TYPE Review
PUBLISHED 17 September 2025
DOI 10.3389/fcomm.2025.1624012

Check for updates

OPEN ACCESS

EDITED BY
Parlan Parlan,
State University of Malang, Indonesia

REVIEWED BY
R. Ahmad Zaky El Islami,
Universitas Sultan Ageng Tirtayasa, Indonesia
Mohammad Taufiq Abdul Ghani,
Universiti Pendidikan Sultan Idris, Malaysia

*CORRESPONDENCE
Elvira G. Galimova
✉ elyagalimowa@yandex.ru

RECEIVED 06 May 2025
ACCEPTED 04 August 2025
PUBLISHED 17 September 2025

CITATION
Galimova EG, Sergeeva OV, Zheltukhina MR, Sokolova NL, Zakharova VL and Drobysheva NN (2025) Mobile learning in science education to improve higher-order thinking skills and communication skills: scoping review.
Front. Commun. 10:1624012.
doi: 10.3389/fcomm.2025.1624012

COPYRIGHT
© 2025 Galimova, Sergeeva, Zheltukhina, Sokolova, Zakharova and Drobysheva. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Mobile learning in science education to improve higher-order thinking skills and communication skills: scoping review

Elvira G. Galimova[1,2]*, Olga V. Sergeeva[3], Marina R. Zheltukhina[4], Natalia L. Sokolova[5], Valeria L. Zakharova[6] and Nataliya N. Drobysheva[7]

[1]Department of Pedagogy of Higher Education, Kazan (Volga region) Federal University, Kazan, Russia, [2]The Federal State Budget Scientific Institution (Federal Scientific Center of Psychological and Multidisciplinary Research), Kazan, Russia, [3]Department of English Philology, Kuban State University, Krasnodar, Russia, [4]Scientific and Educational Center (Person in Communication), Pyatigorsk State University, Pyatigorsk, Russia, [5]Institute of Foreign Languages, Peoples' Friendship University of Russia (RUDN University), Moscow, Russia, [6]Department of Medical and Social Assessment, Emergency and Ambulatory Practice, Sechenov First Moscow State Medical University, Moscow, Russia, [7]Department of English Language and Professional Communication, Financial University under the Government of the Russian Federation, Moscow, Russia

Mobile devices increasingly permeate students' lives, yet their impact on core 21st-century competencies in science classrooms remains unevenly charted. This scoping review maps how mobile-learning interventions influence higher-order thinking skills (HOTS) and communication skills across primary, secondary and tertiary science education. Guided by Arksey and O'Malley's five-stage framework and reported in line with PRISMA-ScR, six databases were searched for English-language studies published between 1 January 2015 and 31 March 2025. Twenty-seven empirical papers met the inclusion criteria and were charted for context, pedagogy, technology and outcomes. Most interventions originated in Indonesia and combined purpose-built science apps or smartphone sensors with ubiquitous chat platforms such as WhatsApp or Viber. Inquiry and problem-based models dominated, typically embedding real-time data collection, instant peer sharing and scaffolded reflection. Across studies, moderate-to-large gains were reported for critical thinking, creativity and, to a lesser extent, verbal and written scientific communication. Designs that coupled evidence gathering with public dissemination (e.g., WeChat science posts) yielded the strongest communication improvements. Recurring challenges included short intervention windows, small intact-class samples, technical glitches and limited teacher preparation. The review concludes that mobile devices can catalyze sustained HOTS and richer scientific discourse when inquiry-rich tasks are buttressed by social-communication channels and graduated scaffolds. Future research should extend trials over full semesters, diversify geographic settings and employ process analytics to trace how mobile interactions translate into durable learning.

KEYWORDS

mobile learning, science education, higher-order thinking skills, communication skills, scoping review

Introduction

Mobile devices have become the most pervasive digital tools in young people's lives, blurring the boundaries between school and everyday contexts. When repurposed for learning, their portability, continuous connectivity and context awareness let students collect data in the field, revisit resources on the move and exchange ideas in real time (Ly and Kearney, 2023; Naveed et al., 2023). The integration of mobile technologies in education has gained significant momentum, with faculty members across technical-engineering disciplines increasingly recognizing their pedagogical potential for enhancing student engagement and learning outcomes (Mohammadi et al., 2020). Socio-constructivist, activity-theory and multimedia-cognitive lenses all predict that such affordances can deepen participation and understanding by coupling action with immediate social dialog and dual-channel processing (Kumar et al., 2021; Mayer, 2024). Science education, which prizes hands-on inquiry and communication of evidence, therefore appears especially suited to mobile learning's strengths (Criollo-C et al., 2021; Gumbheer et al., 2022).

Higher-order thinking skills (HOTS) and clear scientific communication stand at the heart of current science curricula globally. The development of HOTS—encompassing analyzing, evaluating, and creating—has emerged as a critical competency within science and STEM education contexts, recognized as essential for preparing students to navigate complex scientific challenges and contribute meaningfully to technological advancement (Haryadi and Pujiastuti, 2022; Wahono et al., 2020; Wahyuni et al., 2022). These skills equip learners to transfer concepts, solve novel problems and critique claims, while communication skills enable them to articulate findings, question peers and engage public audiences (Affandy et al., 2024; Montgomery et al., 2022; Vilela et al., 2025). The importance of HOTS development extends beyond traditional classroom boundaries, with recent phenomenological studies revealing how e-learning contexts in higher education provide unique opportunities for fostering complex cognitive processes through digital mediation and collaborative knowledge construction (Khadka et al., 2025).

