedural action and attention, while teachers sustain pedagogical intent, sensemaking, and transfer.

Assessment and Feedback as Core Measures of Learning: Assessment and feedback emerged as the crucial mechanism for

[1] [Stakeholder (T = Teacher, E = Expert, S=Student, P=Parent, M = Managerial team) and Number, Source (OPNDED=Open-Ended Questionnaires, INT=Interview, MM = Meeting Minutes (School Management), OBS=Observational Notes), FGRP = Focus Group, SUR = Survey.].

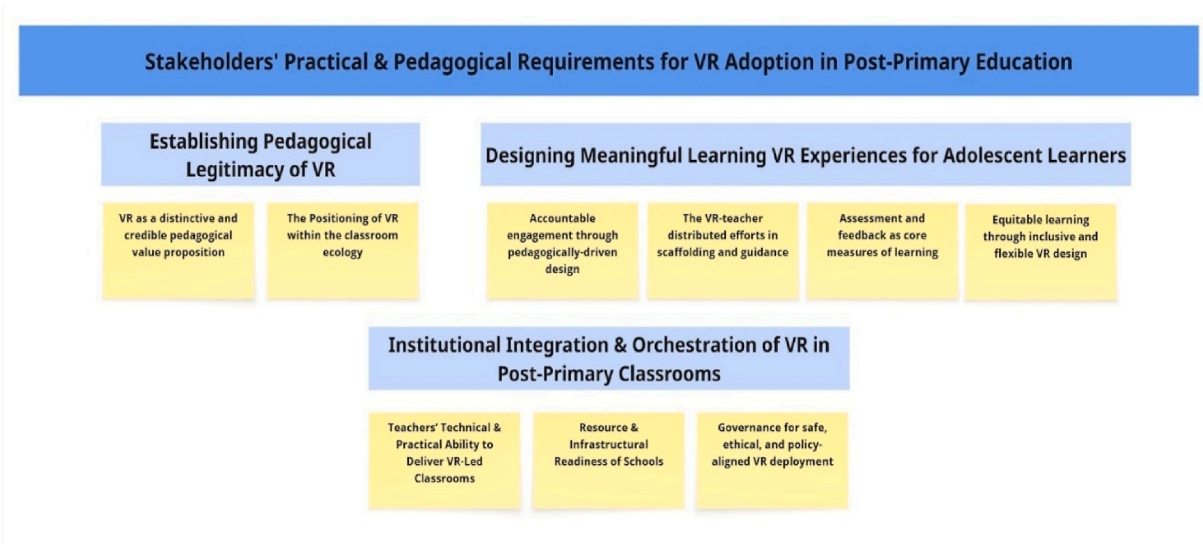


Fig. 2. Affinity map of stakeholders' practical & pedagogical requirements.

converting immersive interaction into demonstrable learning outcomes. Feedback was envisioned along a spectrum, from immediate corrective prompts, *"It should alert them if they do something wrong. Like a virtual warning" [T20, INT]* to richer multimodal supports that accommodate differences in how students attend to and process instruction *"offering information through multiple channels—visual, auditory, and textual cues …" [E10, INT]*. A leadership perspective made the stakes explicit: *"If we're investing in immersive technologies like VR, we must also invest in meaningful assessment—otherwise, we're just entertaining, not educating" [M21, MM]*. As such, there is a need to define a coordinated model in which VR provides step-by-step scaffolding from within, while teachers mediate engagement from outside, ensuring the VR delivers a meaningful learning experience.

Equitable Learning Through Inclusive and Flexible VR Design: Stakeholders emphasised that VR must be designed with inclusivity at its core, ensuring that students with SEN or those with disabilities are not excluded from learning. Subject-experts highlighted that some learners may be unable to participate fully due to visual impairments, motion sickness, or sensory sensitivities, and therefore advocated for alternatives when VR is used *"not suitable for all students … need for inclusive alternatives" [E7, INT]*. This foregrounds an important implication for adolescent classroom contexts: VR cannot be designed around a single "ideal user" pathway. In this framing, inclusivity is tightly coupled with classroom viability, without viable inclusive design, VR risks systematically excluding precisely those learners who are already marginalised by conventional lab environments.

4.3.3. Institutional integration & orchestration of VR in post-primary classrooms

Teachers' Technical & Practical Ability to Deliver VR-Led Classrooms: Across stakeholder groups, successful VR adoption was framed as an implementation challenge rather than a purely technological one *"the main limitation here is the implementation" [T19, INT]*. Many teachers and experts reported absence of actionable guidance is one primary reason why teachers feel reluctant to adopt VR in their everyday classrooms *"I have not had any kind of guidelines" [E1, INT]*. Participants described low confidence and limited prior experience in integrating VR into teaching practice, indicating that institutional uptake depends on building professional capacity through training, implementation scripts, and locally workable guidance that translates the technology into routine pedagogical practice.

Resource and Infrastructural Readiness of Schools: Stakeholders positioned feasibility as a function of both resourcing and infrastructural readiness, arguing that even well-designed VR cannot be routinised unless schools can sustainably finance, house, and support it. Cost was described as a double-edged factor: Some teachers saw VR as a cost-friendly solution *"Even in schools with well-equipped laboratories, concerns around the potential loss or breakage of materials often limit student access to hands-on experiences. Teachers may hesitate to allow students to perform experiments themselves, given the cost and fragility of the equipment or the limited availability of materials" [T02, INT]*. Other co-designers cautioned that expenses related to hardware, software, and ongoing maintenance of VR could make adoption challenging: *"It [VR] holds great promise for education, but the financial demands, covering equipment, software, training, and upkeep, can be a major barrier. Schools need cost-effective models that still deliver strong learning outcomes." [M23, MM]*. Participants also highlighted the infrastructural capacity in their schools: reliable Wi-Fi, adequate physical space, and sufficient device availability were treated as prerequisites for meaningful integration *"Not all schools have good Wi-Fi or enough VR devices" [E9, INT]*, with chemistry laboratories in particular characterised as *"too restricted for immersive use" [E18, INT]*.

Governance for Safe, Ethical, and Policy-Aligned VR Deployment: Participants stressed the need for explicit governance mechanisms that extend beyond individual teacher judgment. Concerns about physical wellbeing (e.g., nausea, dizziness) were frequently raised *"You have to keep these [side effects] in your mind" [T16, INT]*, alongside calls for practical safeguards such as shorter sessions and seated modes. Hygiene practices were also presented as essential *"disinfect between users" [T9, INT]*, yet stakeholders noted an absence of school-level hygiene guidance. These practical risks intersected with broader ethical concerns regarding safeguarding, content quality, and the lack of formal oversight, prompting calls for clearer regulation and standards *"need [for] regulations for the education [sector]" [E19, INT]*.

5. Phase two: initial prototype design & evaluation

In Phase Two, we synthesised Phase One findings into an initial prototype, co-developed with nine co-designers who shaped the prototype's design, content, and features. Once the development of the initial prototype was complete, we invited 83 stakeholders (of which 18 had participated in Phase One) to evaluate the prototype, followed by interviews to obtain in-depth insights into the pedagogical fitness of the prototype, VE content, interaction design, and other relevant areas that contribute to the success of VR deployment post-primary schools.

5.1. Co-design enquiries

A total of nine co-designers, including chemistry experts (n = 3), an HCI researcher (n = 1), a post-primary chemistry teacher (n = 1), VR developers (n = 2), and co-design researchers working with Irish schools (n = 2), participated in eight design enquiry sessions, each lasting up to an hour with 1–3 participants in each session. Following Osborn's classical brainstorming principles (Smith, 1998), sessions began with a presentation of Phase One findings, then moved into iterative discussion to define the prototype design, with sessions audio-recorded, transcribed, and supplemented by researcher notes. Conflicting stakeholder priorities were reconciled iteratively rather than through a formal decision-scoring framework. Design decisions were prioritised according to recurring evidence across stakeholder groups, curricular and pedagogical relevance, safety and accessibility requirements, and practical feasibility for school deployment. Where preferences differed, the team sought design compromises that preserved both pedagogical structure and learner autonomy; for example, tensions between tightly guided and exploratory interaction informed the use of structured scaffolding alongside optional, on-demand support and independent practice.

5.2. Initial prototype design & development

The initial prototype was developed using Unity[2] 6 and the Unity XR Interaction Toolkit. The initial prototype comprised two Virtual Environments (VEs) designed to support complementary learning outcomes for an acid–base titration, a core experiment for determining the concentration of an unknown solution. We selected titration because co-designers consistently identified it—across all teacher interviews—as essential in post-primary chemistry. The VEs were structured using the Cognitive-Affective Model of Immersive Learning (CAMIL) (Makransky & Petersen, 2021), aiming to elicit both procedural knowledge and deeper conceptual understanding by engaging cognitive and affective processes. The initial prototype consisted of two VEs: (1) the Titration Training Lab VE, which provided guided, step-by-step instruction for practising the procedural knowledge of an acid–base titration. (2) the Titration Molecule Exploration VE, which allowed students to experience the titration process from inside the flask, visualising molecular interactions to support conceptual understanding.

5.2.1. Participants

The prototype was evaluated by 83 co-designers (male, 48.2%; female, 51.8%) aged from 14 to 61 years (M = 28.43, SD = 9.12), with the following breakdown:

- Post-primary chemistry teachers (n = 4) and pre-service chemistry teachers (n = 4).
- Subject experts, researchers, and practitioners (n = 63): HCI experts (n = 18), education experts and academics (n = 24), chemistry laboratory senior technical officer (n = 1), VR developers (n = 3), VR developers and computer science practitioners (n = 10), and chemistry education experts (n = 7).
- Post-primary students (n = 10; male = 60%, female = 30%, other/prefer not to say = 10%) with ages ranging from 14 to 18 years (M = 16.7, SD = 1.25).
- Parents of post-primary students (n = 2).

5.3. Prototype evaluation procedure & data collected

Considering the unique cohort of key stakeholders, we needed to adapt the evaluation procedure as follows:

[2] https://unity.com/.

5.3.1. Procedure & data collection – teachers & subject experts

Teachers and subject expert co-designers were invited to engage with the initial prototype. Co-designers were asked to take a break after engaging with the VEs to minimise the risk of any adverse effects, such as motion sickness and disorientation (Szpak et al., 2022, pp. 533–537). Furthermore, one researcher with a background in education and experience in VR design and development recorded qualitative observational notes during co-designers' prototype evaluation sessions, focusing on participants' behaviours, interaction patterns, any technical notes related to hardware or software issues, and any notable moments of confusion or hesitation whilst interacting with the prototype. On average, each testing session generated approximately one to two pages of notes, depending on the length and complexity of the user interaction. After co-designers concluded the evaluation session, they were invited to engage in a semi-structured interview. These interviews aimed to gain rich contextual insights into VR design and the realistic deployment of VR within post-primary settings.

5.3.2. Procedure & data collection – students

Students were invited to engage with the initial prototype, and to minimise potential adverse effects of interacting with VR, students were asked to take frequent breaks between VEs (Szpak et al., 2022, pp. 533–537). Similarly to the approach we undertook with teachers and subject experts, one researcher with a background in education and experience in VR design and development recorded qualitative observational notes during co-designers' prototype evaluation sessions, aiming to analyse students' behaviours, reactions, and interactions with the prototype. On average, each testing session generated approximately one to two pages of notes, depending on the length and complexity of the user interaction. In addition, the prototype captured in-app interaction metrics as students engaged with the prototype, including the time spent in each stage and the number of times the "help" feature was used. Once students completed the VR session, they were asked to participate in a semi-structured interview to explore their perspectives on the initial prototype and its impact on their subject-related knowledge and skills. On average, students took around 50 min to engage with the prototype, while interviews lasted between 15 and 35 min.

5.3.3. Data analysis

As for interviews, we were only able to conduct interviews with 37 co-designers due to their time constraints and availability. Those who took part in interviews included chemistry teachers in post-primary education (n = 2), pre-service chemistry teachers (n = 4), post-primary students (n = 10), Parents (n = 2), researchers in chemistry education (n = 6), researchers in education (n = 9), and researchers within the HCI discipline (n = 4). All interviews were transcribed by one researcher verbatim, then corroborated by another researcher.

The corpus of data, including interview transcripts and observational notes, was analysed using reflexive thematic analysis following Braun and Clarke (2022). One researcher led the initial coding and theme development. A second researcher engaged in critical dialogue by reviewing and challenging developing interpretations, with alternative interpretations discussed reflexively rather than used to establish coding reliability or consensus. Through iterative discussion, the researchers refined the thematic structure, which was then synthesised to identify convergences, divergences, and contextual insights.

5.4. Phase Two findings: initial prototype evaluation

5.4.1. Presence as a vessel for procedural learning and meaningful engagement

Co-designers highlighted that the experience foregrounded *"doing"* through sequential task execution and hands-on manipulation, thereby approximating the felt logic of laboratory work. Several co-designers described the experience as *"like being in the lab" [T4, INT]*, with some noting that it *"didn't feel like a virtual environment" [E2, INT]*.

Furthermore, one student stated that *"It was much better than doing it in a class setting because there were fewer distractions, and I could do it myself." [S3, OPNDED]*. Observational data aligned with these perceptions, indicating that *"users were able to follow procedural workflows with minimal interruption and complete tasks efficiently" [T26, OBS]*, suggesting that interactional fluency reduced interface salience and supported sustained attention to the experiment itself.

