International Union of Scientific and Technological Scholars

Psychology of Education and Learning Sciences
https://journals.cypedia.net/pels

Article

The Impact of Immersive Learning Technologies on Adolescents' Scientific Reasoning: A Mixed-Methods Study Integrating Cognitive Science and Educational Psychology

Rajiv Sharma*
Interdisciplinary Division of Educational Technology, National Institute of Education, Nanyang Technological University, Singapore

Received: 2 July 2025; Revised: 13 July 2025; Accepted: 20 July 2025; Published: 28 July 2025

ABSTRACT
This study investigates how immersive learning technologies (ILTs)—including virtual reality (VR) and augmented reality (AR)—influence adolescents' scientific reasoning skills, by integrating theoretical frameworks from cognitive science and educational psychology. A mixed-methods design was employed, with 320 adolescents (ages 13–16) from 12 middle schools in the southwestern United States randomly assigned to either an ILT-integrated science curriculum group or a traditional textbook-based curriculum group. Quantitative data were collected via pre- and post-tests measuring scientific reasoning (e.g., hypothesis formulation, data analysis, causal inference), while qualitative data included semi-structured interviews and classroom observation notes. Results revealed that the ILT group demonstrated a statistically significant improvement in overall scientific reasoning scores ($M = 76.2$, $SD = 8.9$) compared to the traditional group ($M = 64.5$, $SD = 10.3$; $t(318) = 9.87$, $p < .001$). Cognitive load theory analysis indicated that ILTs reduced extraneous cognitive load by 32% ($p < .01$) by aligning with adolescents' working memory capacities. Qualitative findings further highlighted that ILTs enhanced situational interest and metacognitive awareness, key mediators of learning identified in educational psychology. These findings contribute to the interdisciplinary understanding of how technology can scaffold complex cognitive skills, providing practical implications for science educators and learning technology designers.

Keywords: Immersive Learning Technologies; Scientific Reasoning; Adolescent Cognition; Cognitive Load Theory; Educational Psychology; Learning Sciences

1. Introduction

1.1 Background

Scientific reasoning—the ability to formulate hypotheses, analyze empirical data, and draw evidence-based conclusions—is a foundational skill for adolescents' academic success and lifelong engagement with science (Zimmerman, 2007). However, traditional science instruction often relies on passive textbook reading and teacher-centered lectures, which frequently fail to engage adolescents' developing cognitive systems and limit opportunities to practice complex reasoning (National Research Council [NRC], 2012). In recent years, immersive learning technologies (ILTs), such as VR and AR, have emerged as promising tools to address this gap. By creating interactive, context-rich environments that simulate real-world scientific phenomena (e.g., cellular processes, ecological systems), ILTs have the potential to align with core principles of cognitive science—such as embodied cognition and situated learning—and educational psychology frameworks like cognitive load theory (CLT) and self-determination theory (SDT; Dunleavy & Dede, 2014; Mayer, 2020).

1.2 Theoretical Framework

This study integrates three interdisciplinary theoretical perspectives to guide the investigation of ILTs and scientific reasoning:

1.2.1 Cognitive Load Theory (CLT)

CLT, developed by Sweller (1988), posits that learning is optimized when instructional design aligns with the limitations of human working memory, which can process approximately 5–9 information chunks at a time (Miller, 1956). ILTs may reduce extraneous cognitive load (i.e., unnecessary mental effort spent on irrelevant stimuli) by presenting information in visual, interactive formats that leverage dual-coding theory—simultaneously engaging verbal and visual working memory channels (Paivio, 1971). For adolescents, whose prefrontal cortex (responsible for working memory and executive function) is still developing (Steinberg, 2014), ILTs could scaffold reasoning by reducing cognitive overload and focusing attention on core scientific concepts.

1.2.2 Situated Learning Theory

Lave and Wenger's (1991) situated learning theory argues that knowledge is constructed through participation in authentic, context-rich activities. Traditional science instruction often decouples abstract concepts from real-world applications, whereas ILTs immerse learners in simulated scientific contexts (e.g., conducting virtual experiments, exploring 3D models of ecosystems). This alignment with situated learning may enhance adolescents' ability to transfer scientific reasoning skills to novel problems, a key challenge in science education (Bransford, Brown, & Cocking, 2000).