Reviews and meta-analyses now show moderate-to-large average effects of mobile interventions on critical thinking, creativity and, to a lesser extent, communication (Ansori et al., 2024; Liu and Zhang, 2022). However, the field faces several critical gaps that limit both theoretical understanding and practical implementation. Recent systematic analyses reveal that while STEM enactment effectiveness has been documented across Asian educational contexts (Wahono et al., 2020), the mechanisms underlying technology-mediated HOTS development remain inadequately theorized, particularly regarding how mobile affordances specifically contribute to higher-order cognitive processes.

Three design moves recur across successful studies: real-time data collection, instant sharing through chat or cloud slides and scaffolded prompts that guide reflection (Blackmore and Rønningsbakk, 2023; Kapici et al., 2022; Khery et al., 2020). These patterns suggest that mobile devices can be powerful catalysts when embedded in inquiry-rich pedagogy (Demircioglu et al., 2023; Inel-Ekici and Ekici, 2022). Yet, significant gaps remain in understanding how different pedagogical approaches, from STEM Project-Based Learning models that have shown promise in enhancing pre-service teachers' higher-order thinking capabilities (Haryadi and Pujiastuti, 2022) to innovative assessment methods like question card games that improve students' analytical skills in biology education (Wahyuni et al., 2022), can be optimally integrated with mobile learning technologies.

A critical limitation in current mobile learning research is the uneven geographic distribution of empirical studies, with substantial representation from specific regions while other educational contexts remain underexplored. This geographic concentration creates potential cultural bias in understanding universal versus context-specific effectiveness mechanisms, particularly given that mobile technology adoption and educational integration patterns vary significantly across different cultural and infrastructure contexts (Mohammadi et al., 2020).

While HOTS development through mobile technologies has received increasing attention, the parallel development of communication skills, equally essential for scientific literacy, remains inadequately documented and theoretically underdeveloped. The field lacks comprehensive frameworks that systematically link mobile technology affordances to specific communication skill development pathways, creating a significant gap between technological potential and pedagogical understanding.

The growing prominence of e-learning environments, accelerated by global educational disruptions, has highlighted the need for deeper understanding of how HOTS development occurs in digitally-mediated contexts (Khadka et al., 2025). Mobile learning, as a subset of e-learning approaches, requires specific investigation regarding its unique contributions to higher-order cognitive development, particularly in science education where hands-on inquiry traditionally dominates pedagogical approaches.

Against this backdrop of expanding evidence yet persistent gaps, the present scoping review aims to map the landscape of mobile-learning interventions in school and university science that report outcomes on both higher-order thinking and communication. Guided by Arksey and O'Malley (2005) framework, it addresses three enhanced research questions:

1. What pedagogical approaches and implementation models characterize mobile-supported science lessons that target HOTS and communication, and how do these vary across different cultural and educational contexts?
2. How do these interventions operationalize and measure the development of HOTS and communication skills, with particular attention to assessment adequacy and theoretical framework integration?
3. What contextual, methodological, and cultural patterns and gaps emerge across the evidence base, and what are the implications for cross-cultural transferability and implementation sustainability?

By tracing these dimensions through systematic analysis of empirical studies, the review seeks to clarify where the field now stands, identify critical blind spots that limit both theoretical understanding and practical implementation, and provide evidence-based guidance for how future studies might design, implement and evaluate mobile-science learning interventions for lasting cognitive and communicative impact across diverse educational contexts. The analysis explicitly addresses the need for cultural sensitivity in mobile learning research while maintaining focus on identifying universal mechanisms that transcend specific technological or geographic constraints.

Literature review

Mobile learning

Mobile learning is a form of e-learning. It uses portable devices such as smartphones, tablets, laptops, audio players, and e-books to deliver content and support interaction (Hamidi and Chavoshi, 2018). Students can learn anywhere and anytime because the devices are light and always connected (Gikas and Grant, 2013; Naveed et al., 2023). Portability, connectivity, and context awareness are key features. They let learners collect data in the field, review resources during travel, and share ideas in real time (Ly and Kearney, 2023).

Several theories explain the benefits. Socio-constructivist theory says knowledge grows through talk and shared actions. Mobile chat and cloud apps give fast channels for this talk (Shao and Liu, 2021). Activity theory sees learning as an action that is guided by tools like sensors and apps (Kumar et al., 2021). Multimedia Cognitive Theory adds that dual audio-visual channels in mobile media reduce cognitive load and help select and organize information (Mayer, 2024).

Mobile learning also blurs the line between formal and informal study. It turns the bus ride, the home kitchen, or the school yard into parts of the classroom (Afikah et al., 2022). There are still challenges. Teachers need training, and schools need policy, bandwidth, and clear rules. Money limits, weak infrastructure, and parent worries about screen time also slow progress (Criollo-C et al., 2021; Gumbheer et al., 2022).

Higher-order thinking skills

HOTS are mental processes that rise above remembering facts. Students with HOTS can transfer concepts, link ideas, process information to solve problems, and analyze ideas critically (Sun et al., 2022; Supeno et al., 2019; Seif, 2023). Bloom's original taxonomy placed analysis, synthesis, and evaluation at the peak of cognition. The revised taxonomy by Anderson and Krathwohl (2001) groups HOTS under analyzing, evaluating, and creating, and adds a metacognitive knowledge layer. This layer points to thinking about one's own thinking.

Typical HOTS activities include constructing arguments, asking research questions, comparing alternatives, solving complex non-algorithmic problems, handling controversies, and spotting hidden assumptions (Affandy et al., 2024; Supeno et al., 2019). According to Lu et al. (2021), students' HOTS were directly impacted by learning motivation and peer interaction. Educators value HOTS because they prepare learners for the information age. These skills support critical thinking, creativity, and responsible action in society (Affandy et al., 2024; Ahmad et al., 2020). Learning that targets HOTS also boosts student motivation and knowledge retention (Supeno et al., 2019). Assessment shapes HOTS development. Tests and tasks that demand analysis or evaluation push students toward deeper understanding and longer retention of core facts (Affandy et al., 2024; Anderson and Krathwohl, 2001; Sun et al., 2022).