This procedural fluency also shaped how engagement and presence were interpreted. Rather than treating immersion as an end, co-designers linked enjoyment to the immediacy of action–feedback loops and the clarity of perceptual cues *"so intuitive! I see a colourful highlight" [E10, INT]*. As learners became competent with the interaction mechanism, observations indicated a shift from mere task completion toward exploratory behaviours, such as anticipating outcomes or *"searching for ions" [T36, OBS]*, suggesting that presence operated as a mechanism for sustaining curiosity and epistemic engagement, not simply entertainment.

At the same time, stakeholders were explicit that *"lab-like"* experience does not automatically imply *"lab-equivalent" [T3, INT]* learning. Several co-designers questioned the fidelity required for transfer to real laboratory settings, citing limitations in resistance, hand-movement precision, and the authenticity of feedback. These constraints were treated as consequential, insofar as they may blunt the extent to which procedural knowledge generalises beyond VR. A parallel set of concerns emerged around embodiment: bodily engagement was widely viewed as a key affordance, *"You're just doing it" [E11, INT,]*, because it helps students overcome their limitations in imagining such an abstract concept; however, co-designers cautioned that embodied enactment could remain superficially procedural without deliberate opportunities for sensemaking.

5.4.2. Managing cognitive workload

Most co-designers characterised the prototype as cognitively manageable, with several co-designers attributing its apparent ease to the clarity and attractiveness of the visual presentation, *"the visuals were very attractive and helped me focus" [S8, INT]*. However, stakeholders' descriptions also indicate that cognitive demand was not uniform. Rather than sustained overload, some co-designers reported brief, situational spikes in workload. Students' reflections support this episodic framing of cognitive load. For example, one student described frustration when *"I answered a question and was not brought to the next step" [S3, OPNDED]*, while another noted that feedback moments were sometimes too brief, as *"the answer screen on the quiz disappeared too early" [S7, OPNDED]*. These spikes typically emerged when learners were required to re-establish procedural state *"to remember what we did in the first one" [T4, INT]* or when system transitions introduced ambiguity *"Even I was kind of like, wait!" [T2, INT]*. In other words, moments of uncertainty clustered at task boundaries where the prototype's current state and next action were less explicit, suggesting that cognitive load was episodic and closely tied to state visibility and continuity cues. Minor interactional difficulties, such as object placement, were also acknowledged, *"placing the items was annoying at times", though these were typically framed as manageable and did not overshadow the overall experience, "but overall, it was brilliant." [S9, OPNDED]*.

Co-designers' feedback further shows that these workload dynamics were entangled with the prototype's visual and sensory richness. Many co-designers valued the developers' effort to break down steps and maintain clarity, yet cautioned against excessive visual complexity that could divert attention from task-relevant elements: *"cannot put in too much … that draws attention away from the things that you really need to interact with" [E13, INT]*. Furthermore, some co-designers experienced legibility and comfort issues during prolonged use. For instance, one co-designer noted visual fatigue when reading in-headset instructions *"tired their sight" [T4, OBS]*, indicating that instructional text and interface elements require careful calibration of readability, contrast, and scale.

5.4.3. Calibrating autonomy, scaffolding, and evaluation

Co-designers repeatedly framed effective VR learning as a matter of calibration: the prototype needed to provide enough instructional structure to prevent disorientation, while still preserving meaningful learner agency. Many described the current balance as partly successful, supportive for novices, yet comparatively restrictive for more confident learners, and noted that autonomy is not only a function of "freedom" to explore, but also of whether interaction pacing and system behaviour remain legible and controllable. This concern surfaced in critiques of the prototype's temporal dynamics: some participants pointed to moments where interface changes, such as colour transitions, occurred *"way too fast" [E5, INT],* leaving co-designers feeling they *"didn't have any control" [E3, INT]*. In this framing, over-rapid feedback and transitions can function as a form of unintended over-scaffolding, where the experience effectively "moves on" for the learner and collapses opportunities to act deliberately. An "ideal" experience was articulated as *"a mix of both … some instructions, but you also give them the freedom to explore and make their own mistakes" [T14, INT]*. This led to a consistent recommendation for differentiated supports, such as adjustable scaffolding and optional labels *"beginners show labels, [while] advanced removes them" [E6, INT]*, to accommodate variation in prior knowledge and *pace*.

Within this calibration agenda, stakeholders distinguished between directive guidance and autonomy-supportive scaffolding. Students' accounts indicate that the prototype's on-demand help features supported agency and competence, *"I did it all by myself" [S10, INT]*, and teachers valued that learners *"only accessed the help if they needed it" [T1, INT]*. This sense of autonomy was strongly articulated by students, who valued being able to complete the task independently: *"I was able to do it all by myself" and "it talked me through the steps while I did it" [S1, INT]*. Others highlighted how hands-on manipulation supported conceptual understanding, particularly when *"moving the atoms myself … made me understand the chemical equation better" [S5, INT]*.

Co-designers also credited multimodal feedback, visual highlights and haptic cues for reinforcing action and maintaining orientation; vibrations helped *users "feel that they are touching elements" [E18, INT]*; indicators ensured learners *"can't feel lost" [T14, INT]*. However, they argued that feedback was unevenly implemented: progress indicators, confirmations, and instrument readouts were sometimes subtle or absent, with suggestions such as a persistent *"little pH meter in the corner" [E10, INT]* to stabilise attention and confidence across repetitive steps.

A second gap concerned sensemaking and accountability beyond procedural completion. Co-designers noted that learners could progress through the sequence without articulating observations or reasoning about outcomes, which they saw as insufficiently embedded in the interaction design. They proposed mechanisms to convert action into learning, including a clearer prediction–action–feedback loop that prompts anticipation *"what would happen next" [E17, INT]* and greater reviewability (e.g., replaying key moments, learning from errors). Finally, these dynamics extend to classroom orchestration, where the balance between autonomy and guidance also shapes how teachers coordinate learning activities during VR use: co-designers felt the prototype afforded teachers limited opportunities to actively facilitate during live use, recommending embedded formative prompts and end-of-experience checks (e.g., *"What's the endpoint of the titration?" [E2, INT]* alongside analytics on time-on-task, errors, and procedural accuracy. As one teacher summarised, such features provide *"a brilliant oversight … because in terms of assessment, what we need is data" [T7, INT]*.

These findings are particularly significant in the context of post-primary learners, who occupy a developmental stage characterised by increasing autonomy alongside continued need for instructional support. Unlike younger learners, who typically require tightly guided interaction, or university students, who are expected to self-direct their learning, the participants in this study demonstrated sensitivity to both over-scaffolding and insufficient guidance. This reinforces the need for adaptable systems that can dynamically balance structure and independence, highlighting adolescence as a distinct design context rather

than simply an intermediate stage.

5.4.4. Evaluating the Practicality of VR in everyday classrooms

Beyond individual interaction and learning design, the findings highlight broader shifts in how VR integrates into classroom practice. To better understand these changes, we interpret classroom deployment through the lens of classroom orchestration and sociotechnical systems, focusing on how VR reshapes coordination, participation, and instructional responsibility in real-world settings. Co-designers framed classroom viability as dependent on pedagogical wraparound and safe, inclusive participation conditions, features that sit largely outside the core prototype design by itself but determine whether the prototype can be practically rolled out in routine school practice. First, many co-designers noted that the current iteration offered limited support for how learners "onboard" and "offboard" the experience as a learning activity. Several recommended brief preparatory scaffolds to establish purpose and expectations before immersion, such *as "a 10–15 min snippet as a teaser" [E14, INT]*. Similarly, co-designers emphasised the absence of structured offboarding: without explicit prompts, learners could leave VR having completed steps but without articulating what they observed or why outcomes occurred. Many teachers commented on the need for a guided debrief to support learning consolidation, with one participant suggesting that teachers should *"sit with them [students] and ask them some questions" [T18, INT]*. In this framing, onboarding and offboarding are not peripheral logistics but necessary pedagogical infrastructure for connecting immersive action to classroom sense-making. From a classroom orchestration perspective, these onboarding and offboarding phases play a critical role in structuring learning across different activity modes (e.g., instruction, immersion, reflection). Rather than being peripheral, they function as coordination mechanisms that enable teachers to sequence and integrate VR within broader pedagogical flows.

Second, stakeholders stressed that VR must be deployable under school constraints of safety, wellbeing, and hygiene. While several co-designers described the prototype as a *"very safe environment" [T5, INT]*, they simultaneously emphasised the need to anticipate side effects *"You have to keep these [side effects] in your mind" [E70, SUR]* and to operationalise mitigations (i.e., short sessions, seated interaction mode). Hygiene was consistently raised as an implementation requirement, *"disinfect between users" [E16, SUR]*, particularly in shared-device classroom use. Importantly, safety was also conceptualised pedagogically: co-designers valued VR's capacity to rehearse hazardous chemistry procedures in a controlled way, approximating *"a real-life experience" while keeping experimentation safe [E9, INT]*. These considerations highlight that VR deployment operates within a broader sociotechnical context, where technical design must align with institutional constraints, safety practices, and classroom norms. The effectiveness of VR is therefore contingent not only on its instructional features, but on how well it integrates into existing educational infrastructures. Students explicitly valued VR as a safe rehearsal space, noting that *"being able to do it safely helped me understand the titration better," [S7, OPNDED]* particularly when considering hazardous procedures.

Finally, co-designers argued that classroom viability depends on ergonomic and accessibility calibration that enables equitable participation across diverse learners. Participants stressed the need to accommodate additional needs and disabilities, *"crucial to include those who have additional needs or disabilities" [E16, INT]*, and raised concerns about sensory and pacing demands, noting that learners on the autism spectrum *"might find this nauseating" [E1, INT]*. These findings position real-world classroom deployment as contingent on designing beyond the VR experience itself. From the student perspective, this potential extends beyond the immediate lesson, with learners describing the approach as *"an excellent idea" that "could seriously change how science is taught in schools" and expressing confidence that it "can become very useful for students in the future" [S5, OPNDED]*. Taken together, these findings suggest that VR should not be understood as a standalone instructional tool, but as part of a broader classroom infrastructure that redistributes how learning is coordinated and supported. From a sociotechnical perspective, the system shifts aspects of instructional responsibility, such as guidance, pacing, and feedback, across teachers, learners, and technology. This reframing moves beyond a descriptive account of deployment to highlight how immersive systems reshape classroom organisation and pedagogical practice.

6. Phase three: advanced prototype design & final evaluation

Following feedback from Phase Two, an advanced prototype was developed. 42 students, six teachers, and 2 parents in two schools in Ireland evaluated the advanced prototype.

6.1. Advanced prototype design & development

The advanced prototype was developed using Unity 6 and the Unity XR Interaction Toolkit. Aiming to meet co-designers' feedback as realistically as possible, Table 1 presents the features embedded in the advanced prototype that were revised and the new features and VEs that were added (see Figs. 3 and 4), and their area of application for each feature. At the end, the advanced prototype comprised four VEs, where students engaged with these VEs in the following sequence: the Tutorial

Table 1
Advanced prototype features.

Area of Application	Prototype Feature	Details
Support VR Experience Management	Health, Safety & Hygiene Protocol (New)	Written protocol that covers practices before, during, and after VR use tailored for real-world classroom rollout.
	Teacher-Data Dashboard (New)	Students' time spent per titration stage and frequency of the "help" button.
Measure or Add Feedback on Student Progress	Progress bar (New)	Demonstrates where learners are in the process.
	Quizzes (New)	Short interactive quizzes were added to the VEs, providing immediate feedback on students' learning.
	Student Data Dashboard (New)	Time spent on each stage and the number of times the "help" button was used.
VE-Specific Enhancements	Tutorial VE (New)	A new VE was developed to introduce users to essential VR interactions, addressing varying levels of prior VR experience.
	Titration Procedure VE (Revised)	The VE was redesigned to precisely match the Oide Chemistry Laboratory Manual for Teachers (*Oide Chemistry Laboratory Manual for Teachers (2023)*).
	Molecule Exploration VE (Revised)	Replaced passive molecule observation in this VE with interactive molecules to encourage active exploration.
	Free Titration Exploration VE (New)	Learners perform acid–base titrations independently, following the procedure learned in the Titration Training Lab VE. No guidance is provided by default; however, learners can access help via a "help" button.
Other Enhancements	Scaffolding & Guidance (Revised & New)	Added a more structured, step-by-step guidance with multimodal feedback, including atom labelling, visual cues, scaffolded prompts, audio guidance, and a progress bar.
	Spatial Awareness (New)	Incorporated spatial sound within VEs.

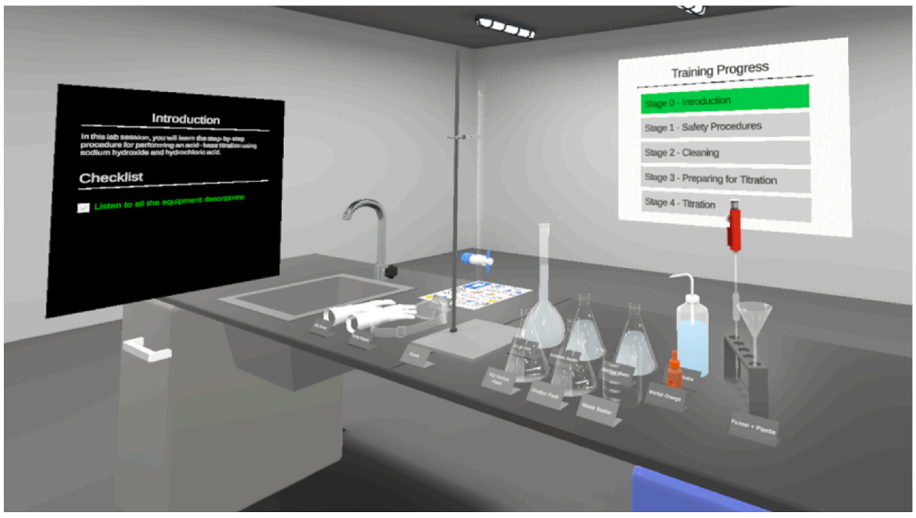


Fig. 3. Snapshot of the Free Exploration VE featuring a progress bar and Guideline Bar.