1.2.3 Self-Determination Theory (SDT)

SDT (Ryan & Deci, 2000) identifies autonomy, competence, and relatedness as basic psychological needs that drive intrinsic motivation. ILTs provide opportunities for adolescents to explore scientific phenomena at their own pace (autonomy), receive immediate feedback on their reasoning (competence), and collaborate with peers in virtual environments (relatedness). Enhanced intrinsic motivation, in turn, may increase engagement with scientific reasoning tasks, a critical factor given adolescents' declining interest in science during middle school (Osborne, Simon, & Collins, 2003).

1.3 Research Gaps and Objectives

Despite growing interest in ILTs, three key gaps remain in the literature: (1) Most studies focus on short-term knowledge acquisition (e.g., memorization of facts) rather than complex cognitive skills like scientific reasoning (Huang et al., 2020); (2) Few studies integrate cognitive science and educational psychology to explain *why* ILTs may influence reasoning, limiting theoretical generalizability; (3) Mixed-methods designs that combine quantitative measures of reasoning with qualitative insights into learning processes are rare, leading to incomplete understanding of ILT effectiveness.

To address these gaps, this study aims to:

Compare the impact of ILT-integrated versus traditional science curricula on adolescents' scientific reasoning skills;

Examine how ILTs influence cognitive load and intrinsic motivation, using CLT and SDT as explanatory frameworks;

Explore adolescents' and teachers' perceptions of ILTs as tools for scaffolding scientific reasoning.

2. Methodology

2.1 Participants

A total of 320 adolescents (ages 13–16, M = 14.2, SD = 0.9) participated in this study, recruited from 12 public middle schools in Arizona and California, United States. Schools were selected to represent diverse socioeconomic backgrounds (42% of participants eligible for free/reduced-price lunch) and ethnicities (45% Hispanic/Latino, 30% White, 15% Asian American, 10% Black/African American). Participants were enrolled in 8th-grade life science courses, as this grade level focuses on complex biological concepts (e.g., evolution, ecology) that require sophisticated scientific reasoning (NRC, 2012).

Participants were randomly assigned to either the experimental group (ILT-integrated curriculum, n = 160) or the control group (traditional textbook-based curriculum, n = 160). Randomization was conducted at the classroom level to avoid within-classroom contamination, with 6 classrooms assigned to each group. Teachers in both groups had at least 5 years of teaching experience (M = 7.3, SD = 2.1) and received 8 hours of training on the respective curriculum prior to the study.

2.2 Materials

2.2.1 Immersive Learning Technology (ILT) Curriculum

The experimental group used a 10-week ILT-integrated life science curriculum developed in collaboration with learning technology designers at Arizona State University. The curriculum included three core VR/AR modules:

Cellular Processes VR: A fully immersive VR module where students explore cell organelles, simulate cellular respiration, and test hypotheses about how environmental factors (e.g., temperature, oxygen levels) affect cell function.

Ecosystem AR: An AR module that overlays digital models of ecological food webs onto real-world classroom objects (e.g., plants, rocks), allowing students to manipulate variables (e.g., removing a predator species) and observe resulting changes.

Evolution Simulation: A hybrid VR/AR module where students "travel back in time" to observe fossil records, compare anatomical features of species, and construct evidence-based explanations for evolutionary relationships.

All modules included embedded scaffolds: (1) Real-time feedback on hypothesis formulation (e.g., "Your hypothesis includes a clear independent variable—great job!"); (2) Metacognitive prompts (e.g., "What data do you need to support your conclusion?"); (3) Collaborative tools (e.g., virtual whiteboards for group data analysis).

2.2.2 Traditional Curriculum

The control group used the same 10-week life science curriculum (aligned with Next Generation Science Standards) but delivered via traditional methods: textbook readings (Pearson Life Science, 2020), teacher lectures, and paper-based worksheets. No digital tools beyond basic PowerPoint presentations were used, and activities were structured to match the experimental group's content sequence (e.g., cellular processes taught in Week 2, ecosystems in Week 5).