Communication skills

Communication skills let students share ideas and receive feedback. They cover speaking, listening, writing, reading, and visual display. Good communication means a student can articulate, explain, describe, clarify, listen, question, and share in ways that others understand (Afikah et al., 2022; Elkot et al., 2025). Communication belongs to the core set of 21st-century skills. Alongside information and media literacy it helps young people work in teams, solve problems, and use technology effectively (Ahmad et al., 2020; Sergeeva et al., 2023; Vilela et al., 2025). Science teachers therefore aim to weave communication into every lesson.

Research links communication to deep learning. When students talk or write about a concept they reorganize knowledge, test ideas, and build shared meaning. Mobile devices expand this process because they connect learners to content, peers, and teachers at any place or time (Vilela et al., 2025).

Many studies still report weak student communication. Supeno et al. (2019) found that high-school physics students struggled to express scientific ideas in writing even after inquiry tasks. Afikah et al. (2022) note similar gaps in oral clarity and confidence. Effective pedagogy gives structured talk and writing time. Collaborative learning invites group explanation and peer feedback. Problem-based and project-based learning ask students to debate evidence, draft reports, and present solutions, all of which raise verbal and written fluency (Afikah et al., 2022; Montgomery et al., 2022; Vilela et al., 2025). Teachers who combine these approaches with mobile chat, shared slides, and video tools create frequent, authentic practice. Communication skills are broad, essential, and teachable. They underpin knowledge construction and workplace readiness. Technology and inquiry methods provide new chances to rehearse these skills, but students still need clear guidance, feedback, and time to practice.

Relation among mobile learning, HOTS and communication skills

Mobile learning uses portable digital devices such as smartphones and tablets so that students can reach learning materials at any place and time (Ahmad et al., 2020). In science lessons these devices add cameras, sensors, and social apps. Such tools fit the goals of the 21st century. Teachers wish to grow higher-order thinking skills (HOTS) like analysis, evaluation, and creation, and also clear communication (Ahmad et al., 2020; Ansori et al., 2024; Hamidi and Chavoshi, 2018).

Research shows that mobile devices help HOTS (Ahmad et al., 2020; Khery et al., 2020; Liu and Zhang, 2022; Liu et al., 2024). Ahmad et al. (2020) designed an inquiry cycle in which learners photograph plants, log data in a cloud sheet, and explain patterns in a chat room. Each step forces students to ask why, compare results, and build new ideas. The authors report large gains in analysis and synthesis scores.

Field studies give similar messages. Supeno et al. (2019) observed 120 Indonesian students who used a physics data-logging app during field trips. Students first solved routine tasks. They then faced ill-structured tasks that needed new plans and teamwork. Post-tests show a rise in creative problem-solving and scientific writing. Interviews add that the phone screen makes talk easy because everyone can see the same graph. This shared view supports communication.

Systematic reviews give a wider view. Afikah et al. (2022) screened 30 studies from 2012 to 2021. They list six common teaching models: collaborative, inquiry, project-based, problem-based, game-based, and flipped classroom. All models raise HOTS, but collaborative and project designs also lift verbal explanation ability. Many studies use group chat, voice notes, or shared whiteboards for that purpose. A later review by Ansori et al. (2024) covers work from 2018 to 2023. Smartphones form 78% of devices. Inquiry and problem-based learning appear most often. Average effect sizes are moderate for critical thinking (g = 0.56) and strong for creativity (g = 0.82). Communication gets a smaller but steady effect (g = 0.42). The review warns that most trials last less than 6 weeks and rely on self-report rubrics.

Recent case reports add useful detail. Afikah et al. (2022) describe a chemistry augmented-reality app where students scan lab tools and watch 3-D animations. Users had to narrate steps to peers through voice messages. Scores on clarity of explanation doubled. Another study by Kapici et al. (2022) shows that virtual lab lifts both reasoning and argumentative writing in science.

Across this body of work, three design features recur. First, real-time data collection lets students test ideas in context. Second, instant sharing through chat or cloud slides keeps talk active. Third, scaffolded prompts guide reflection. These features push students beyond recall, toward HOTS, and give them many chances to shape and refine language. Nevertheless, gaps remain. Many studies use small samples and short tasks. Few track long-term retention or follow actual classroom schedules. Future research should run full-semester programs and study how teachers manage device use, privacy, and student talk during routine lessons. Longer trials will also help to measure true gains in communication confidence.

Methodology

To capture the breadth and methodological diversity of evidence on how mobile-learning interventions influence higher-order thinking and communication skills in science education, we adopted a scoping-review design. Scoping reviews are recommended when the purpose is to map key concepts, sources, and knowledge gaps within a heterogeneous body of literature (Arksey and O'Malley, 2005). Our procedure followed the refinements proposed by Levac et al. (2010) emphasizing iterative team reflexivity, transparent decision logs, and stakeholder consultation—and the step-by-step guidance of the Joanna Briggs Institute Manual for Evidence Synthesis. Accordingly, we undertook the five canonical stages: (1) clarifying the review question, (2) systematically identifying relevant studies, (3) applying explicit selection criteria, (4) charting data with a calibrated extraction form, and (5) collating, summarizing and reporting the results.

Data collection process

Following the PRISMA-ScR guidelines (Tricco et al., 2018) and Arksey and O'Malley (2005) five-stage scoping-review framework, we searched literature published in English between 1 January 2015 and 31 March 2025 with the following search query in scopus and WoS databases.