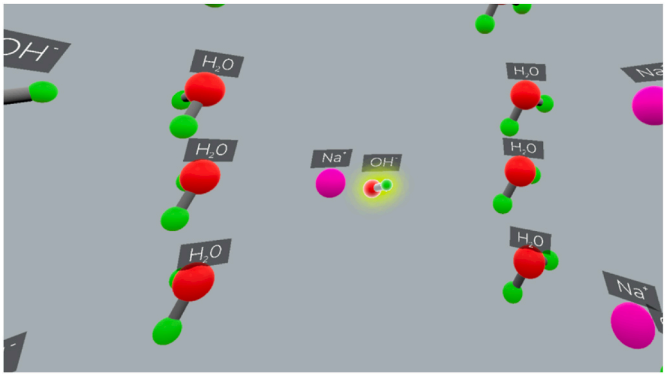


Fig. 4. Snapshot of the Molecule Exploration VE featuring a conceptual design of the molecules inside the titration flask, illustrating user interactions with chemical molecules.

VE, the Titration Training Lab VE, the Titration Molecule Exploration VE, and finally, the Free Exploration VE.

Table 1 summarises the implemented features. Many features emerged directly from specific challenges and opportunities identified during earlier phases. For example, the need for clearer procedural guidance, identified in Phase One workshops, led to the integration of step-by-step scaffolding and visual cues in the VR environment. Similarly, observations of cognitive overload at task transitions informed the introduction of more explicit state indicators and feedback mechanisms. By explicitly linking design features to stakeholder-identified needs, the prototype reflects a grounded, evidence-driven design process rather than a purely technology-driven implementation. Core design decisions in the final prototype can be directly traced to findings from earlier phases:

- **Procedural scaffolding and guidance:** Co-designers emphasised the importance of clarity and step-by-step support, leading to the integration of structured prompts, visual highlights, and guided sequencing within the VR experience.
- **Adaptive autonomy:** Feedback from both students and teachers highlighted the need to balance guidance with independence. This informed the inclusion of on-demand help features and adjustable scaffolding elements.
- **Feedback and state visibility:** Observed moments of confusion during prototype evaluation motivated the addition of clearer feedback mechanisms, including confirmations, indicators, and persistent visual cues to support orientation.
- **Classroom integration features:** Teachers' concerns regarding assessment and orchestration led to the inclusion of analytics, progress tracking, and opportunities for formative evaluation within the system.

6.1.1. *Structure of the virtual environments (VEs)*

The final prototype is organised into four interconnected Virtual Environments (VEs), each designed to support a distinct pedagogical function identified through the co-design process. These environments were not arbitrarily defined, but emerged in response to stakeholder feedback regarding onboarding, conceptual understanding, procedural guidance, and independent practice.

- **VE1 – Orientation and onboarding:** Introduces the experimental context, familiarises learners with controls, and establishes expectations for the task. This directly responds to stakeholder feedback highlighting the need for structured onboarding to reduce initial cognitive load and disorientation.
- **VE2 – Guided procedural learning:** Supports step-by-step execution of the experiment through structured prompts, visual cues, and feedback mechanisms. This design was informed by co-design findings emphasising the importance of clarity, sequencing, and scaffolding for novice learners.
- **VE3 – Molecule Exploration Environment:** Designed to support conceptual understanding through interactive molecular representations, replacing earlier passive observation approaches. This change was directly motivated by stakeholder feedback that active manipulation, rather than passive viewing, was necessary to support deeper engagement and meaning-making.
- **VE4 – Independent practice and performance tracking:** Enables learners to apply procedural knowledge with reduced guidance, supporting autonomy and reinforcing procedural knowledge and independent procedural rehearsal. In addition, this environment captures learning analytics, including time spent on each stage and frequency of "help" feature usage. These features were introduced in response to teacher feedback highlighting the need for assessment data and insight into student progress and difficulties.

Together, these environments reflect a progression from orientation and guided interaction to conceptual exploration and independent practice, aligning immersive interaction with pedagogical sequencing

and classroom needs. The structure of the four VEs directly reflects insights from earlier co-design phases. For example, the transition from passive molecular visualisation to interactive manipulation in VE3 was motivated by stakeholder feedback emphasising the importance of active engagement for conceptual understanding. Similarly, the inclusion of performance tracking features in VE4, such as time-on-task and help usage, was informed by teachers' requests for greater visibility into student progress and assessment. These examples illustrate how the final prototype embodies a traceable link between stakeholder-identified needs and concrete design decisions. To provide a clearer representation of the learning experience as encountered by students, Fig. 5 summarises the sequential flow through the four VEs. It illustrates the progression from onboarding and guided procedural practice to molecular exploration and independent titration, while also showing how feedback, scaffolding, progress support and assessment were embedded within relevant VEs rather than operating as separate stages.

6.2. Advanced prototype evaluation procedure

6.2.1. Participants

The prototype was evaluated by 42 post-primary students from two schools. The majority of students had prior VR experience (71.4%). Students ranged in age from 15 to 18 years (M = 16.81, SD = 0.92), with 64.3% identifying as male, 33.3% as female, and 2.4% selecting prefer not to say. Students' confidence in chemistry varied, with the majority reporting that they were "somewhat confident" (n = 24, 57.1%). Others described themselves as "confident" (n = 11, 26.2%), "not confident" (n = 4, 9.5%), or "very confident" (n = 3, 7.1%). The Phase Three sample also included six teachers (n = 6; male = 2, female = 4) and two parents (n = 2; male = 1, female = 1). One parent reported no prior experience with VR, while the other reported having prior experience using VR.

6.2.2. Procedure

Classroom Preparation: Prior to the evaluation, one of the researchers prepared the classroom environment to accommodate groups of 5–8 students simultaneously. All furniture that could obstruct movement or pose potential hazards was removed. The Meta Quest 2 boundary-setting feature was employed to establish a virtual perimeter, thereby alerting students when/if crossing outside the designated area. In addition, all headsets were thoroughly cleaned prior to use. Before deployment, teachers received approximately 10 min of familiarisation with the VR system and classroom procedures; additional researcher support during implementation was approximately 15 min per school.

Students' Onboarding: Upon arrival, students were welcomed and seated, and the study was introduced through a brief verbal briefing outlining the purpose of the research, the VR experience, and participation expectations. Ethical procedures were followed, including confirmation of informed assent. Students were informed of their right to withdraw at any time. Before using the VR headset, they completed a pre-evaluation questionnaire and received an orientation covering safe and appropriate headset use, interaction within the VE, and procedures to follow in case of discomfort.

Students' Engagement in VR: Students then engaged individually with the VR application at their own *pace*. The experience comprised four VEs in which students explored a virtual chemistry laboratory and observed and performed key steps of the titration process using virtual tools and guided prompts. The full VR session lasted approximately 35–50 min, including short breaks between environments to minimise fatigue, cognitive overload, and physical discomfort.

Students' Offboarding: Following completion of the VR experience, students completed a post-evaluation questionnaire and then participated in a semi-structured focus group discussion facilitated by the researcher. The discussion encouraged reflection on learning, engagement, and the perceived educational value of the VR experience, as well as its potential role in formal chemistry education.

No cases of motion sickness, visual discomfort, or other significant adverse effects were reported during Phase Three. However, three participants experienced temporary headset discomfort, primarily related to the perceived weight of the headset during longer use; these participants had previously used the application without experiencing symptoms. Risk was mitigated through the study's health and safety protocol, the design requirements identified during earlier phases, established VR-use guidance, scheduled breaks, and a locomotion design restricted to natural physical movement without artificial movement within the virtual environment.

6.2.3. Data collection & analysis

Data collection for students included pre- and post-VR surveys, observational notes, and focus group discussions. The pre-evaluation survey gathered demographic information, self-reported confidence in chemistry, and baseline measures of conceptual knowledge, procedural knowledge, and procedural self-evaluation. Procedural self-evaluation was assessed using five self-report items concerning students' perceived confidence in carrying out key stages of the titration procedure. Each item was rated on a four-point response scale, and item scores were summed to produce a total score ranging from 5 to 20, with higher scores indicating greater perceived procedural confidence. Procedural self-evaluation and procedural knowledge were treated as distinct constructs: procedural self-evaluation captured students' perceived confidence in carrying out the titration procedure, whereas procedural knowledge assessed their curriculum-aligned knowledge of the procedural steps and decisions required to perform the titration correctly.

Conceptual knowledge was assessed using five multiple-choice items related to the acid–base titration content. Responses were scored against

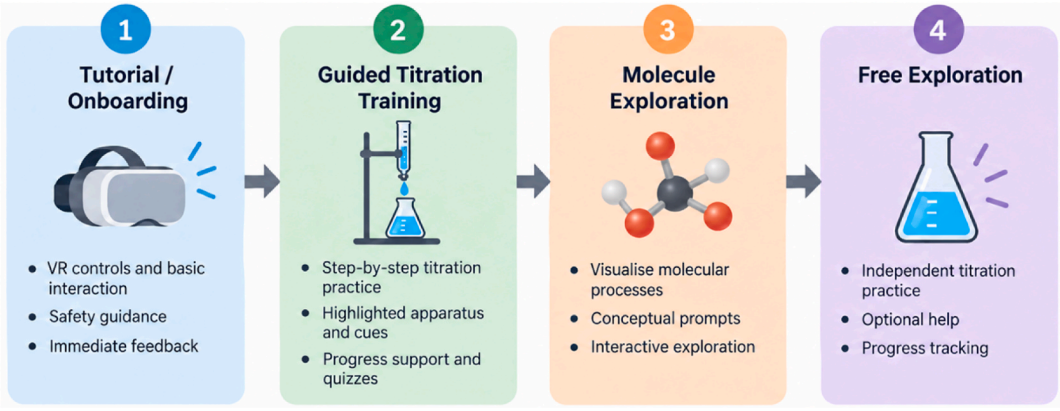


Fig. 5. Student learning flow across the four virtual environments in the advanced VR chemistry experience.

a predetermined answer key (each correct answer = 2 points), with item scores summed to produce a total conceptual knowledge score ranging from 0 to 10; higher scores indicated greater conceptual understanding. A representative conceptual item asked: "Why is it important to rinse the burette and pipette with the solutions they will contain before starting the titration?"

Procedural knowledge was assessed using five structured multiple-choice and short written-answer questions taken from Oide Chemistry Laboratory Manual for Teachers (2023) for the relevant acid–base titration procedure. These questions assessed students' knowledge of the procedural steps and decisions required to conduct the titration correctly. Responses were scored against predetermined standard answers and item-specific marking criteria and summed to produce a total procedural knowledge score ranging from 0 to 10, with higher scores indicating greater procedural knowledge. A representative short written-answer item asked: "Describe, briefly, how you would rinse the apparatus before starting a titration." All short written responses were scored by one researcher using the predetermined answer key and item-specific marking criteria; therefore, inter-rater reliability was not calculated. The assessment measured curriculum-aligned procedural knowledge rather than self-reported skill, in-system interaction logs, or direct observational ratings of physical laboratory performance. The use of questions from the national laboratory manual also supported alignment between the procedural assessment and the titration procedure represented within the VR environment. The complete conceptual and procedural knowledge assessment items and associated scoring criteria are provided in Supplementary Appendix Y. Students were not subsequently observed performing the titration in a physical laboratory; therefore, transfer from VR-based procedural knowledge to hands-on laboratory performance was not evaluated in this study.

The post-evaluation survey repeated the conceptual knowledge and procedural knowledge measures and additionally captured students' cognitive load, engagement and enjoyment, perceived learning, usability, sense of presence, acceptance, and future intention to use VR. These measures were adapted from established instruments and made developmentally appropriate for the student cohort (Makransky & Petersen, 2021). Post-use student experience was assessed using brief, developmentally adapted measures of engagement/enjoyment, perceived learning, usability, presence, acceptance/future-use intention, and cognitive load. Engagement/enjoyment comprised three four-point items assessing enjoyment, interest, and focus, which were summed to produce a score ranging from 3 to 12. Perceived learning, usability, presence, and acceptance/future-use intention were each assessed using a single four-point item, with higher scores indicating more positive perceptions. Cognitive load comprised four four-point items; the positively worded ease-of-understanding and instructional-clarity items were reverse-coded and the four items were averaged, such that higher scores indicated greater cognitive load. Internal consistency in the present sample was $\alpha = .63$ for the three-item engagement/enjoyment measure and $\alpha = .54$ for the four-item cognitive-load index. Because perceived learning, usability, presence, and acceptance/future-use intention were assessed using single items, internal-consistency coefficients were not applicable. Given the brief adapted format, particularly for cognitive load, these measures were interpreted descriptively rather than as full validated psychometric scales. Full item composition, response ranges, scoring procedures, and measurement sources for the student post-use measures are provided in Supplementary Appendix Z. The survey also included an open-ended feedback section.

Qualitative data were collected through focus groups held immediately after the VR activity. Each group included 5–8 students and lasted approximately 24–35 min, following a semi-structured guide adapted from (Krueger & Casey, 2015). The discussions addressed knowledge acquisition, procedural knowledge, engagement, enjoyment, and perceived educational value. Sessions were audio-recorded with consent and transcribed verbatim. The transcripts were analysed using reflexive thematic analysis following Braun and Clarke (2022), with an inductive orientation to coding and theme development.