2.2.3 Measurement Tools

Scientific Reasoning Test (SRT): A 30-item multiple-choice and open-response test adapted from the Lawson Classroom Test of Scientific Reasoning (LCTSR; Lawson, 2000) and validated for middle school students (Cronbach's α = .87). The SRT measures five subskills: hypothesis generation (α = .82), data interpretation (α = .85), causal inference (α = .83), control of variables (α = .81), and argument construction (α = .84). Pre-tests were administered 1 week before curriculum implementation, and post-tests 1 week after completion.

Cognitive Load Assessment (CLA): A 12-item Likert-scale questionnaire (1 = "Strongly Disagree" to 7 = "Strongly Agree") adapted from Paas, Tuovinen, Tabbers, and Van Gerven (2003) to measure extraneous (e.g., "The instruction included unnecessary information"), intrinsic (e.g., "The scientific concepts were complex"), and germane (e.g., "The activities helped me understand how to reason scientifically") cognitive load (Cronbach's α = .89).

Intrinsic Motivation Scale (IMS): A 15-item Likert-scale questionnaire (1 = "Never" to 5 = "Always") based on SDT (Ryan & Deci, 2000) measuring autonomy (α = .86), competence (α = .88), relatedness (α = .85), and situational interest (α = .87).

Semi-Structured Interviews: 40 participants (20 from each group) and 12 teachers were interviewed post-study. Interview questions focused on perceptions of curriculum effectiveness (e.g., "How did the curriculum help you practice scientific reasoning?") and challenges (e.g., "What was difficult about using the VR/AR tools?"). Interviews lasted 20–30 minutes, were audio-recorded, and transcribed verbatim.

Classroom Observation Notes: Researchers conducted 24 classroom observations (2 per classroom) using a structured protocol to document student engagement (e.g., time spent on reasoning tasks) and teacher scaffolding (e.g., number of metacognitive prompts).

2.3 Procedure

The study was approved by the Institutional Review Board (IRB) of Arizona State University (Protocol #2023-0456). Parental consent and student assent were obtained for all participants.

Pre-Test Phase (Week 1): All participants completed the SRT, CLA (baseline), and IMS (baseline) during regular class time. Researchers also conducted pre-study interviews with teachers to document existing instructional practices.

Curriculum Implementation (Weeks 2–11): Both groups completed the 10-week life science curriculum. The experimental group used ILTs for 2–3 class periods per week (45 minutes per period), while the control group used traditional materials for the same duration. Researchers conducted classroom observations during Weeks 4 and 8 to document implementation fidelity.

Post-Test Phase (Week 12): All participants completed the post-test SRT, post-test CLA, and post-test IMS. Semi-structured interviews with participants and teachers were conducted during Weeks 12–13.

2.4 Data Analysis

Quantitative data were analyzed using SPSS 28.0. Independent samples t-tests compared pre- and post-test SRT scores between groups, while repeated-measures ANOVAs examined changes in cognitive load and intrinsic motivation over time. Effect sizes (Cohen's d) were calculated to determine the practical significance of group differences.

Qualitative data (interviews, observation notes) were analyzed using thematic analysis (Braun & Clarke, 2006). Two researchers independently coded the data using a deductive framework based on the study's theoretical models (CLT, SDT, situated learning) and inductive codes emerging from the data (e.g., "VR/AR usability challenges"). Inter-coder reliability was assessed using Cohen's κ, with a threshold of $\kappa > .80$ considered acceptable ($\kappa = .86$ for participant interviews, $\kappa = .88$ for teacher interviews). Discrepancies were resolved through discussion.

3. Results

3.1 Quantitative Results

3.1.1 Scientific Reasoning (SRT)

Pre-test SRT scores showed no significant difference between the experimental group ($M = 58.3$, $SD = 9.2$) and the control group ($M = 57.8$, $SD = 8.9$; $t(318) = 0.42$, $p = .675$), indicating groups were equivalent at baseline.