Search query

(("mobile learning" OR "m-learning" OR "mobile education" OR "mobile technology" OR "mobile device*" OR "smartphone*") AND ("science education" OR "science teaching" OR "science learning") AND (("higher-order thinking" OR "HOTS" OR "critical thinking" OR "problem solving" OR "analytical thinking" OR "creative thinking" OR "evaluative thinking") OR ("communication skill*" OR "communicative competenc*" OR "verbal communication" OR "scientific communication" OR "science communication")))

As shown in Figure 1, the initial searches returned 39 records in Web of Science and 34 in Scopus (73 total). 21 articles are duplicated publications. Two independent reviewers applied predefined inclusion criteria—empirical studies that applied mobile learning in a science-education context and reported outcomes on higher-order thinking and/or communication skills—and exclusion criteria (pre-2015 works, inaccessible full texts, reviews, book chapters). After review, 8 publications remained. To broaden coverage, concept-equivalent searches were repeated in ERIC (yielding five additional papers), IOPscience (six papers) and Google Scholar, produced eight more publications. Total 27 publications are selected for scoping review analysis.

Quality assessment protocol

Although quality assessment is not mandatory for scoping reviews, publications were assessed according to the Mixed Methods Assessment Tool (MMAT) to increase methodological rigor and contribute to the interpretation of the findings. As indicated in Table 1 in the findings, there is methodological diversity in the studies (experimental, quasi-experimental, quantitative, qualitative and mixed studies). On the other hand, as stated in the data collection process, the publications were selected by two researchers. In this selection process, the studies with differences were included in the publications that both researchers agreed on. In addition, as stated in Tables 2, 3, there is a diversity of pedagogical approaches and information technology. The diversity in educational level presented in Table 4 supports the diversity of the target audience. In terms of geographical distribution, Indonesia may be partly to blame. This may be due to the high emphasis on higher order thinking skills in the Indonesian curriculum and the high number of publications on this subject. Since one of the aims of the scoping review study is to reveal the general situation, it is a valuable result because it identifies the existing situation. Considering the items expressed as a result, it can be stated that there is heterogeneity in the studies.

Data analysis

After the final pool of 27 articles had been identified, both reviewers downloaded the full texts and imported them into a shared Google Sheet that functioned as the data-charting form. Working independently, each researcher read every paper from beginning to end and copied verbatim quotations that answered the review questions (for example, descriptions of the pedagogical approach, of how mobile tools were used, and of measured outcomes for higher-order thinking or communication). Each quotation was pasted into the worksheet beside its bibliographic details and the guiding question it addressed.

The researchers then met online and compared every entry line by line. Wherever the same passage had been selected they confirmed a match; where one researcher had chosen a quotation the other had missed, they returned to the source article together and decided whether the passage was relevant. This adjudication stage continued until both researchers agreed on every quotation, yielding full consensus with no unresolved disagreements.

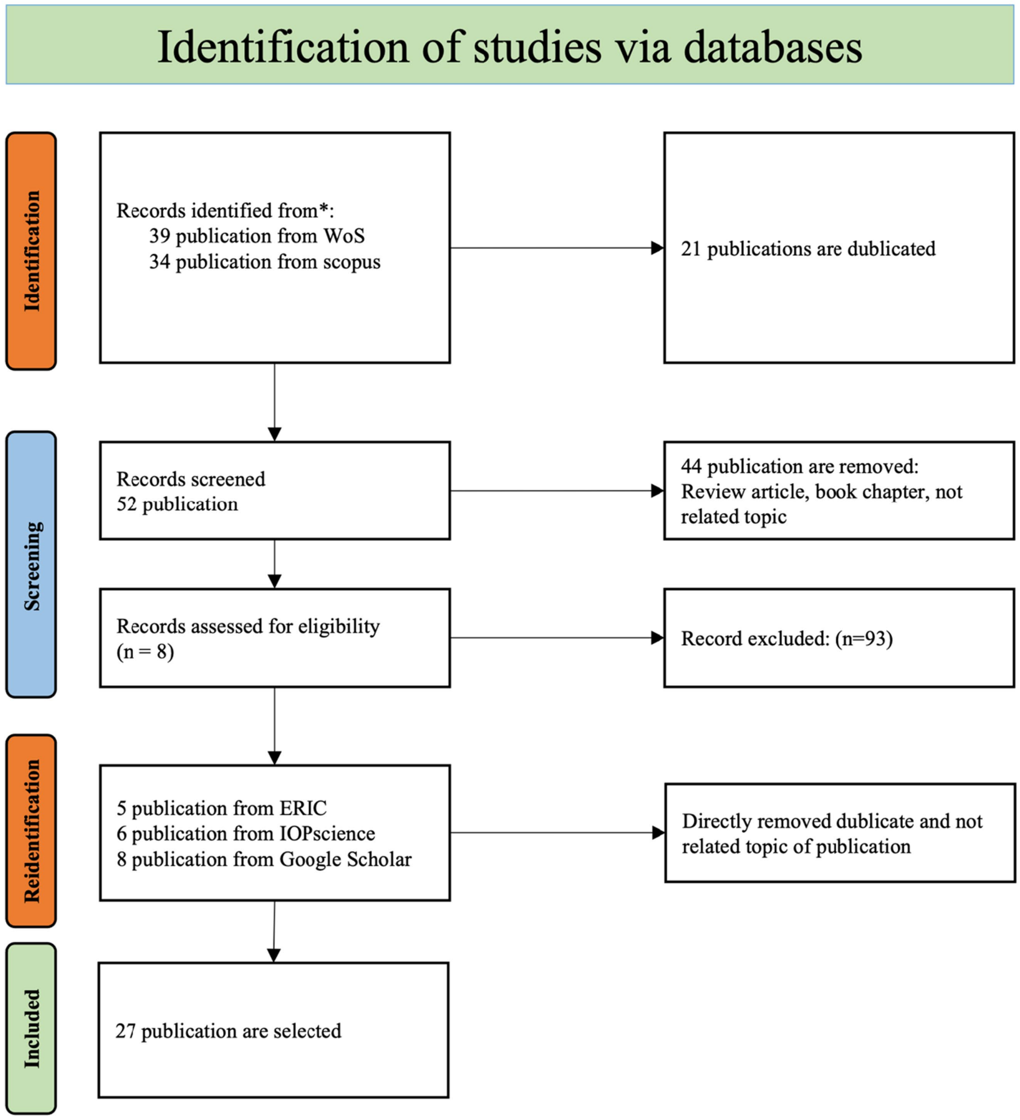

FIGURE 1
Data collection process.