The teacher post-use questionnaire included the User Experience Questionnaire (UEQ) (Schrepp et al., 2017), an adapted NASA Task Load Index (NASA-TLX) (Hart & Staveland, 1988), engagement items, health and hygiene items, and open-ended feedback. The standard 26-item UEQ assessed six dimensions—Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty—with bipolar items scored from −3 to +3 and dimension scores calculated according to standard UEQ procedures. The adapted NASA-TLX assessed six dimensions—mental demand, physical demand, time pressure, perceived success, effort, and frustration—using 1–7 response scales. Rather than calculating a weighted overall workload score, the six dimensions were analysed descriptively as separate indicators of teachers' perceived task load. Four additional 1–7 items assessed enjoyment, absorption, perceived control, and willingness to use the VR experience again, while four 1–7 items assessed perceptions of the health and hygiene procedures. Higher scores indicated more of the construct represented by each item. Given the small teacher sample (n = 6), these measures were analysed descriptively rather than used for population-level psychometric inference. Full details of the teacher post-use measures, including item composition, response scales, scoring, and sources or adaptations, are provided in Supplementary Appendix Z. Teachers then participated in focus groups exploring the educational value and practical classroom implementation of the VR experience. Discussions were audio-recorded with consent, transcribed verbatim, and analysed using reflexive thematic analysis following Braun and Clarke (2022), with an inductive orientation to coding and theme development. Parents first observed their children interacting with the advanced prototype and were subsequently invited to explore the prototype themselves to gain first-hand experience of VR. Finally, parents participated in semi-structured interviews to share their perspectives on using VR in classrooms, including its potential benefits and challenges for student learning. Interviews lasted approximately 10–15 min.

6.2.4. *Qualitative and quantitative data analysis methods*

The study employed a mixed-methods analytical approach. Descriptive statistics (means and standard deviations) were used to summarise cognitive load, engagement/enjoyment, perceived learning, usability, presence, and acceptance/future-use intention. Exploratory subgroup analyses were conducted using independent-samples t-tests to examine post-intervention differences across school, gender, and prior VR experience in chemistry confidence, cognitive load, engagement/enjoyment, perceived learning, usability, presence, and acceptance/future-use intention. Because these analyses involved 21 subgroup comparisons, Holm–Bonferroni correction was applied across the family of tests to control the family-wise error rate. Where the assumption of homogeneity of variance was violated, the unequal-variances test was used. Gender comparisons were restricted to students identifying as male or female because only one participant selected "prefer not to say." Given the modest sample and unequal subgroup sizes, these analyses were treated as exploratory and interpreted cautiously. In addition, pre–post changes in conceptual knowledge, procedural knowledge, and procedural self-evaluation were examined using paired-samples t-tests, with change scores defined as post-test minus pre-test. Mean differences, 95% confidence intervals, and Cohen's d effect sizes were calculated. The normality of paired difference scores was assessed using Shapiro–Wilk tests. As significant departures from normality were identified for the three outcomes, Wilcoxon signed-rank tests were additionally conducted as sensitivity analyses to assess the robustness of the parametric findings. All tests were two-tailed with an alpha level of 0.05. A summary of the post-use student and teacher measures and their scoring procedures is provided in Supplementary Appendix Z. Qualitatively, data from interviews, focus groups, observational notes, and open-ended survey responses were analysed using reflexive thematic analysis (Braun & Clarke, 2022).

6.3. Phase Three findings: advanced prototype findings

To move beyond a purely descriptive account of classroom changes, we interpret our findings through established perspectives from HCI and the learning sciences, particularly classroom orchestration, distributed cognition, and sociotechnical systems. These frameworks allow us to understand how the introduction of VR not only affects individual learners, but also reshapes coordination, interaction, and the distribution of cognitive and instructional responsibilities within the classroom.

6.3.1. Learning gains & transferable competencies

Following participation in the VR experience, students showed significant pre–post increases in conceptual knowledge, procedural knowledge, and procedural self-evaluation. Procedural knowledge and procedural self-evaluation are reported separately, reflecting assessed procedural understanding and students' perceived confidence in performing the procedure, respectively. Conceptual knowledge increased significantly from pre-test ($M = 7.81$, $SD = 2.49$) to post-test ($M = 9.52$, $SD = 0.97$), $t(41) = 4.79$, $p < .001$, with a mean increase of 1.71 points, 95% CI [0.99, 2.44], Cohen's $d = 0.74$. Procedural knowledge also increased significantly in scores on the curriculum-aligned procedural knowledge assessment from pre-test ($M = 4.71$, $SD = 3.32$) to post-test ($M = 9.14$, $SD = 1.07$), $t(41) = 10.40$, $p < .001$, with a mean increase of 4.43 points, 95% CI [3.57, 5.29], Cohen's $d = 1.61$. Students' procedural self-evaluation scores similarly increased from pre-test ($M = 11.62$, $SD = 4.47$) to post-test ($M = 16.83$, $SD = 2.89$), $t(41) = 11.36$, $p < .001$, with a mean increase of 5.21 points, 95% CI [4.29, 6.14], Cohen's $d = 1.75$. Shapiro–Wilk tests indicated departures from normality in the difference scores for conceptual knowledge ($W = 0.741$, $p < .001$), procedural knowledge ($W = 0.904$, $p = .002$), and procedural self-evaluation ($W = 0.933$, $p = .016$). Wilcoxon signed-rank sensitivity analyses nevertheless confirmed significant pre–post differences for all three outcomes (knowledge: $Z = 4.12$; procedural knowledge: $Z = 5.66$; procedural self-evaluation: $Z = 5.59$; all $p < .001$). Beyond the pre–post learning measures, Students also reported positive post-intervention experiences across the broader experience measures (Table 2). Engagement/enjoyment was high ($M = 10.98$, $SD = 1.20$; possible range = 3–12), while perceived learning ($M = 3.67$, $SD = 0.48$), usability ($M = 3.71$, $SD = 0.51$), presence ($M = 3.24$, $SD = 0.73$), and acceptance/future-use intention ($M = 3.83$, $SD = 0.44$) were also positively rated on their respective 1–4 scales. Cognitive load was comparatively low ($M = 1.67$, $SD = 0.36$; range = 1–4), suggesting that students generally experienced the VR activity as cognitively manageable.

Students' qualitative reflections reinforced these gains, with one describing the VR experience as *"the clearest way the topic has ever been explained to me"* [S30, OPNDED] and noting that it is *"a better use than simply looking at the textbook" [*S26, FGRP]. Teachers similarly emphasised the difficulty of teaching abstract, molecular-level concepts through traditional methods and described VR visualisation *"a lot of chemistry, same with physics, is using your imagination to try to make sense of what you're being told"[T3, INT],* while parents highlighted *"how introducing concepts through VR can allow students to focus on understanding before refining techniques in live laboratory settings"* [P2, INT].

Beyond direct measures of learning outcomes, we also observed that students themselves recognised transferable competencies they had developed throughout the experience, aligning with previous work, suggesting that experiential learning can support deeper skill internalisation beyond rote memorisation (C. Liu et al., 2025). For instance, one parent reported that their child could *"she got to really see the procedure … she intuitively knew what to do"* [P1, INT], which points to a shift from remembering verbal instructions to remembering performed actions. Such enactment was not just a matter of memory recall; several students were proud that they *"did it all by myself"* [S7, OPNDED], further suggesting that procedural memory reinforced through first-person enactment also had a positive knock-on effect on how they see themselves in relation to the subject. This sense of reward is especially important for students who struggle with the subject: *"I really enjoyed my experience as I find chemistry a hard subject" [S16, OPNDED]*. We speculate that one of the reasons why such sense of pride surfaced may be due to a healthy balance between guidance and control; teachers noticed how students' learning shifted toward independent *"problem solving, rather than relying on hints"* [T4, FGRP], which was reiterated by students—"*… and if you need help, you could get help from the system"* [S16, OPNDED].

Particularly, parents shared an interesting take as they highlighted how the exposure to VR itself is an opportunity on its own *"I think it's a great opportunity ….. I believe it's important that our kids are exposed to these kinds of technologies early in life."[P2, INT]*. We therefore explored whether students' prior VR experience was associated with how they perceived the VR experience. Students with prior VR experience reported slightly higher usability scores ($M = 3.77$, $SD = 0.43$) than those without prior VR experience ($M = 3.58$, $SD = 0.67$); however, this difference was not statistically significant, $t(14.79) = -0.88$, $p = .393$, 95% CI [−0.63, 0.26], Cohen's $d = -0.36$. Similarly, students with prior VR experience reported higher acceptance and future-use intention ($M = 3.90$, $SD = 0.31$) than those without prior VR experience ($M = 3.67$, $SD = 0.65$), but this difference was also not statistically significant, $t(12.98) = -1.19$, $p = .255$, 95% CI [−0.66, 0.19], Cohen's $d = -0.54$. These findings suggest that, within this sample, prior VR experience was not significantly associated with students' usability or acceptance ratings. Nevertheless, the parent perspective highlights the broader educational value of providing students with opportunities to engage with emerging digital technologies, particularly as digital competence is increasingly relevant to participation in technology-rich educational and professional contexts (Bejaković & Mrnjavac, 2020; Boulhrir et al., 2026; Hamash et al., 2024; Milakovich, 2012).

Exploratory subgroup analyses examined whether post-intervention responses differed according to school, gender, or prior VR experience across chemistry confidence, cognitive load, engagement/enjoyment, perceived learning, usability, presence, and acceptance/future-use intention. The smallest unadjusted p-values were observed for chemistry confidence across school groups, $t(39.05) = -2.91$, $p = .006$, Cohen's $d = -0.90$, and across gender groups, $t(39) = 2.83$, $p = .007$, Cohen's $d = 0.93$. However, neither comparison remained statistically significant after Holm–Bonferroni correction across the 21 exploratory tests (adjusted $p = .126$ and 0.140, respectively), and no other subgroup comparison was significant after adjustment. Accordingly, no robust evidence was identified that post-intervention responses differed systematically by school, gender, or prior VR experience. Given the modest overall sample and unequal subgroup sizes, these findings should be interpreted cautiously. Full exploratory subgroup results are provided in Supplementary Appendix X.

Teachers also reported positive experiences when interacting with the VR system despite limited prior familiarity with immersive technology: four of the six teachers had never previously used VR, and five reported never gaming. User experience ratings were high across all six UEQ dimensions, with mean scores of 2.97 for Attractiveness, 2.50 for Perspicuity, 2.71 for Efficiency, 2.54 for Dependability, 2.92 for Stimulation, and 2.96 for Novelty on the −3 to +3 UEQ scale. On the 1–7 task-load measures, teachers reported low time pressure ($M = 1.33$) and frustration ($M = 1.33$), modest mental demand ($M = 3.50$), physical demand ($M = 2.67$), and effort ($M = 2.83$), alongside high perceived

Table 2
Overall post-intervention student experience measures (N = 42).

Measure	Possible range	M	SD
Engagement/enjoyment	3–12	10.98	1.20
Perceived learning	1–4	3.67	0.48
Usability	1–4	3.71	0.51
Presence	1–4	3.24	0.73
Acceptance/future-use intention	1–4	3.83	0.44
Cognitive load	1–4	1.67	0.36

success (M = 6.33). Engagement ratings were near ceiling, with all six teachers giving the maximum rating (7/7) for enjoyment, absorption, perceived control, and willingness to use the simulation again. Health and hygiene procedures were also rated positively by most teachers, with five of six selecting the maximum rating across all four items; however, one teacher provided substantially lower ratings (1–3/7), indicating that perceptions were not uniformly positive. These ratings describe teachers' own experience of interacting with the VR system and should not be interpreted as measures of classroom-orchestration workload. Given the small teacher sample, these findings are intended to provide contextual implementation insights rather than population-level estimates.

6.3.2. Reworking the classroom ecology

In post-primary chemistry laboratories, practical work is typically organised around shared benches, limited apparatus, and time-bounded sessions, which often necessitate group work and turn-taking. Our participants have highlighted this: equitable participation is difficult to guarantee. Teachers described how, in group experiments, *"inevitably one person may dominate"* [T5, FGRP], while students reported that constraints of time and roles can mean they *"wouldn't get to do everything"* [S35, FGRP], resulting in only partial procedural engagement and uneven learning. In the light of this, participants positioned VR as offering a form of access that physical laboratories cannot reliably provide: one-to-one procedural rehearsal. In the VEs, students could complete the full workflow independently and *"learn the entire experiment just by yourself"* [S35, FGRP], rather than observing peers or performing only a subset of steps. Teachers emphasised that this individualised enactment ensured *that "every student gets the experience of titration individually and does not become a bystander watching another student dominate the procedure"* [T2, FGRP].

Building on this, stakeholders argued that VR not only individualised *participation*, but also individualised *pace* and support. In this, students described the VR experience as enabling self-paced progression—*"allows you to learn at your own pace and learn in your own way rather than with the whole group"* [S28, FGRP]—and as providing immediate, on-demand guidance that is difficult to guarantee in busy classrooms, where *"if you needed help … you have to wait till the teacher can come over to you"* [S26, FGRP]. Several students characterised this as effectively providing *"one person for every single individual person"* [S40, FGRP], reducing dependence on teacher availability while maintaining momentum through the procedure. Teachers likewise emphasised that the system allowed students to *"move at a good pace"* [T1, FGRP], suggesting that one-to-one pacing and support functioned as mechanisms for sustaining engagement and ensuring that all learners could complete the experiment with appropriate guidance.