Post-test results revealed a significant main effect of group ($F(1, 318) = 97.43$, $p < .001$, $\eta^2 = .23$). The experimental group achieved significantly higher post-test SRT scores ($M = 76.2$, $SD = 8.9$) than the control group ($M = 64.5$, $SD = 10.3$; Cohen's $d = 1.21$, indicating a large effect size).

Subskill analysis showed the experimental group outperformed the control group across all five scientific reasoning subskills (all $p < .001$):

- Hypothesis generation: Experimental ($M = 78.5$, $SD = 9.1$) vs. Control ($M = 65.2$, $SD = 10.4$; $d = 1.38$)
- Data interpretation: Experimental ($M = 77.3$, $SD = 8.7$) vs. Control ($M = 63.8$, $SD = 9.8$; $d = 1.42$)
- Causal inference: Experimental ($M = 75.9$, $SD = 9.3$) vs. Control ($M = 64.9$, $SD = 10.1$; $d = 1.15$)
- Control of variables: Experimental ($M = 74.8$, $SD = 8.5$) vs. Control ($M = 62.7$, $SD = 9.6$; $d = 1.31$)
- Argument construction: Experimental ($M = 76.7$, $SD = 8.8$) vs. Control ($M = 65.5$, $SD = 10.2$; $d = 1.19$)

3.1.2 Cognitive Load (CLA)

Repeated-measures ANOVA showed a significant group × time interaction for extraneous cognitive load ($F(1, 318) = 45.67$, $p < .001$, $\eta^2 = .13$). The experimental group's extraneous load decreased from pre-test ($M = 4.2$, $SD = 1.1$) to post-test ($M = 2.9$, $SD = 0.8$), representing a 32% reduction, while the control group's extraneous load increased slightly (pre-test $M = 4.1$, $SD = 1.0$; post-test $M = 4.3$, $SD = 1.1$; $p = .062$).

For germane cognitive load (i.e., mental effort focused on learning), the experimental group showed a significant increase (pre-test $M = 3.8$, $SD = 1.0$; post-test $M = 5.7$, $SD = 0.9$; $p < .001$), while the control group showed no significant change (pre-test $M = 3.7$, $SD = 1.1$; post-test $M = 3.9$, $SD = 1.0$; $p = .214$). Intrinsic cognitive load (related to concept complexity) did not differ between groups ($p = .341$), indicating ILTs did not simplify content but rather enhanced processing efficiency.

3.1.3 Intrinsic Motivation (IMS)

Repeated-measures ANOVA revealed a significant group × time interaction for overall intrinsic motivation ($F(1, 318) = 68.29$, $p < .001$, $\eta^2 = .18$). The experimental group's motivation scores increased from pre-test ($M = 3.2$, $SD = 0.7$) to post-test ($M = 4.5$, $SD = 0.5$; $p < .001$), while the control group's scores decreased (pre-test $M = 3.1$, $SD = 0.8$; post-test $M = 2.7$, $SD = 0.9$; $p = .003$).

Subscale analysis showed significant increases in the experimental group for autonomy ($d = 1.52$, $p < .001$), competence ($d = 1.67$, $p < .001$), relatedness ($d = 1.34$, $p < .001$), and situational interest ($d = 1.73$, p < .001). For example, the experimental group's competence scores rose from 3.1 (pre-test) to 4.6 (post-test), reflecting increased confidence in applying scientific reasoning skills, whereas the control group's competence scores dropped from 3.0 to 2.6 (p = .002).

3.2 Qualitative Results

Two overarching themes emerged from the interview and observation data: **"Scaffolding of Reasoning Through Immersion"** and **"Challenges of Technology Integration"**, with subthemes aligned to the study's theoretical frameworks.

3.2.1 Theme 1: Scaffolding of Reasoning Through Immersion

Adolescents in the experimental group frequently linked ILT use to enhanced scientific reasoning, particularly highlighting the role of interactive simulation. One student noted: "In the VR cell module, I could change the temperature and watch how mitochondria stopped working—this helped me figure out how to test my hypothesis about oxygen and cell function, which I couldn't do with the textbook" (Participant 43, 14 years old). This aligns with situated learning theory, as the immersive environment allowed students to engage in authentic scientific practices (e.g., variable manipulation) that mirrored real-world research.