The agreed quotations were then reorganized inside the Google Sheet under the four analytic categories derived from the review questions: (1) pedagogical approaches and implementation models, (2) strategies for developing higher-order thinking skills, (3) approaches to communication-skill development, and (4) contextual or methodological features. Using this structure, the reviewers jointly wrote short analytic summaries for each category, synthesizing the quotations into coherent findings rather than counting frequencies. These narrative syntheses form the basis of the Results and Discussion sections that follow.

Findings

A clear geographic skew emerges: Indonesia dominates the evidence base, contributing more than half of all studies (14 of 27) and publishing consistently every year from 2018 to 2022 (Table 5). A second-tier cluster of work comes from Turkey, Germany, and Hong Kong, each with two or three studies, while the remaining 11 countries appear only once. Temporally, the dataset spans *2016 → 2024,* but activity accelerates after 2018, reflecting the surge of interest in mobile and immersive technologies for science education during—and after—the COVID-19 era. The sparse representation of regions such as Africa, South America, or mainland Europe (outside the German-speaking context) highlights an opportunity for more geographically diverse research.

Purpose-built learning apps are becoming the dominant strategy (Table 2). Eleven of the 27 studies developed their own Android, iOS or hybrid applications, arguing that "ownership of mobile technology has been very widespread, from children to the elderly" (Made Tegeh et al., 2020) and that a custom environment can "accommodate teachers, students, sources and other learning media" (Hariadi et al., 2022). These apps typically weave together

TABLE 1 Methodological approaches.

ID	Research design (labels used by authors)	Data-collection and analysis highlights	Sample size/study duration
1	Pre-/post single-group (4 Cs awareness)	3-point Likert + open-ended; H-gain; content analysis (inter-rater = 0.90)	*n* = 10; 5-month TLS (12 weeks active)
2	4-D R&D model → Pre-/post control-group	MC-HOTS test, validation questionnaires; ANOVA-mixed; IRT	*n* = 1,070 across 5 regions; 1 semester
3	One-group pre-/post (ICT-PBL)	Quizizz HOTS (30 items), product rubrics; n-gain; descriptive	*n* = 36; unspecified but unit-length
4	Classroom action research (2 cycles)	Cognitive tests, ARCS motivation survey, observation	*n* = 86; two action cycles (≈1–2 months)
5	Non-equivalent control-group	Pre/post-tests, five psychosocial scales, ANCOVA; interviews	*n* = 44; ≈4 h intervention (235 min)
6	Quasi-experiment (online forums)	Forum posts scored with Facione CTS rubric; qualitative levels	*n* = 37; 3-week forum inside 14-week course
7	One-shot case study (5 treatments)	Metacognitive rubric; Kruskal-Wallis	*n* = 126; several lessons (duration not given)
8	Mixed-methods competition study	Anonymous survey (α = 0.87), χ^2; open-response content analysis	*n* = 1,095 pool, 47 valid surveys; 1 semester
9	One-group pre–/post	Process-skills and collaboration tests; paired-t, n-gain	*n* = 30; even semester 2021
10	Qualitative action research	Video, discussions, Google Forms; NVivo-directed content	*n* = 11; 14-week course, 2 weeks recorded
11	Holistic multiple-case (exploratory)	6-week IBL; open-ended survey; open/axial coding	*n* = 80 (4 groups); 6 weeks
12	R&D (Luther + Dick and Carey)	Expert & user questionnaires; descriptive stats	Validators + 3 pilot students; prototype cycle
13	Post-test only control-group	NOS, concept, literacy tests (α > 0.8); t-test	*n* = 44; single lesson
14	Pre-experimental (post-test only)	Two-tier HOTS test; multiple regression	*n* = 312; end-of-semester snapshot
15	2 × 2 quasi-experimental ANOVA	HOTS test (16 items), motivation survey (38); two-way ANOVA	*n* = 36; one semester
16	Quasi-experiment (BWML vs. control)	Pre/post, Wilcoxon, n-gain, Mann–Whitney	*n* = 137; trial after model refinement
17	One-group pre–/post (co-blended)	HOTS test; paired-t, descriptive	*n* = 35; short unit
18	Borg-and-Gall development + trials	Expert validation; ANOVA between groups	*n* = 34 limited, wider n > 100; multi-stage
19	3-year longitudinal case study	40 h obs., pupil work, 24 interviews; thematic analysis (NVivo)	~60 pupils/yr.; 3 academic years
20	Case study with process-mining	m-Orchestrate logs → BupaR & FOMM; LA visualization	*n* = 35; 2 weeks (4 lessons + 2 h group work)
21	Case study with control group	MC & open tests; obs. Checklist; Likert; t-test	*n* = 30; 4 × 45 min sessions
22	Quasi-exp. repeated-measures	Scales (interest, curiosity), concept tests; ANCOVA/ regression	*n* = 154 complete; 4 lessons + 6–16 wk. follow-up
23	Descriptive technical paper	Demonstrative experiments; no formal stats	—
24	Mixed-methods (AR vs. text)	Pre/post, SPSS; semi-structured interviews; thematic	*n* = 50; 6-week Telegram-based intervention
25	Mixed-methods quasi-exp.	Pre/post reasoning & comm. Tests; teacher interviews; t-tests	*n* = 434 (5 schools); 3-week post-training
26	Qualitative case study	Video/ audio obs.; Toulmin and argument-level coding	*n* = 79 cohort, 26 focal; 3-week (19 h)
27	Pre-/post non-equivalent groups	HOTS test; normality, homogeneity, t-tests, effect size	*n* = 36; unit in 2018–19 year

multimedia explanations, interactive tasks and instant feedback, scaffolding analysis, evaluation and creation—the core layers of HOTS.