Another aspect that VR reshaped the wider classroom ecology by altering how teachers viewed experiential learning. Teachers reported being surprised by how readily students could use VR to progress through the titration VEs independently *"Doing it this way allows each student to individually experience all the procedural steps and they're doing it themselves."* [T3, FGRP], which reduced the need for constant instructional intervention. As a result, several described a shift in their role from instructor to moderator, with VR providing procedural guidance while they monitored progress, managed turn-taking, and targeted support where needed. The introduction of VR reconfigured patterns of classroom orchestration, shifting the teacher's role from direct instruction toward facilitation and coordination. Rather than delivering procedural knowledge step-by-step, teachers focused on monitoring progress, managing pacing across students, and intervening selectively when difficulties arose. This redistribution of instructional responsibility reflects a transition from teacher-centred delivery toward a more decentralised orchestration model (Mõttus et al., 2025). This shift in how students learned through the VR experience prompted teachers to reconsider their teaching approaches *"even the way I teach titration, I'd be changing how I teach it,"* [T6, FGRP]. Teachers also perceived potential workload benefits beyond live orchestration. One teacher suggested that VR *"would hugely decrease teachers prep time for experiments and give more time for theory" [T4, OPNDED]*. This should be interpreted as a perceived benefit rather than as evidence of reduced teacher workload, which was not directly measured in the present study. Teacher workload during classroom orchestration was not directly quantified. Although teachers completed a task-load measure after interacting with the VR application themselves, this captured workload associated with their own use of the system rather than cognitive load, monitoring effort, or intervention frequency while supervising students. The observed redistribution of teacher roles should therefore be interpreted as a qualitative change in orchestration rather than evidence of reduced teacher workload. Future classroom deployments should directly measure teacher orchestration workload through perceived cognitive load, monitoring time, intervention frequency, and the amount of individual support required during VR and conventional laboratory activities.

From a distributed cognition perspective, procedural knowledge was no longer held solely by the teacher but was partially externalised into the VR system itself. The system guided learners through experimental steps, effectively acting as a cognitive scaffold that reduced reliance on teacher-led instruction. As a result, learning became distributed across the student, the immersive environment, and the teacher, altering how knowledge was accessed and enacted during the lesson. This is aligned with Lindgren and Johnson-Glenberg (2013) where they argue that immersive systems offload cognitive processes into the environment, meaning that knowledge is enacted through interaction with the system, not solely transmitted by the teacher.

These shifts highlight that the impact of VR extends beyond individual learning gains to the broader sociotechnical configuration of the classroom. The effectiveness of the system depended not only on its instructional design, but also on how it aligned with classroom norms, teacher practices, and institutional constraints. This reinforces the need to design immersive technologies as components of larger educational systems rather than isolated tools. This perspective is consistent with work showing that VR reshapes the sociotechnical fabric of classroom activity (Mõttus et al., 2025), that the effects of technology emerge through contextual enactment rather than inherent features (Orlikowski, 2007), and that educational technologies must be understood as systemic components embedded within wider organisational and cultural structures rather than isolated interventions (Selwyn, 2022).

6.3.3. VR as a laboratory bridge: preparing & rehearsing

Participants consistently perceived VR as a potentially valuable complementary tool for preparing students for physical laboratory work and supporting ongoing revision. Students emphasised the role of VR in pre-lab preparation, characterising it as *"great preparation" [S40, FGRP],* which helped them familiarise themselves with procedures before entering the real lab. Teachers echoed this view, highlighting the pedagogical value of students *"learn by doing before even covering it in chemistry class" [T2, OPNDED]*. At the same time, students were clear that VR was not seen as a replacement for hands-on laboratory work, but rather as a scaffold that supports it. One student noted a preference for completing the experiment physically first and then using VR *later "as a tool to revise" [S11, FGRP]*, suggesting a flexible integration of VR across different stages of the learning cycle. The capacity for repetition emerged as a key affordance, with students valuing the ability to repeat experiments independently and revisit them at home prior to class: "*You can repeat it as much as you can also, and I'll do it at home before coming" [S36, OPNDED]*. Furthermore, recent studies in immersive learning show that VR can spark students' curiosity and initial interest in science by providing highly engaging, exploratory entry points into unfamiliar subjects (Conrad et al., 2024). Consistent with these findings, participants highlighted VR's potential to stimulate student interest and curiosity by offering an engaging introduction to subjects, especially before they select their specialist subjects for the next year. Parents described

VR as a way to give students a *"taste of what a subject might look like, suggesting that virtual laboratories could allow learners to explore techniques in physics, biology, and chemistry before formal instruction" [P1, INT]*. Teachers similarly emphasised VR's value as a motivational tool, noting that when used alongside appropriate theory and laboratory work, it could be *"really used to promote the subject." [T3, FGRP]*.

7. Discussion

The results demonstrate that significant pre–post increases were observed in conceptual understanding and curriculum-aligned procedural knowledge following participation in the VR-supported experience. However, these findings should be interpreted cautiously. Because the study did not include a control or comparison condition, the observed pre–post improvements cannot be attributed solely to the VR intervention. Alternative explanations, including repeated exposure to the assessment content, increased familiarity with the testing procedure, teacher facilitation and classroom orchestration, and the novelty associated with immersive VR, cannot be ruled out. The findings therefore provide evidence of learning gains following participation in a VR-supported chemistry learning experience, rather than definitive evidence of a causal effect of VR itself. This distinction is particularly important because the intervention was intentionally designed as an integrated classroom experience in which immersive interaction was combined with procedural scaffolding, multimodal feedback, independent practice, and teacher support.

Within these boundaries, the three-phase co-design investigation showed significant short-term pre–post increases in both conceptual and procedural knowledge following participation in the VR-supported chemistry experience, alongside changes in classroom roles and opportunities for individual practice. Importantly, the study was not designed to establish whether VR was superior to conventional laboratory instruction, but to examine the suitability, feasibility, and educational value of a co-designed VR environment when enacted within authentic post-primary chemistry classrooms. A controlled comparison addresses a complementary but different question: whether learning outcomes under VR differ from those achieved through an alternative instructional condition. For example, Hu-Au and Okita (2021) used a between-subjects design to compare a VR chemistry laboratory with a real-life chemistry laboratory and found comparable outcomes for general chemistry knowledge, laboratory skills, and procedure-related safety behaviours, although differences were observed in some laboratory behaviours. In contrast, Machfudin et al. (2025) used a pretest–posttest control-group design with 78 high-school students and reported positive learning gains for students using a VR-based adaptive chemistry learning application, while the control groups showed negative learning gains. Together, these studies illustrate how controlled designs can provide evidence concerning the comparative effectiveness of VR, with findings that may range from comparable outcomes to positive effects depending on the learning context and intervention. The present study addresses a complementary question by examining whether a stakeholder co-designed VR environment can function as a meaningful educational tool within authentic post-primary chemistry classrooms. Rather than positioning VR as inherently superior to conventional instruction, the findings suggest that its value may lie in how its affordances can be integrated with existing pedagogical practices to provide opportunities for repeated, self-paced procedural rehearsal and visualisation of otherwise abstract molecular processes. Our qualitative and quantitative evidence points to two interrelated mechanisms that may help explain the observed improvements: (1) perceptual externalisation of sub-microscopic processes and (2) embodied, low-risk procedural rehearsal supported by immediate multimodal feedback. These interpretations are consistent with the Cognitive-Affective Model of Immersive Learning (Makransky & Petersen, 2021), which suggests that VR affordances such as immersion, agency, and fidelity can influence motivation, presence, and self-efficacy when appropriately aligned with instructional design.

At the same time, the absence of a comparison condition means that these mechanisms should be regarded as plausible explanations for how learning may have occurred, rather than as experimentally established mechanisms of VR effectiveness. The qualitative findings indicate that students valued the opportunity to enact procedures independently, receive immediate feedback, and visualise molecular processes, while teachers described changes in how they could facilitate and monitor learning. These observations provide contextual evidence for the role of these features, but they do not allow their individual contributions to the observed learning gains to be isolated. Future controlled and component-level studies will therefore be important for determining whether the benefits observed here arise primarily from immersion, embodied interaction, molecular visualisation, scaffolding, feedback, teacher orchestration, or the combination of these elements.

The exploratory subgroup analyses indicated no robust differences in post-intervention responses across school, gender, or prior VR experience after correction for multiple comparisons across the participating students. After correction for multiple comparisons, no significant differences in post-intervention responses were identified by school context, gender, or prior VR experience. Although the modest sample size and unequal subgroup sizes mean that these findings should be interpreted cautiously, the pattern suggests that the co-designed VR experience did not appear to advantage only particular groups of learners. This may reflect the value of the participatory co-design process, through which the environment was iteratively refined to support diverse learner needs through features such as structured scaffolding, self-paced progression, multimodal feedback, and on-demand support. Rather than demonstrating equivalent effectiveness across all groups, the absence of robust subgroup differences provides preliminary evidence that the experience was received similarly across the examined demographic groups; it does not establish accessibility equivalence.

At the same time, the findings emphasise that presence and novelty alone are insufficient to support learning. VR must be structured with clear instructional design, including sequencing, on-demand scaffolding, and reviewable evidence of learner actions (e.g., quizzes, progress indicators, dashboards) to translate engagement into understanding. This is consistent with multimedia and cognitive load principles (Mayer, 2009), which stresses the importance of limiting extraneous information, sequencing content, and coupling visualisation with explanatory cues to support comprehension and avoid episodic overload. Participants also raised practical concerns, including motion effects, hygiene in shared headset use, infrastructure constraints, and maintenance, which must be addressed for sustainable classroom deployment. Phase Three teacher feedback also identified remaining inclusive-design considerations. One teacher questioned the representativeness of the virtual hands, suggesting that the design should not default to a white male representation, and also recommended to *"lower the funnel"* to improve apparatus accessibility [T5, OPNDED]. The iterative incorporation of safety protocols, short session designs, and teacher-facing supports in the advanced prototype highlights the importance of aligning technical design with pedagogical intent and classroom orchestration to be deployable at scale.

A contribution of this work is its examination of curriculum-aligned procedural knowledge and repeated procedural rehearsal within authentic adolescent chemistry classrooms, which remains comparatively underexplored in prior VR-based chemistry education research. The substantial pre–post increase in procedural knowledge indicates that students demonstrated stronger curriculum-aligned procedural knowledge following participation in the VR experience. In parallel, the design provided opportunities for repeated, self-paced, step-by-step rehearsal of the titration procedure, an opportunity that can be difficult to provide consistently in physical laboratory settings because of time, safety, and resource constraints. This positions VR not only as a tool for conceptual visualisation, but as a medium for supporting procedural rehearsal and curriculum-aligned procedural knowledge.

Beyond these learning outcomes, our findings reveal important implications for how immersive systems should be designed for specific learner groups. A central insight is the need to recognise adolescence as a distinct design context for immersive learning technologies. Prior research shows that adolescent learners occupy a unique developmental stage characterised by emerging executive functions, uneven self-regulation, and heightened sensitivity to motivational cues, factors that shape how they engage with immersive environments (Makransky & Petersen, 2021; Parong & Mayer, 2021; Van Dinther et al., 2023; Yousfi, 2022). The autonomy–scaffolding tension we observed in post-primary learners reflects a developmental shift in immersive learning design. In early childhood and primary education, VR experiences are typically characterised by tightly structured, adult-mediated interaction with limited opportunities for autonomous exploration (Dubach et al., 2025). By contrast, post-primary learners benefit from guided autonomy, where exploratory freedom must be balanced with instructional support to prevent cognitive overload (Parong & Mayer, 2018). At the higher-education level, immersive learning environments tend to emphasise self-directed navigation and open exploration, with comparatively minimal scaffolding (Luo et al., 2021; Radianti et al., 2020). This developmental gradient helps explain why the autonomy–scaffolding tension is particularly salient in our post-primary sample.

In contrast, our findings suggest that post-primary learners require a more flexible balance between these modes. Evidence from VR-supported language and STEM learning shows that adolescents benefit from adjustable scaffolding, pacing control, and on-demand support, which help manage cognitive load while still promoting agency and exploratory behaviour (Bacca-Acosta et al., 2022; Pflieger et al., 2026; Xiuli, Juhou, Yushuai, Guangixin, & Wei, 2018). These design features align with research demonstrating that learners in this age group possess a developing, but not yet fully stable, capacity for self-regulation, and therefore rely on clear structure and timely feedback to sustain productive engagement (Makransky et al., 2019). Taken together, this positions VR not merely as a medium for content delivery, but as a technology capable of supporting developmental transitions in learning behaviour, enabling designers to scaffold autonomy in ways that are responsive to adolescents' evolving cognitive and motivational needs.

Beyond individual learning outcomes, our findings point to a broader shift in how VR functions within the classroom. Rather than operating as a standalone instructional tool, the VR system became integrated into the classroom ecology, reshaping how teaching and learning were organised. Specifically, the system enabled: One-to-one procedural rehearsal, allowing students to progress at their own *pace*; A shift in teacher roles from direct instruction toward facilitation and orchestration; Increased learner autonomy, as students could independently engage with experimental procedures. These changes suggest that the value of VR lies not only in its instructional affordances but in its ability to redistribute roles, attention, and interaction within the classroom. In this sense, VR can be understood as a sociotechnical infrastructure that supports new forms of pedagogical practice, rather than simply a digital substitute for existing laboratory activities. Fig. 6 synthesises these findings into a conceptual framework of the classroom ecology emerging from the co-designed VR environment, illustrating the reciprocal relationships among student autonomy, teacher orchestration, technology affordances, and the sociotechnical constraints that shape classroom enactment.