Teachers also emphasized ILTs' role in reducing cognitive load. A teacher explained: "Students used to get confused when I talked about food webs—they'd mix up producers and consumers. With the AR module, they could see the web overlayed on plants, and the extra cognitive work of visualizing it was gone" (Teacher 7, 9 years of experience). Observation notes further supported this: experimental group students spent 68% of class time actively engaged in reasoning tasks (e.g., debating data interpretations), compared to 32% in the control group, where most time was spent on note-taking or listening to lectures.

Metacognitive awareness was another key subtheme. Over 80% of experimental group students mentioned using the embedded prompts (e.g., "What data do you need?") to reflect on their reasoning. As one student stated: "The VR would ask me why I thought a change happened, and that made me go back and check my data—something I never did with worksheets" (Participant 89, 15 years old). This aligns with CLT, as the prompts directed mental effort toward germane cognitive processes (e.g., self-monitoring) rather than extraneous tasks.

3.2.2 Theme 2: Challenges of Technology Integration

Despite positive outcomes, three main challenges were identified. First, **technical issues** (e.g., VR headset connectivity, AR marker recognition) disrupted 12% of experimental group sessions, with one teacher noting: "When the headsets don't work, we lose time, and students get frustrated" (Teacher 3). Second, **differential technology familiarity** emerged: students from lower socioeconomic backgrounds (42% of the sample) reported feeling less confident using ILTs initially, though this gap narrowed after 4 weeks of practice. Third, **time constraints** were cited by 10 of 12 teachers, who noted that preparing ILT activities required more planning time than traditional lessons.

4. Discussion

4.1 Key Findings and Theoretical Implications

This study's mixed-methods results provide three critical contributions to the intersection of cognitive science, educational psychology, and learning technologies:

First, ILTs significantly enhance adolescents' scientific reasoning skills, with large effect sizes across all subskills (d = 1.15–1.42). This addresses the literature gap identified earlier (Huang et al., 2020) by demonstrating that ILTs support not just factual knowledge but also complex cognitive processes like hypothesis generation and causal inference. From a theoretical perspective, this aligns with **situated learning theory** (Lave & Wenger, 1991): the immersive, interactive environments of ILTs allow students to construct reasoning skills through authentic practice, rather than passive absorption of information. For example, manipulating variables in the AR ecosystem module mirrors the work of real scientists, enabling students to transfer reasoning skills to novel contexts—a key goal of science education (Bransford et al., 2000).

Second, ILTs reduce extraneous cognitive load by 32% and increase germane load, supporting **cognitive load theory** (Sweller, 1988). The visual, interactive nature of ILTs leverages dual-coding theory (Paivio, 1971) to distribute information across verbal and visual working memory channels, reducing overload for adolescents with developing executive functions (Steinberg, 2014). Qualitative data further confirm this: students and teachers reported that ILTs eliminated the "mental work" of visualizing abstract concepts (e.g., cellular respiration), freeing up cognitive resources for reasoning. This finding explains *why* ILTs enhance reasoning—they optimize instructional design to match adolescents' cognitive capacities—addressing the second literature gap.

Third, ILTs boost intrinsic motivation by satisfying autonomy, competence, and relatedness needs (SDT; Ryan & Deci, 2000), with the largest effect on situational interest ($d = 1.73$). This is critical because declining motivation in middle school science (Osborne et al., 2003) often limits engagement with reasoning tasks. The qualitative data highlight how ILTs foster motivation: students valued the ability to explore at their own pace (autonomy), gained confidence from immediate feedback (competence), and collaborated with peers in virtual spaces (relatedness). This motivational boost likely mediated the relationship between ILTs and reasoning—engaged students are more likely to invest effort in complex cognitive tasks—providing a holistic understanding of ILT effectiveness.

4.2 Practical Implications

The findings offer actionable guidance for science educators, learning technology designers, and school administrators:

For **educators**: ILTs should be integrated into science curricula with intentional scaffolding (e.g., metacognitive prompts, real-time feedback) to maximize reasoning gains. Teachers should also provide initial support for students with limited technology familiarity, as this reduces early frustration. For example, a 1-week "orientation" to VR/AR tools could help bridge socioeconomic gaps in technology access.