Low-friction communication platforms act as the social glue. Even when an intervention centers on a bespoke app, researchers bolt on familiar channels so that students can deliberate together. In Greece, a Viber group offered "an extra channel for information, communication, and cooperation… also outside the classroom environment" (Kousloglou et al., 2023); Indonesian biology students submitted reflections and presentations through WhatsApp because "during the pandemic … questionnaires were given through the WA-app group chat" (Isnaeni et al., 2021). China's medical-science project pushed this logic further, publishing student work on a WeChat public account that quickly

TABLE 2 Distribution of the technologies.

Technology category	# of studies (*n* = 27)*	Share	Representative examples
Custom learning apps (stand-alone Android/iOS or web-mobile hybrids)	11	41%	IPMLM, MoLearn, electrolyte mobile app, cooperative-blended app
Communication apps & platforms (messaging, video-conf., social media)	7	26%	WhatsApp, Viber, WeChat, Zoom + Google Drive/Forms
Smartphone/tablet sensors & probes	4	15%	Smart Cart + sensors, phyphox, SPARKvue experiments
Augmented Reality (AR)	3	11%	NutricARd + Zappar, EdLab AR lab, multiple free AR apps on tablets
Virtual Reality (VR)	1	4%	EduVenture® spherical-video VR on iPads
Quiz/Assessment-centered apps	1	4%	Quizizz (HOTS item bank)
Interactive e-book / multimedia presentation tools	2	7%	EPUB + Moon Reader, Focusky interactive slides
Simulation software	1	4%	Five-E Inquiry + computer simulations
General tablet creativity apps	1	4%	iMovie, Popplet, Explain Everything, etc.

*Percentages rounded; one study can appear in more than one category when it truly spans categories, but double-counting inside a single study was avoided.

TABLE 3 Pedagogical approach.

Pedagogical-approach family	≈ Share*	Representative studies (ID—references, year)	Typical implementation moves
Inquiry-based learning (IBL/IBSE, open, guided, argument-driven)	10 / 27	1—Kousloglou et al. (2023); 10—Arabacioglu and Unver (2016); 11—Inel-Ekici and Ekici (2022); 26—Demircioglu et al. (2023)	Phased inquiry cycles (orientation → investigation → conclusion); mobile sensors/AR/VR supply data; teacher as facilitator
Problem-based learning (PBL)	4	3—Isnaeni et al. (2021); 11—Inel-Ekici and Ekici (2022); 16—Hariadi et al. (2022); 17—Sulisworo et al. (2018)	Ill-structured problems anchor lessons; learners research and present solutions via mobile apps or LMS
Scaffolding/guided instruction	3	2—Dasilva et al. (2019); 18—Agustihana and Suparno (2018); 1—Kousloglou et al. (2023)	Teacher modeling → gradual release; apps embed hints, worked examples, self-tests
Collaborative/cooperative learning	4	9—Dwikoranto et al. (2021); 17—Sulisworo et al. (2018); 20—Song et al. (2022); 21—Kuanbayeva et al. (2024)	Small-group tasks, chat or AR environments; emphasis on division of labor and shared artifacts
Blended and flipped models	4	12—Made Tegeh et al. (2020); 16—Hariadi et al. (2022); 17—Sulisworo et al. (2018); 5—Chang et al. (2020)	Videos/worksheets before class; in-class discussion or VR design studio; mobile app as hub
Student-centered, self-paced m-learning	5	4—Hasbiyati et al. (2019); 14—Putranta et al. (2021); 15—Cahyana et al. (2019); 24—Ziden et al. (2022); 7—Damopolii and Kurniadi (2019)	E-books, quizzes, or AR cards that can be accessed "anytime, anywhere"; teacher adopts facilitator role
Hands-on sensor/experiential learning	4	22—Hochberg et al. (2018); 23—Staacks et al. (2018); 1—Kousloglou et al. (2023); 5—Chang et al. (2020)	Smartphones or probes collect real-time data; results visualized on the spot; fosters analysis and argument
Augmented/virtual reality focus	4	5—Chang et al. (2020); 21—Kuanbayeva et al. (2024); 24—Ziden et al. (2022); 26—Demircioglu et al. (2023)	AR/VR scenes embed 3-D manipulatives; location-based or immersive tasks; often paired with inquiry or argumentation
Simulation-centered Five-E model	1	25—Temsah and Safa (2021)	Computer simulations mapped to Engage-Explore-Explain-Elaborate-Evaluate cycle
Competition/new-media production	1	8—Zhang et al. (2024)	Students create videos/posters, receive one-to-one feedback, publish on WeChat for real audiences

*Percentages rounded; one study can appear in more than one category.

attracted "2,400 followers and 25,000 hits" (Zhang et al., 2024), thus extending communication from peer discussion to a true public audience.

Sensor-rich smartphones and external probes enable authentic data inquiry. When learners handled a PASCO Smart Cart, they could measure "force, position, velocity, acceleration along three axes, and

TABLE 4 Target audience and participant demographic characteristics.