Beyond the specific titration application, the longitudinal co-design process generated transferable design knowledge for immersive STEM learning. Eight principles emerged across the three phases. First, immersive technology should be anchored in a clear pedagogical and curricular purpose and used where its affordances provide distinctive value, while maintaining disciplinary and representational accuracy. Second, immersive learning should be structured as a progression from onboarding and guided enactment toward conceptual exploration, increasingly independent practice, and subsequent reflection or debriefing. Third, autonomy and scaffolding should be calibrated rather than treated as competing design choices, enabling self-paced

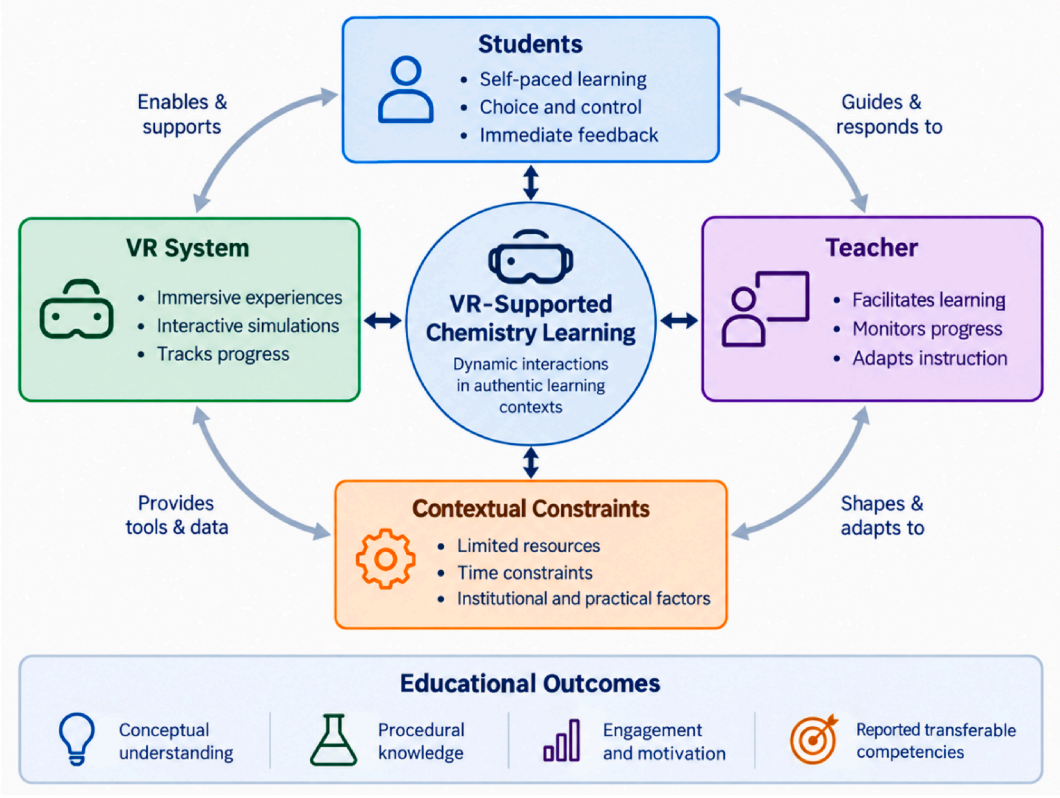


Fig. 6. Conceptual framework of the classroom ecology shaped by the co-designed VR environment.

progression while making support available on demand. Fourth, embodied interaction should be coupled with representations that help learners connect their actions to otherwise invisible scientific mechanisms, so that procedural enactment remains linked to conceptual sensemaking.

Fifth, system state, progress, consequences of actions, and learning expectations should remain visible through multimodal feedback, persistent cues, and embedded assessment, particularly at task transitions where uncertainty and cognitive demand may increase. Sixth, immersive systems should be designed explicitly for classroom orchestration as well as individual interaction, distributing guidance appropriately between the technology and teacher and providing teachers with opportunities to monitor progress, intervene, assess, and debrief. Seventh, immersive environments can provide equitable one-to-one practical rehearsal and individual pacing, but this individualisation should remain embedded within a wider social learning ecology rather than replacing peer interaction and teacher-mediated sensemaking. Eighth, safety, accessibility, ergonomics, and deployment feasibility should be treated as first-order design requirements, encompassing locomotion, physical comfort, sensory and representational inclusivity, hygiene, teacher preparation, shared-device management, infrastructure, and cost.

These principles were produced through an iterative process in which stakeholder requirements were translated into prototype features, evaluated in practice, and revised across successive phases. When stakeholder priorities conflicted, decisions were reconciled according to recurring evidence across groups, curricular and pedagogical relevance, safety and accessibility, and practical feasibility. Accordingly, these principles should be understood as empirically grounded design propositions emerging from an authentic post-primary chemistry context rather than universal prescriptions; their transferability should be examined across other STEM domains and learning activities. Together, these principles translate the empirical findings into reusable design knowledge, while their practical value ultimately depends on whether immersive systems can be implemented and sustained within the organisational realities of schools.

Achieving meaningful integration of VR in schools, however, depends on more than technical functionality or learning effectiveness. Successful deployment requires attention to a range of sociotechnical factors, including: Structured scaffolding to guide student interaction; Teacher support for orchestration and classroom management; Alignment with curriculum and assessment expectations; Practical considerations such as time constraints and classroom logistics. The extended co-design process, involving a broad stakeholder ecosystem, was critical in surfacing these requirements. This reinforces the importance of designing immersive learning technologies not only for users, but for the systems in which they are embedded. For sustainable implementation, however, qualitative identification of deployment barriers should be complemented by measurable implementation indicators. In the present study, stakeholders identified practical concerns relating to classroom setup, technical infrastructure, teacher preparation, hardware and software maintenance, and resource requirements. These factors suggest that the sustainability of VR deployment should be evaluated not only in terms of learning effectiveness but also through operational measures such as setup and preparation time, teacher training requirements, technical support and maintenance burden, device utilisation, infrastructure requirements, and cost per student or learning session. The school deployments in the present study provide some indicative implementation observations in this regard. The eight headsets used across the two participating schools required approximately 10 min of preparation before each session, while cleaning and resetting the equipment between users required approximately 5 min. Teacher familiarisation was supported through a short 10-min introductory session, and researcher support during deployment was approximately 15 min per school. One teacher also perceived a potential preparation-time benefit, suggesting that VR could *"hugely decrease teachers prep time for experiments and give more time for theory"* [T4, OPNDED]. This represents a perceived benefit rather than a directly measured reduction in teacher workload. No major technical problems requiring substantial intervention were encountered during the deployments. Implementation estimates are based on two school deployments and may differ in other infrastructure/resource contexts. The hardware investment was also relatively contained at approximately €200 per headset, with eight headsets available for the study, while the software environment was researcher-developed. For the 42 students participating in the final evaluation, this corresponds to an indicative hardware allocation of approximately €38 per student if the initial headset cost is divided across this cohort alone; however, this should not be interpreted as a lifecycle cost per student because the devices are reusable across multiple classes and cohorts. Although these figures were recorded informally rather than through a formal implementation-cost analysis, they provide an initial indication that, under the conditions of this study, routine preparation, hygiene, and teacher familiarisation did not impose substantial operational demands. Teachers' prior digital competence was not systematically assessed; therefore, its influence on implementation outcomes cannot be determined from the present study. Future deployments should examine whether teacher digital competence affects preparation time, confidence, orchestration, and support requirements.

At the scale used in this study, the eight headsets represented an approximate hardware cost of €1600 at the time of the study. This estimate excludes software development, staff time, maintenance, replacement, storage, and longer-term technical support costs. For under-resourced schools, adoption may therefore be more feasible through small shared headset sets used across classes rather than full one-device-per-student provision. However, feasibility in under-resourced settings remains dependent on local funding, charging and storage capacity, technical support, and teacher availability, and should be evaluated through longer-term implementation studies. Nevertheless, these observations should not be interpreted as evidence of long-term cost-effectiveness or scalability. Future studies should systematically quantify setup and cleaning time, teacher training and technical-support requirements, maintenance and replacement costs, device utilisation, and cost per student or learning session across longer deployment periods and larger numbers of schools.

This study did not directly compare VR with lower-cost alternatives such as desktop simulations or video-supported instruction. Such approaches may be more practical where the learning goal can be achieved without immersive interaction. The potential comparative value of VR lies in affordances such as embodied manipulation, spatial visualisation, and repeated first-person procedural rehearsal; however, these benefits should be weighed against higher hardware and implementation requirements. Future comparative studies should therefore examine whether these affordances produce sufficient added educational value to justify the additional cost and complexity. While prior research has demonstrated the potential of virtual laboratories and immersive environments to support chemistry learning, much of this work has focused on controlled settings and conceptual outcomes. Our findings extend this body of work by providing evidence from a real-world classroom deployment, with a particular emphasis on procedural knowledge development, classroom integration, and stakeholder-informed design. By linking learning outcomes with classroom ecology and deployment constraints, this study contributes a more holistic understanding of how immersive technologies can move beyond experimental prototypes toward sustainable educational infrastructure.

8. Limitations and future work

This study has several limitations that should be considered when interpreting the findings. First, the absence of a control or comparison group limits the extent to which the observed pre–post learning gains can be attributed specifically to the VR intervention. Although students

demonstrated significant improvements in conceptual knowledge and procedural knowledge following engagement with the VR experience, alternative explanations cannot be ruled out. In particular, improvements may have been influenced by repeated exposure to the assessment content, familiarity with the testing procedure, teacher facilitation and classroom orchestration, or the novelty and motivational effects associated with experiencing VR. Moreover, because the intervention was implemented as an integrated classroom experience combining immersive interaction, procedural scaffolding, feedback, independent practice, and teacher support, it is not possible to disentangle the contribution of the VR system itself from these complementary pedagogical and contextual factors. Accordingly, the findings should be interpreted as evidence of learning gains following participation in the VR-supported learning experience rather than as evidence of a causal effect of VR in isolation.

The absence of a control condition should also be understood in relation to the primary purpose of this study. The study was not designed to compare VR with conventional laboratory instruction or to establish the superiority of VR as an instructional approach. Rather, it sought to examine the suitability, feasibility, and educational value of a co-designed VR environment when enacted within authentic post-primary chemistry classrooms. The evaluation therefore prioritised ecological validity and deployment realism, allowing the co-designed environment to be examined as part of a broader classroom ecology involving students, teachers, pedagogical scaffolding, classroom routines, and institutional constraints. This design enabled us to investigate whether the VR environment could support targeted conceptual and procedural learning, meaningful student engagement, and classroom integration under realistic conditions. However, this approach necessarily limits causal attribution and does not allow the contribution of VR to be separated from the wider pedagogical and contextual conditions in which it was implemented.

Future research should therefore complement this real-world, deployment-oriented evaluation with controlled or quasi-experimental designs that include appropriate comparison conditions. Such studies could compare the co-designed VR environment with conventional laboratory instruction, alternative digital or simulation-based approaches, or other pedagogical conditions. They could also examine the contribution of individual components of the intervention, such as immersion, procedural interaction, molecular visualisation, scaffolding, feedback, and teacher facilitation. This would allow subsequent research to move beyond establishing the suitability and feasibility of co-designed VR environments towards determining their comparative effectiveness and causal contribution to chemistry learning.

Second, the evaluation focused on a single chemistry experiment, acid–base titration, and assessed learning outcomes immediately following the VR experience. This limits the extent to which the findings can be generalised to other chemistry topics, learner groups, or longer-term learning trajectories. Future studies should examine a broader range of chemistry experiments and conceptual domains and assess learning over extended periods. Longitudinal research would be particularly valuable for determining whether learning gains are retained after the initial intervention and whether procedural knowledge developed in VR transfers to physical laboratory contexts. A further limitation concerns the scope of the Phase Three implementation and the small adult samples. Teacher ($n = 6$) and parent ($n = 2$) perspectives were intended to provide contextual implementation insights rather than population-level estimates, and the Phase Three deployment occurred in only two schools. These findings should therefore be interpreted within the specific contexts of the participating schools and stakeholders.

Third, novelty effects may have influenced students' engagement, enjoyment, acceptance, and perceptions of the educational value of VR. However, 71.4% of the Phase Three students had prior VR experience, and exploratory subgroup analyses identified no robust differences by prior VR experience in engagement/enjoyment or the other post-intervention outcomes after correction for multiple comparisons. These findings reduce, but do not eliminate, the possibility of a novelty effect, as prior VR exposure does not necessarily equate to familiarity with this specific educational application. Longitudinal investigations involving repeated VR use across a school term or academic year could help distinguish sustained pedagogical value from short-term novelty and identify how students' engagement, usability perceptions, and acceptance evolve with experience.

A further limitation concerns the quantitative assessment of implementation sustainability. Although the co-design process identified important practical requirements and barriers, including teacher preparation, classroom setup, infrastructure, equipment availability, maintenance, and associated costs, these factors were not systematically measured using quantitative implementation metrics. Consequently, although the study provides indicative observations of setup, cleaning, teacher familiarisation, technical support, and hardware costs, these were not collected through a formal implementation or cost-effectiveness evaluation and should not be interpreted as generalisable estimates of long-term deployment burden. Future research should incorporate implementation measures such as time required for device preparation and classroom setup, teacher training and support hours, frequency and duration of technical interventions, hardware and software maintenance requirements, device utilisation rates, and total cost per student or learning session. Collecting these measures alongside learning and usability outcomes would provide a stronger basis for evaluating the scalability and sustainability of VR deployment across schools with different levels of infrastructure and resources.