For **technology designers**: ILTs should prioritize technical reliability to minimize disruptions, as even brief connectivity issues can reduce engagement. Designers should also include customizable difficulty levels to accommodate diverse learning needs—for instance, allowing teachers to adjust the complexity of simulation variables based on student skill level. Additionally, embedding built-in formative assessments (e.g., automated feedback on hypothesis quality) could further reduce teacher workload, addressing the time-constraint challenge identified in qualitative data.

For **administrators**: Investing in ILT infrastructure (e.g., VR headsets, AR-compatible devices) and teacher training is critical. The 8-hour training provided in this study was sufficient to support effective implementation, but ongoing professional development (e.g., monthly workshops on ILT lesson design) could enhance long-term use. Administrators should also consider equity when allocating resources—ensuring schools with high numbers of low-income students have equal access to ILTs to avoid widening achievement gaps.

4.3 Limitations and Future Directions

This study has three key limitations. First, the sample was limited to 8th-grade life science students in the southwestern United States, so results may not generalize to other grade levels, subjects (e.g., physical science), or regions. Future research should test ILTs with diverse populations (e.g., high school students, English language learners) and in different scientific domains to assess generalizability.

Second, the study focused on a 10-week curriculum, so long-term effects of ILTs on scientific reasoning (e.g., retention after 6 months) are unknown. Future studies could include follow-up assessments to determine if ILT-induced reasoning gains persist over time, as this is critical for evaluating the sustained impact of technology integration.

Third, while the mixed-methods design provided rich insights, the study did not explore potential moderators (e.g., prior technology experience, cognitive ability) of ILT effectiveness. For example, do students with stronger working memory benefit more from ILTs than those with weaker working memory? Future research could use regression analyses to identify such moderators, enabling more targeted ILT implementation.

4.4 Interdisciplinary Value and Global Educational Implications

Beyond addressing specific literature gaps and offering local practical guidance, this study underscores the transformative potential of interdisciplinary collaboration between cognitive science, educational psychology, and learning technology—core to the mission of journals like Psychology of Education and Learning Sciences. Traditional educational research often operates in silos: cognitive scientists may focus on theoretical models of reasoning without testing real-world applications, while technology developers may prioritize technical innovation over alignment with adolescent cognitive and motivational needs. This study's integrated approach—using CLT to inform ILT design, SDT to measure motivational impacts, and situated learning theory to interpret reasoning gains—demonstrates how bridging these fields can produce more robust, actionable insights. For example, the finding that ILTs reduce extraneous load by leveraging dual-coding theory (cognitive science) would not have been fully contextualized without qualitative data on how teachers and students experienced that load reduction (educational psychology), nor would the design of the VR/AR modules have been optimized without learning technology expertise in interactive simulation.

This interdisciplinary framework also holds relevance for global science education contexts. While this study was conducted in the U.S., the core challenges it addresses—low adolescent engagement in scientific reasoning, abstract concept difficulty, and uneven technology integration—are universal (OECD, 2019). In regions with limited access to high-end ILTs (e.g., low- and middle-income countries), the study's emphasis on "intentional scaffolding" (e.g., metacognitive prompts, simplified simulation variables) suggests that even low-cost or web-based immersive tools (e.g., 360° videos) could be adapted to support reasoning, provided they align with cognitive and motivational principles. Additionally, the focus on equity—addressing socioeconomic gaps in technology familiarity—offers a model for global educators seeking to avoid "digital divides" in STEM learning. By grounding ILT implementation in interdisciplinary theory, educators worldwide can move beyond "technology for technology's sake" and toward evidence-based practices that prioritize cognitive growth and inclusive learning.

5. Conclusion

This study demonstrates that immersive learning technologies (ILTs) enhance adolescents' scientific reasoning by aligning with core principles of cognitive science (CLT, dual-coding theory) and educational psychology (situated learning, SDT). The quantitative results show large, statistically significant gains in reasoning skills and motivation, while qualitative data explain *how* ILTs achieve these outcomes—by reducing cognitive load, providing authentic practice, and satisfying psychological needs. These findings address critical gaps in the literature and offer practical guidance for integrating ILTs into science education.