Educational level/age band	# studies (*n* = 27)	Typical age (yrs)	Notes on context and selection
Upper-primary (Grades 4–6/9–12 yrs)	4	10–12	UK and Norway (Blackmore and Rønningsbakk, 2023), Hong Kong (Song et al., 2022), Lebanon (Temsah and Safa, 2021), Taiwan (Chang et al., 2020)
Lower-secondary/junior-high (Grades 7–9/12–15 yrs)	6	12–15	Greece (Kousloglou et al., 2023), Turkey (Demircioglu et al., 2023), Malaysia (Ziden et al., 2022), Indonesia (Hasbiyati et al., 2019; Putri and Aznam, 2019), Kazakhstan (Kuanbayeva et al., 2024)
Upper-secondary (Grades 10–12/15–18 yrs)	9	15–18	Heavy Indonesian presence [(Agustihana and Suparno, 2018; Cahyana et al., 2019; Dasilva et al., 2019; Hariadi et al., 2022; Made Tegeh et al., 2020; Putranta et al., 2021; Sulisworo et al., 2018) plus German Gymnasium sample (Hochberg et al., 2018)]
Undergraduate/pre-service teachers	7	18–24	Turkey (Arabacioglu and Unver, 2016; Inel-Ekici and Ekici, 2022), Indonesia (Dwikoranto et al., 2021; Khery et al., 2020; Maryuningsih et al., 2019), China (Zhang et al., 2024),
Mixed/informal and museum visitors	1	Broad	Germany app demo (Staacks et al., 2018)

*Percentages rounded; one study can appear in more than one category.

TABLE 5 Distribution of study based on country and years.

Country	Years represented (*n* = studies)	Total studies
Indonesia	2018 (1) · 2019 (7) · 2020 (2) · 2021 (3) · 2022 (1)	14
Turkey	2016 (1) · 2022 (1) · 2023 (1)	3
Germany	2018 (2)	2
Hong Kong	2020 (1) · 2022 (1)	2
Greece	2023 (1)	1
Taiwan	2020 (1)	1
United Kingdom	2023 (1)	1
Norway	2023 (1)	1
Singapore	2022 (1)	1
People's Republic of China	2024 (1)	1
Kazakhstan	2024 (1)	1
Switzerland	2018 (1)	1
Malaysia	2022 (1)	1
Lebanon	2021 (1)	1

rotational velocity" in real time (Kousloglou et al., 2023). Likewise, the phyphox app turned ordinary handsets into portable laboratories, showcasing how "numerous variables … can be measured anywhere and anytime" (Staacks et al., 2018). Such activities force students to interpret noisy data, critique methods and defend conclusions—classic higher-order behaviors—while talking through findings with classmates.

Immersive spatial technologies are gaining traction but remain emergent. Augmented-reality projects like NutricARd required "learning cards with markers … for the human digestive system" (Ziden et al., 2022), while Kazakhstan's EdLab let pupils manipulate virtual voltmeters, prompting comments such as "it helped me to visualize and use the voltmeter … in a way that made [it] easier to understand and remember" (Kuanbayeva et al., 2024). A single VR study reported that hands-on design of spherical-video environments deepened geomorphology understanding (Chang et al., 2020). These tools excel at model-building and explanation, but infrastructural and training barriers keep adoption below 15% of the sample.

Assessment is shifting from end-point testing to formative analytics. By embedding Quizizz, one biology study could deliver HOTS-aligned items that "make students excited … and able to increase student independence" (Isnaeni et al., 2021). Several custom apps likewise collect usage traces so teachers can intervene in the moment rather than after the unit ends.

Geography and chronology matter. More than half of the studies were run in Indonesia, where mobile phones bridge limited desktop access and where educators note that "almost all students bring and use smartphones in high school" (Putranta et al., 2021). Chronologically, first-generation work (2016–2019) focused on single-function apps and sensor probes; newer studies (2020–2024) add AR/VR and outward-facing social-media publication, signaling a move from private classroom experimentation toward public science communication.

Taken together, the evidence suggests that the greatest learning gains come when content-specific mobile apps are tightly coupled with ubiquitous messaging or social channels, allowing students to collect authentic data, argue about it, and broadcast their explanations to real audiences—each step feeding directly into higher-order thinking and richer communication skills.

Across the 27 studies, a clear pattern emerges: mobile technologies are consistently woven into inquiry- and problem-oriented pedagogies that position students as active investigators (Table 3). Greek secondary pupils, for example, stepped through a full inquiry cycle in which they "develop questions, formulate appropriate hypotheses … analyze, understand, and explain the results of their experiments" (Kousloglou et al., 2023), while Kazakhstani learners used augmented reality to pursue "student-centered, open-ended, and inquiry-based activities aimed at promoting critical thinking, problem-solving, and knowledge construction" (Kuanbayeva et al., 2024). Even when the formal label shifts to Problem-Based Learning or Collaborative Inquiry, the recurring design is the same: a real or simulated problem anchors the lesson, mobile devices supply data or content, and students must negotiate meaning together.

A second, equally pervasive theme is scaffolding—often technological, sometimes human, usually both. Indonesia's IPMLM project pairs an Android app with teacher modeling, forming "a synergistic scaffolding system to support science learning" (Dasilva et al., 2019). In chemistry, a flipped-class M-learning environment walks students through observing, asking, gathering information, associating and communicating, because "the strategies used were flipped classroom, learning directed to follow the steps of the scientific approach" (Made Tegeh et al., 2020). Such staged supports fade as learners gain competence, echoing Vygotskian notions of assistance that is gradually withdrawn.