A compelling next step is therefore to move from in-VR learning gains to transfer and durability across the full chemistry learning trajectory. Future work could conduct longitudinal, school-term studies that examine whether VR-supported preparation improves hands-on laboratory accuracy, safety compliance, and independence in live laboratory settings, and whether these gains persist over time and generalise across experiments through near- and far-transfer. Such research would provide stronger evidence regarding whether procedural knowledge developed in immersive environments translates into authentic laboratory performance. A second line of future work is to treat classroom orchestration as a primary design concern rather than an implementation afterthought. Building on the findings of this study, researchers could develop and evaluate teacher-facing orchestration features, including real-time progress and difficulty indicators, adaptive grouping, opportunities for teachers to pause and debrief the experience, and lightweight formative assessments at points where students experience increased cognitive demand. This also creates opportunities to investigate co-regulation and collaborative learning: future systems could support peer dialogue, collaborative troubleshooting, and reflective discussion while preserving the benefits of individual procedural rehearsal through paired VR/non-VR activities, comparison of student attempts, or mixed-reality views that make aspects of students' activity visible to teachers and peers. The aim would be to design VR systems that redistribute guidance and accountability productively, supporting teachers as facilitators and sensemaking leaders rather than replacing social learning with isolated headset use.

Looking ahead, Artificial Intelligence (AI) could also be explored as a means of supporting both adaptive learning and classroom orchestration. The deployment findings from our study suggest a more focused role for AI in supporting classroom orchestration. Teachers reported the challenge of monitoring individual students while simultaneously managing the wider class, while students valued the immediate, individualised guidance provided by the VR environment. These observations point to a potential role for AI in translating learner interaction data into actionable orchestration support. For learners, future systems could personalise pacing, feedback timing, and challenge levels using interaction traces to generate just-in-time prompts that preserve agency and reduce superficial task completion (Hamash, Ghreir, et al., 2025b). For teachers, AI could provide concise indicators of students requiring

additional support, summarise common difficulties for post-VR discussions, and recommend targeted follow-up activities, thereby reducing monitoring demands while supporting pedagogical responsiveness (Boulhrir et al., 2025). Although these possibilities remain prospective and were not directly evaluated in the present study, future research should investigate whether AI-enabled orchestration can successfully scale support while preserving learner autonomy and meaningful teacher-student interaction. Realising these possibilities requires careful consideration of privacy-preserving design, transparent decision-making, equity, and robust safeguards appropriate for adolescent learners to ensure intelligent support remains trustworthy and feasible within classroom contexts (Ghreir et al., 2025, pp. 292–297).

9. Conclusion

This paper presents a comprehensive investigation into the design, development, and classroom deployment of a VR-based chemistry learning system for post-primary education. Through a 25-month participatory co-design process involving a diverse stakeholder ecosystem, students showed significant pre–post increases in conceptual understanding and procedural knowledge following participation in the VR-supported chemistry experience. Given the absence of a comparison condition, these changes should not be interpreted as causal effects attributable specifically to VR. A central contribution of the study is therefore not a claim of VR effectiveness, but evidence that measurable learning changes can be observed alongside the authentic classroom deployment of a stakeholder co-designed immersive learning environment. The deployment also illustrates how immersive environments can provide repeated, self-paced opportunities for step-by-step procedural rehearsal in a safe and controlled setting. Qualitative findings suggest that perceptual externalisation of sub-microscopic processes and embodied, low-risk procedural rehearsal with immediate feedback may have contributed to students' learning experiences, although the individual causal contribution of these features was not experimentally tested.

Beyond learning outcomes, the study highlights how VR can reshape classroom practice. Rather than functioning as a standalone instructional tool, the system operated as part of a broader classroom ecology, enabling self-paced learning, redistributing instructional roles, and supporting new forms of teacher facilitation and orchestration. This positions VR as a sociotechnical component of classroom infrastructure, rather than a simple digital substitute for laboratory activities. Importantly, the findings identify adolescent learners as a distinct design context. The results demonstrate that this group requires a careful balance between autonomy and scaffolding, reflecting their transitional stage between guided and independent learning. Design features such as adjustable support, pacing control, and on-demand assistance were critical in supporting this balance, highlighting the importance of developmentally informed design in immersive learning environments. The study also underscores the importance of sociotechnical alignment for sustainable deployment. Successful integration depended not only on system design, but also on teacher practices, assessment requirements, classroom orchestration, and practical constraints such as time, safety, and accessibility. These findings reinforce the need to design immersive technologies as part of broader educational systems rather than isolated interventions. Beyond the specific chemistry application, the study contributes transferable design knowledge for immersive STEM learning centred on pedagogical legitimacy, progressive learning architecture, calibrated autonomy and scaffolding, embodied conceptualisation, state visibility and feedback, teacher orchestration, equitable individual participation within a social learning ecology, safety, accessibility, and deployment-aware design.

Several limitations should be acknowledged. The absence of a control condition limits causal interpretation of learning gains, and the focus on a single experimental context and short-term outcomes constrains generalisability. Future work should therefore (i) conduct longitudinal transfer studies that measure effects on hands-on laboratory accuracy, safety, and retention; (ii) run component-level trials to disentangle the contributions of visualisation, haptics, and scaffolding to procedural versus conceptual outcomes; and (iii) develop and evaluate scalable teacher-centred workflows and rollout models using measurable implementation indicators, including procurement and infrastructure requirements, classroom setup time, teacher training and support demands, maintenance burden, device utilisation, and cost per student or learning session across diverse school contexts. For the HCI community, the study underscores that immersive learning systems become educationally meaningful when interaction design is tightly coupled with classroom practice, stakeholder participation, and the practical realities of deployment. Overall, this work contributes a theoretically grounded and empirically supported account of how VR can move beyond experimental use toward becoming a viable and pedagogically meaningful component of post-primary education, with scalability requiring further evaluation.

CRediT authorship contribution statement

Mahmoud Hamash: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Luma Tabbaa:** Writing – review & editing, Validation, Formal analysis. **Peter Tiernan:** Writing – review & editing, Supervision, Funding acquisition. **Gareth W. Young:** Supervision, Funding acquisition, Conceptualization.

Ethics declaration

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. This study was approved by the DCU Research Ethics Committee (REC). (Approval No. DCUREC/2025/010)

Data availability

The data generated in this study are not publicly available because they involve minor participants and are subject to the conditions of the study's ethical approval and participant-consent procedures.

Ethics approval and consent to participate

This study was reviewed and approved by the Research Ethics Committee of [University name and reference number omitted for anonymisation]. All procedures were conducted in accordance with relevant institutional requirements and ethical guidelines. Written informed consent was obtained from all adult participants and from the parents or legal guardians of participating students. Written assent was additionally obtained from each student participant.

Privacy and data protection. The privacy rights of all participants were protected throughout the study. Participant data were anonymised and stored securely in accordance with institutional data-protection and research-data-management requirements, with access restricted to the research team.

Declaration of generative AI and AI-assisted technologies in the manuscript

During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language refinement, editing, and improving the clarity of responses to reviewers. After using this tool, the authors carefully reviewed and edited all content as needed and take full

responsibility for the content of the published article.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mahmoud Hamash reports financial support was provided by Research Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was conducted with the financial support of the Research Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cexr.2026.100184.

References

ASTI Research & Surveys. (2025). *School leaders: Workload, wellbeing and curriculum change* (782425/CM).

Atabek, O. (2019). Challenges in integrating technology into education. https://doi.org/10.48550/ARXIV.1904.06518.

Bacca-Acosta, J., Tejada, J., Fabregat, R., Kinshuk, & Guevara, J. (2022). Scaffolding in immersive virtual reality environments for learning English: An eye tracking study. *Educational Technology Research and Development, 70*(1), 339–362. https://doi.org/10.1007/s11423-021-10068-7

Batamuliza, J., Habinshuti, G., & Nkurunziza, J. B. (2024). Exploring lower secondary school teachers' perceptions of integrating simulations into chemistry instruction. *Frontiers in Education, 9*, Article 1484495. https://doi.org/10.3389/feduc.2024.1484495

Bejaković, P., & Mrnjavac, Ž. (2020). The importance of digital literacy on the labour market. *Employee Relations: The International Journal, 42*(4), 921–932. https://doi.org/10.1108/ER-07-2019-0274

Benton, L., & Johnson, H. (2014). Structured approaches to participatory design for children: Can targeting the needs of children with autism provide benefits for a broader child population? *Instructional Science, 42*(1), 47–65. https://doi.org/10.1007/s11251-013-9297-y

Boulhrir, T., Ghreir, H., Hamash, M., & Robert, M. (2026). Artificial intelligence in education: Mapping adaptive learning and learning analytics in K–12 Online, virtual, and distance learning. *The International Review of Research in Open and Distributed Learning, 27*(2), 122–148. https://doi.org/10.19173/irrodl.v27i2.9370

Boulhrir, T., Hamash, M., & Ghreir, H. M. A. (2025). The dual edge of large Language models: Innovation in education and emerging ethical implications. In S. Ul Rehman (Ed.), *Innovations and ethical dimensions of large Language models* (pp. 273–316). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-5017-2.ch009.

Bower, M., Howe, C., McCredie, N., Robinson, A., & Grover, D. (2014). Augmented Reality in education – Cases, places and potentials. *Educational Media International, 51*(1), 1–15. https://doi.org/10.1080/09523987.2014.889400

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.

Cao, J., B, A. S., Das, A., & Emami-Naeini, P. (2024). Understanding parents' perceptions and practices toward children's security and privacy in virtual reality. *2024 IEEE symposium on security and privacy (SP)*. https://doi.org/10.1109/SP54263.2024.00222

Cecotti, H., Huisinga, L., & Peláez, L. G. (2024). Fully immersive learning with virtual reality for assessing students in art history. *Virtual Reality, 28*(1), 33. https://doi.org/10.1007/s10055-023-00920-x

Chatain, J., Kapur, M., & Sumner, R. W. (2023). Three perspectives on embodied learning in virtual reality: Opportunities for interaction design. *Extended abstracts of the 2023 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3544549.3585805

Chen, J., Zhao, L., Li, Y., Xie, Z., Wilensky, U., & Horn, M. (2024). "Oh My God! It's Recreating Our Room!" understanding children's experiences with A room-scale augmented reality authoring toolkit. In *Proceedings of the CHI conference on human factors in computing systems* (pp. 1–17). https://doi.org/10.1145/3613904.3642043

Christopoulos, A., Mystakidis, S., Cachafeiro, E., & Laakso, M.-J. (2023). Escaping the cell: Virtual reality escape rooms in biology education. *Behaviour & Information Technology, 42*(9), 1434–1451. https://doi.org/10.1080/0144929X.2022.2079560

Christopoulos, A., Styliou, M., Ntalas, N., & Stylios, C. (2024). The impact of immersive virtual reality on knowledge acquisition and adolescent perceptions in cultural education. *Information, 15*(5), 261. https://doi.org/10.3390/info15050261

Chua, W. L., Tan, R. X. R., Liaw, S. Y., Koh, Y., Chan, G. W. H., Mikkonen, K., Neo, N. W. S., & Seah, B. (2025). Blending of virtual reality with high-fidelity simulation for interprofessional team training: A mixed methods study. *Nurse Education Today, 155*, Article 106857. https://doi.org/10.1016/j.nedt.2025.106857

Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computers & Education: X Reality, 4*, Article 100053. https://doi.org/10.1016/j.cexr.2024.100053

Damyana, G. (2024). The role of stem lab experiments in building science literacy in chemistry education. *International Journal of Multidisciplinary Research in Arts, Science and Technology, 2*(8), 42–50. https://doi.org/10.61778/ijmrast.v2i8.76

DiSalvo, B. (2016). Participatory design through a learning science lens. *Proceedings of the 2016 CHI conference on human factors in computing systems*. https://doi.org/10.1145/2858036.2858405

Docter, M. W., De Vries, T. N. D., Nguyen, H. D., & Van Keulen, H. (2024). A proof-of-concept of an integrated VR and AI application to develop classroom management competencies in teachers in training. *Education Sciences, 14*(5), 540. https://doi.org/10.3390/educsci14050540

Dubach, J., Anzeneder, S., Tempelmann, S., & Cacchione, T. (2025). Effects of school-based immersive virtual reality interventions on learning in the K-6 range: A systematic literature review. *Computers & Education: X Reality, 7*, Article 100117. https://doi.org/10.1016/j.cexr.2025.100117

Fang, X. (Ed.). (2021). *HCI in games: Serious and immersive games: Third international conference, HCI-Games 2021, held as part of the 23rd HCI international conference, HCII 2021, virtual event, July 24–29, 2021, proceedings, part II, 12790*. Springer International Publishing. https://doi.org/10.1007/978-3-030-77414-1.

Fisser, P., & Thijs, A. (2015). *Integration of 21st century skills into the curriculum of primary and secondary education*.

Ghreir, H., Hamash, M., Boulhrir, T., Megat, M. A. Z., Tiernan, P., & Donlon, E. (2025). Bridging theory and practice in AI-Driven assessments for UAE K-12 education. *2025 IEEE 14th international conference on engineering education (ICEED)*. https://doi.org/10.1109/ICEED66794.2025.11400464

Hamash, M., Ghreir, H., Tiernan, P., & Boulhrir, T. (2025a). From NPCs to AI assistants: A scoping review of AI-Driven agents in immersive STEM learning. In B. I. Edwards, H. Abuhassna, D. Olugbade, O. A. Ojo, & W. A. Jaafar Wan Yahaya (Eds.), *Advances in computational intelligence and robotics* (pp. 211–244). IGI Global. https://doi.org/10.4018/979-8-3373-0847-0.ch008.