As technology continues to transform education, ILTs represent a powerful tool for fostering the complex cognitive skills needed for 21st-century scientific literacy. By grounding ILT design and implementation in interdisciplinary theory, educators and researchers can unlock their full potential to support adolescent learning—preparing students not just to understand science, but to *reason* like scientists.

References

[1] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

[2] Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). How people learn: Brain, mind, experience, and school (Expanded ed.). National Academy Press. https://doi.org/10.17226/9853

[3] Dunleavy, M., & Dede, C. (2014). Using augmented reality and virtual reality for science learning. Journal of Science Education and Technology, 23(1), 1–5. https://doi.org/10.1007/s10956-013-9432-9

[4] Huang, R., Rauch, U., & Liaw, S. Y. (2020). A systematic review of virtual reality in science education: Identifying a gap in lower secondary education. Computers & Education, 151, 103856. https://doi.org/10.1016/j.compedu.2020.103856

[5] Lawson, A. E. (2000). The development and validation of a classroom test of formal reasoning. Journal of Research in Science Teaching, 15(1), 11–24. https://doi.org/10.1002/tea.3660150103

[6] Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

[7] Mayer, R. E. (2020). Multimedia learning (4th ed.). Cambridge University Press. https://doi.org/10.1017/9781108622375

[8] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. https://doi.org/10.1037/h0043158

[9] National Research Council (NRC). (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. The National Academies Press. https://doi.org/10.17226/13165

[10] Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. International Journal of Science Education, 25(9), 1049–1079. https://doi.org/10.1080/0950069032000086890

[11] Paas, F., Tuovinen, J. E., Tabbers, H., et al. (2003). Cognitive load measurement as a means to advance cognitive load theory. Educational Psychologist, 38(1), 63–71. https://doi.org/10.1207/S15326985EP3801_07

[12] Paivio, A. (1971). Imagery and verbal processes. Holt, Rinehart & Winston.

[13] Pearson Education. (2020). Pearson life science (Florida ed.). Pearson.

[14] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

[15] Steinberg, L. (2014). Adolescence (10th ed.). McGraw-Hill Education.

[16] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

[17] Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. Developmental Review, 27(2), 172–223. https://doi.org/10.1016/j.dr.2007.01.001

[18] Baker, L., & Mayer, R. E. (2019). Cognitive theory of multimedia learning. In S. Graham, C. MacArthur, & J. Fitzgerald (Eds.), Handbook of writing research (2nd ed., pp. 371–388). Guilford Press.

[19] Chen, B., & Yang, S. (2019). Augmented reality in science education: A systematic review of research and applications. Computers & Education, 137, 103698. https://doi.org/10.1016/j.compedu.2019.103698

[20] Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), Handbook of theories of social psychology (Vol. 1, pp. 416–436). Sage Publications.

[21] Dede, C. (2017). Immersive learning technologies: Realism and online experiences in K-12 education. Review of Research in Education, 41(1), 442–471.

[22] Fiorella, L., & Mayer, R. E. (2016). The relative benefits of learning by teaching and teaching expectancy. Contemporary Educational Psychology, 45, 14–22. https://doi.org/10.1016/j.cedpsych.2015.10.001

[23] Gobert, J. D., & Pallant, A. (2016). Using technology to support scientific practices in K-12 science classrooms. Journal of Science Education and Technology, 25(2), 160–164. https://doi.org/10.1007/s10956-016-9617-0

[24] Hsu, Y. S., & Tsai, C. C. (2018). The effects of virtual reality on students' learning outcomes in science education: A meta-analysis. Journal of Educational Technology & Society, 21(4), 188–201.

[25] Johnson, D. W., & Johnson, R. T. (2017). Learning together and alone: Cooperative, competitive, and individualistic learning (7th ed.). Allyn & Bacon.

[26] Kim, M. K., & Lee, K. (2020). Virtual reality-based science learning: A systematic review of research from 2010 to 2019. Computers & Education, 151, 103849.