Hamash, M., & Mohamed, H. (2023). Identifying the level of teacher's maturity in the Impactful Technology Use (ITU) in Abu Dhabi public schools. *Innovative Teaching and Learning Journal, 7*(2), 60–71. https://doi.org/10.11113/itlj.v7.135

Hamash, M., Mohamed, H., & Tiernan, P. (2024). Developing a new model for achieving flow state in STEAM education: A mixed-method investigation. *Sains Humanika, 16*(3), 101–111. https://doi.org/10.11113/sh.v16n3.2163

Hamash, M., & Tiernan, P. (2024). Trends in virtual reality research for post-primary education (2013–2024). *2024 IEEE 13th international conference on engineering education (ICEED)*. https://doi.org/10.1109/ICEED62316.2024.10923821

Hamash, M., Tiernan, P., & Young, G. (2025b). Virtual reality in post-primary education research trends from 2013 to 2024: A systematic-narrative review. *Computers & Education: X Reality, 7*, Article 100123. https://doi.org/10.1016/j.cexr.2025.100123

Hamerská, L., Matěcha, T., Tóthová, M., & Rusek, M. (2024). Between symbols and particles: Investigating the complexity of learning chemical equations. *Education Sciences, 14*(6), 570. https://doi.org/10.3390/educsci14060570

Han, X., Luo, H., Wang, Z., & Zhang, D. (2025). Using virtual reality for teacher education: A systematic review and meta-analysis of literature from 2014 to 2024. *Frontiers in Virtual Reality, 6*, Article 1620905. https://doi.org/10.3389/frvir.2025.1620905

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task load index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hendriks, N., Truyen, F., & Duval, E. (2013). Designing with Dementia: Guidelines for participatory design together with persons with dementia. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, & M. Winckler (Eds.), *Human-Computer Interaction – interact 2013, 8117* pp. 649–666). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40483-2_46.

Hu-Au, E., & Okita, S. (2021). Exploring differences in student learning and behavior between real-life and virtual reality chemistry laboratories. *Journal of Science Education and Technology, 30*(6), 862–876. https://doi.org/10.1007/s10956-021-09925-0

Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. *Proceedings of the ACM SIGCHI conference on human factors in computing systems*. https://doi.org/10.1145/258549.258715

Jegstad, K. M. (2024). Inquiry-based chemistry education: A systematic review. *Studies in Science Education, 60*(2), 251–313. https://doi.org/10.1080/03057267.2023.2248436

Jiang, H., Zhu, D., Chugh, R., Turnbull, D., & Jin, W. (2025). Virtual reality and augmented reality-supported K-12 STEM learning: Trends, advantages and challenges. *Education and Information Technologies, 30*(9), 12827–12863. https://doi.org/10.1007/s10639-024-13210-z

Jin, Q., Liu, Y., Yarosh, S., Han, B., & Qian, F. (2022). How will VR enter university classrooms? Multi-stakeholders investigation of VR in higher education. In *CHI conference on human factors in computing systems* (pp. 1–17). https://doi.org/10.1145/3491102.3517542

Jin, Q., Liu, Y., Yuan, Y., Han, B., Qian, F., & Yarosh, S. (2024). Virtual reality, real pedagogy: A contextual inquiry of instructor practices with VR video. In *Proceedings of the CHI conference on human factors in computing systems* (pp. 1–21). https://doi.org/10.1145/3613904.3642510

Kablo, E., Kleber, M., & Arias Cabarcos, P. *Artifacts for "PrivaCI in VR: Exploring Perceptions and Acceptability of Data Sharing in Virtual Reality Through Contextual Integrity"*. (2025). Zenodo https://doi.org/10.5281/ZENODO.15608186.

Kaimara, P., Oikonomou, A., & Deliyannis, I. (2022). Could virtual reality applications pose real risks to children and adolescents? A systematic review of ethical issues and concerns. *Virtual Reality, 26*(2), 697–735. https://doi.org/10.1007/s10055-021-00563-w

Kim, H. N. (2024). Guidance for best practices in participatory design involving people with visual impairment. *Ergonomics in Design: The Quarterly of Human Factors Applications, 32*(2), 48–54. https://doi.org/10.1177/10648046211058248

Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research* (5th ed.). SAGE.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 445–452. https://doi.org/10.3102/0013189X13511661

Liu, C., Meng, S., Zheng, W., & Zhou, Z. (2025). Research on the impact of immersive virtual reality classroom on student experience and concentration. *Virtual Reality, 29*(2), 82. https://doi.org/10.1007/s10055-025-01153-w

Liu, J., Liu, Z., Wang, C., Xu, Y., Chen, J., & Cheng, Y. (2024). K-12 students' higher-order thinking skills: Conceptualization, components, and evaluation indicators. *Thinking Skills and Creativity, 52*, Article 101551. https://doi.org/10.1016/j.tsc.2024.101551

Lucero, A. (2015). Using affinity diagrams to evaluate interactive prototypes. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction – interact 2015, 9297* pp. 231–248). Springer International Publishing. https://doi.org/10.1007/978-3-319-22668-2_19.

Luo, H., Li, G., Feng, Q., Yang, Y., & Zuo, M. (2021). Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning, 37*(3), 887–901. https://doi.org/10.1111/jcal.12538

Machfudin, M. F., Setiawan, E. I., Halim, K. J., Santoso, J., Putra Utama, A. B., Gunawan, G., Wardana Kusum, S. B., Singh, V., & Tuan Vu, T. N. (2025). Evaluating student learning outcomes in virtual reality adaptive chemistry. *Jurnal Ilmu Pendidikan, 31*(2), 388–398. https://doi.org/10.17977/um048v31i22025p388-398

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Maun, R., Fabri, M., & Trevorrow, P. (2021). Adapting participatory design activities for autistic adults: A review. In M. M. Soares, E. Rosenzweig, & A. Marcus (Eds.), *Design, user experience, and usability: Design for diversity, well-being, and social development, 12780* pp. 300–314). Springer International Publishing. https://doi.org/10.1007/978-3-030-78224-5_21.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Milakovich, M. E. (2012). *Digital governance*, 0 ed. Routledge. https://doi.org/10.4324/9780203815991

Mills, K. A., & Brown, A. (2022). Immersive virtual reality (VR) for digital media making: Transmediation is key. *Learning, Media and Technology, 47*(2), 179–200. https://doi.org/10.1080/17439884.2021.1952428

Moon, H., Bautista Isaza, C. A., Gallagher, M., McDaniel, C., Vernier, A., Ican, L., Springer, K., Cohn, M., Bennett, S., Nair, P., Ricard, A., Pochiraju, N., Enriquez, D., Lee, S. W., Ogle, J. T., Newbill, P., & Jeon, M. (2025). "Look at my planet!": How handheld virtual reality shapes informal learning experiences. *Proceedings of the extended abstracts of the CHI conference on human factors in computing systems*. https://doi.org/10.1145/3706599.3720020

Mõttus, P., Bardone, E., & Opermann, S. (2025). Opening the Black box: The role of tinkering with virtual reality in materialising the future classroom. *Postdigital Science and Education, 7*(3), 788–812. https://doi.org/10.1007/s42438-025-00568-7

Oide chemistry laboratory manual for teachers.(2023) (Oide).

Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organization Studies, 28*(9), 1435–1448. https://doi.org/10.1177/0170840607081138

Orosz, G., Németh, V., Kovács, L., Somogyi, Z., & Korom, E. (2023). Guided inquiry-based learning in secondary-school chemistry classes: A case study. *Chemistry Education: Research and Practice, 24*(1), 50–70. https://doi.org/10.1039/D2RP00110A

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785–797. https://doi.org/10.1037/edu0000241

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37*(1), 226–241. https://doi.org/10.1111/jcal.12482

Pflieger, L. C. J., Baumann, T., Hartmann, M., Koch, L., Kern, F., Kullmann, P., Grafe, S., Latoschik, M. E., & Bannert, M. (2026). Exploring scaffolding mechanisms for learning regulation in immersive collaborative VR environments. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-026-09959-6

Piaget, J., & Inhelder, B. (2013). *The growth of logical thinking from childhood to adolescence: AN ESSAY on the construction of formal operational structures*. Taylor and Francis.

Qorbani, S., Dalili, S., Arya, A., & Joslin, C. (2024). Assessing learning in an immersive virtual reality: A curriculum-based experiment in chemistry education. *Education Sciences, 14*(5), 476. https://doi.org/10.3390/educsci14050476

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778. https://doi.org/10.1016/j.compedu.2019.103778

Rahman, H., Wahid, S. A., Ahmad, F., & Ali, N. (2024). Game-based learning in metaverse: Virtual chemistry classroom for chemical bonding for remote education. *Education and Information Technologies, 29*(15), 19595–19619. https://doi.org/10.1007/s10639-024-12575-5

Sabrie, N., Khan, R., Plahouras, J., Johnston, B. C., Scaffidi, M. A., Grover, S. C., & Walsh, C. M. (2025). Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. *Cochrane Database of Systematic Reviews, 2025*(9). https://doi.org/10.1002/14651858.CD008237.pub4

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(4), 40. https://doi.org/10.9781/ijimai.2017.445

Selwyn, N. (2022). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury Academic.

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays, 29*(2), 58–69. https://doi.org/10.1016/j.displa.2007.09.005

Smith, G. F. (1998). Idea-Generation techniques: A formulary of active ingredients. *Journal of Creative Behavior, 32*(2), 107–134. https://doi.org/10.1002/j.2162-6057.1998.tb00810.x

Southgate, E. (2022). Teachers facilitating student virtual reality content creation: Conceptual, curriculum, and pedagogical insights. In P. MacDowell, & J. Lock (Eds.), *Immersive education* (pp. 189–204). Springer International Publishing. https://doi.org/10.1007/978-3-031-18138-2_12.

Stracke, C. M., Bothe, P., Adler, S., Heller, E. S., Deuchler, J., Pomino, J., & Wölfel, M. (2025). Immersive virtual reality in higher education: A systematic review of the scientific literature. *Virtual Reality, 29*(2), 64. https://doi.org/10.1007/s10055-025-01136-x

Szpak, A., Richards, A., Michalski, S. C., & Loetscher, T. (2022). Getting the Most out of virtual reality: Evaluating short breaks to reduce cybersickness and cognitive aftereffects. *2022 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*. https://doi.org/10.1109/VRW55335.2022.00119

Üce, M., & Ceyhan, İ. (2019). Misconception in chemistry education and practices to eliminate them: Literature analysis. *Journal of Education and Training Studies, 7*(3), 202. https://doi.org/10.11114/jets.v7i3.3990

Van Dinther, R., De Putter, L., & Pepin, B. (2023). Features of immersive virtual reality to support meaningful chemistry education. *Journal of Chemical Education, 100*(4), 1537–1546. https://doi.org/10.1021/acs.jchemed.2c01069

Walsh, A., O'Brien, R., McGuire, K., & Power, D. (2025). Exploring virtual reality as a tool for enhancing teaching and learning anatomy to medical students: A feasibility and acceptability study. *The Clinical Teacher, 22*(5), Article e70191. https://doi.org/10.1111/tct.70191

Wiedbusch, M., Brosnihan, A., Delgado, T., Dever, D., Marano, C., Patel, M., & Azevedo, R. (2025). From game-based simulation to practice: The challenges of capturing, modeling, and transferring multimodal data for chemistry skill mastery. In J. L. Plass, & X. Ochoa (Eds.), *Serious games, 15259* pp. 447–454). Nature Switzerland: Springer. https://doi.org/10.1007/978-3-031-74138-8_37.

Wiedbusch, M. D., Dever, D. A., Park, S., Romero, S., Jannotti, R., Garcia-Gali, D., & Azevedo, R. (2026). Performance and learning differences inside and outside immersive virtual learning (IVR): Investigating the CAMIL model for chemistry lab skill development in high school students. *Learning and Instruction, 104*, Article 102344. https://doi.org/10.1016/j.learninstruc.2026.102344

Woodward, J., Alemu, F., E. López Adames, N., Anthony, L., Yip, C., J, & Ruiz, J. (2022). "It would be cool to get stampeded by dinosaurs": Analyzing children's conceptual model of AR headsets through Co-Design. In *CHI conference on human factors in computing systems* (pp. 1–13). https://doi.org/10.1145/3491102.3501979

Woźniak, M. P., Lewczuk, A., Adamkiewicz, K., Józiewicz, J., Malaya, M., & Ladonski, P. (2020). ARchemist: Aiding experimental chemistry education using augmented reality technology. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1–6). https://doi.org/10.1145/3334480.3381441

Wu, Y., You, S., Guo, Z., Li, X., Zhou, G., & Gong, J. (2023). MR.Brick: Designing A remote mixed-reality educational game System for promoting children's social & collaborative skills. *Proceedings of the 2023 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3544548.3581041

Xiuli, H., Juhou, H., Yushuai, L., Guangixin, H., & Wei, W. (2018). The effect of learning styles and scaffolding strategy on students' achievement in a VR learning environment. In *iLRN 2018 Montana*. Verlag der Technischen Universität Graz. https://doi.org/10.3217/978-3-85125-609-3-34.

Yang, C., Zhang, J., Hu, Y., Yang, X., Chen, M., Shan, M., & Li, L. (2024). The impact of virtual reality on practical skills for students in science and engineering education: A meta-analysis. *International Journal of STEM Education, 11*(1), 28. https://doi.org/10.1186/s40594-024-00487-2

Ye, R., & Fan, M. (2025). Exploring the design of a multimodal head-mounted virtual reality for earthquake escape education for primary school students. *Proceedings of the extended abstracts of the CHI conference on human factors in computing systems*. https://doi.org/10.1145/3706599.3719919

Yousfi, J. J. (2022). *Immersive learning through virtual reality: A literature review on the cognitive and motivational impact of VR in education for students aged 10 to 16*. https://doi.org/10.5281/ZENODO.16909946. Zenodo Version 1.0. .