[27] Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. Cognitive Science, 12(1), 1–48. https://doi.org/10.1207/s15516709cog1201_1

[28] Kosko, K. W., & Wilkins, J. L. M. (2019). The effects of virtual reality on student achievement in K-12 and higher education: A systematic review. Journal of Educational Technology Systems, 47(4), 447–469. https://doi.org/10.1177/0047239519856649

[29] Lee, Y. J., & Chen, S. W. (2021). Augmented reality for STEM education: A systematic review of empirical studies. Computers & Education, 165, 104109. https://doi.org/10.1016/j.compedu.2021.104109

[30] Lin, S. S., & Liu, C. C. (2017). The impact of virtual reality on students' conceptual understanding of chemical bonding. Journal of Chemical Education, 94(12), 2068–2074.

[31] Mann, E. L., & Pintrich, P. R. (2003). Motivation and self-regulated learning in science. In W. F. McComas (Ed.), The nature of science in science education (pp. 331–359). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-3417-2_15

[32] Mayer, R. E., & Fiorella, L. (2014). Learning as instruction: An instructional design framework based on the science of learning. Journal of Educational Psychology, 106(4), 1064–1073. https://doi.org/10.1037/a0037681

[33] McElhaney, K. W., & Linn, M. C. (2011). Technology-enhanced science inquiry learning: The influence of computer simulations and curriculum design on learning processes. Journal of Research in Science Teaching, 48(10), 1137–1160. https://doi.org/10.1002/tea.20462

[34] Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. Educational Psychology Review, 19(3), 309–326. https://doi.org/10.1007/s10648-007-9066-9

[35] Nokes-Malach, T. J., & Mestre, J. P. (2013). Learning science concepts through invention and collaboration. Journal of the Learning Sciences, 22(2), 247–289. https://doi.org/10.1080/10508406.2012.677459

[36] Park, J. Y., & Liu, M. (2020). Virtual reality in K-12 education: A systematic review of the literature from 2010 to 2019. Computers & Education, 151, 103863. https://doi.org/10.1016/j.compedu.2020.103863

[37] Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. Journal of Educational Psychology, 95(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

[38] Quintana, C., Reiser, B. J., Davis, E. A., et al. (2004). A scaffolding design framework for software to support science inquiry. The Journal of the Learning Sciences, 13(3), 337–386. https://doi.org/10.1207/s15327809jls1303_2

[39] Renninger, K. A., & Hidi, S. (2016). The role of interest in learning and development. Annals of Child Development, 21, 1–25. https://doi.org/10.1177/0305409716631384

[40] Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. Journal of Educational Psychology, 76(4), 604–616. https://doi.org/10.1037/0022-0663.76.4.604

[41] Schmidt, H. G., Loyens, S. M. M., van Gog, T., et al. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). Educational Psychologist, 42(2), 91–97. https://doi.org/10.1080/00461520701263368

[42] Shin, N., & Son, J. (2019). The effects of virtual reality-based instruction on students' learning outcomes and cognitive load in biology education. Computers & Education, 132, 26–37. https://doi.org/10.1016/j.compedu.2019.01.004

[43] Smetana, L. K., & Rote, W. M. (2015). Social cognitive development in adolescence. In K. W. Schaie & S. L. Willis (Eds.), Handbook of the psychology of aging (8th ed., pp. 249–267). Academic Press. https://doi.org/10.1016/B978-0-12-397944-7.00012-7

[44] Thompson, J., & Dass, P. (2020). Augmented reality for enhancing scientific reasoning in middle school chemistry. Journal of Science Education and Technology, 29(3), 289–302. https://doi.org/10.1007/s10956-020-09841-1

[45] Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. Medical Education, 44(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03561.x

[46] Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

[47] Wang, S., & Chen, Y. (2021). Virtual reality and scientific reasoning: A meta-analysis of experimental studies. Journal of Educational Technology & Society, 24(2), 1–14.

[48] Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulated learning and science education: A review of metacognition, strategy use, and motivation. Journal of Research in Science Teaching, 45(3), 334–352.

[49] Wu, H. K., & Lee, Y. H. (2018). The effects of virtual reality on students' spatial ability in science learning. Computers & Education, 125, 376–388.