Third, blended and boundary-crossing implementation models turn learning into a seamless, anytime-anywhere process. The MoLearn platform exemplifies "learning that can be done virtually and without face-to-face unlike classroom learning" (Hariadi et al., 2022), while cooperative-blended studies remind us that "some interactions between teacher and students or among students still need to be undertaken in class" (Sulisworo et al., 2018). Personal devices extend these interactions far beyond school walls: university students appreciate that they can "conduct research anywhere and anytime and can easily complete out-of-school inquiry tasks" (Inel-Ekici and Ekici, 2022), and secondary students note how a simple Viber group "promoted cooperation and communication between them" (Kousloglou et al., 2023).

Underlying all this is a shift toward student-centered, socially mediated learning in authentic contexts. Teachers increasingly act as facilitators—"teacher as facilitator, student as learning center" (Hariadi et al., 2022)—providing just-in-time guidance such as "one-on-one guidance on the works and suggested revisions" during a science-communication competition (Zhang et al., 2024). Successful designs put problems in everyday language—"the context on the article should be close to student daily life" (Khery et al., 2020)—and move learners into "authentic, real-life situation[s] apart from the classroom and independent of time" (Arabacioglu and Unver, 2016). In short, the most effective mobile-enhanced pedagogies fuse structured inquiry with graduated support, blend physical and virtual spaces, and cultivate collaborative meaning-making around problems that resonate with students' lives (Table 6).

HOTS grow in problem-rich inquiry spaces (Table 6). From Greek physics lessons where pupils "develop questions, formulate appropriate hypotheses … and communicate their findings" (Kousloglou et al., 2023) to Kazakh students who tackle open-ended AR lab tasks "aimed at promoting critical thinking, problem-solving, and knowledge construction" (Kuanbayeva et al., 2024), the dominant recipe is to anchor learning in a puzzle that cannot be resolved with recall alone. Whether the format is inquiry-based, problem-based or case-based, the common denominator is that learners must identify variables, sift evidence and justify choices—directly exercising the Analyze-Evaluate-Create trio at the apex of Bloom.

Strategically layered scaffolds keep cognitive load productive. Android IPMLM combines teacher modeling with in-app hints so that "the teacher gives a problem from the simplest … to a complex problem" (Dasilva et al., 2019). In blended statistics, MoLearn's five-phase script steers novices through investigation, analysis and presentation while still positioning "teacher as facilitator, student as learning center" (Hariadi et al., 2022). The support is explicit early on, then gradually fades—echoing Vygotsky's zone and ensuring that mental effort is spent on meaning-making rather than floundering.

Collaboration is more than feel-good rhetoric—it is engineered for argument quality. AR-based sky observation asked pupils to weigh competing theories; every team "engaged in the argumentation and produced quality arguments" (Demircioglu et al., 2023). Primary learners in Hong Kong's m-Orchestrate app cycled through WeEngage–WeReflect stages that map group moves against a "Matrix of collaborative problem-solving skills" (Song et al., 2022). Such structures channel talk toward evidence, claims and rebuttals—corner-stones of critical and creative thinking.

Metacognitive discipline turns inquiry into learning-how-to-learn. After a 12-week sensor-rich sequence, nine out of ten Greek students "reported pausing … to reflect" versus three before (Kousloglou et al., 2023). Biology students using mobile modules practiced metacognition "in each learning process" so frequently that skill growth was described as continuous (Damopolii and Kurniadi, 2019). Reflection prompts, self-checks and flipped pre-readings all cultivate the Evaluate tier of Bloom—thinking about one's own thinking.

Doing science with devices tightens the loop from observation to inference. Hands-on VR designers in Taiwan had to understand geomorphology deeply enough to build 360° tours, which "cultivated their problem-solving and metacognitive skills" (Chang et al., 2020). Turkish preservice teachers, armed with pH and conductivity probes, pinpointed variables impossible to isolate by eye, thereby sharpening hypothesis formation (Arabacioglu and Unver, 2016). Real-time data streamed to screens lets students analyze, evaluate and iterate on the spot.

Rich visual and multimedia scaffolds ignite creation and originality. Focusky's zooming canvases let Indonesian pupils "see, hear, talk, and write" the science—an experience estimated to yield 70% retention (Putri and Aznam, 2019). A smartphone e-book's video-augmented pages boosted the otherwise elusive originality strand of creativity by 14% across action-research cycles (Hasbiyati et al., 2019). Such multimodal inputs push learners beyond rote to synthesize and invent.

Feedback-rich assessment locks gains into place. ICT-PBL biology embedded Quizizz HOTS items so students could "complete tests more independently" while teachers saw instant analytics (Isnaeni et al., 2021). The BWML model capped each phase with tailored evaluation instruments "made about the high-level cognitive domains" (Hariadi et al., 2022), ensuring that Analyze/Evaluate/Create are not optional extras but measured outcomes (Table 7).

Mobile-first science lessons nurture communication in several, mutually reinforcing ways (Table 7). To begin with, persistent chat spaces such as Viber and WhatsApp let talk spill past the bell. Greek pupils reported that their class Viber group became "an extra channel for information, communication, and cooperation… outside the classroom environment" (Kousloglou et al., 2023), a setting where they practiced encouragement, active listening and respectful critique—sub-skills the study lists under the 4 Cs (p. 11). Because the medium is familiar and always on, learners feel safe asking questions and offering quick feedback, behaviors that textbook lessons rarely provoke.

A second thread is the shift from consuming to producing scientific messages. When Taiwanese students designed 360° VR field guides they had to phrase explanations so that others could follow the tour; this hands-on process meant they could "interact with peers and discuss with the teacher" about clarity and accuracy (Chang et al.,

TABLE 6 HOTS development strategies.

Strategy family (not mutually exclusive)	# studies*	Representative mechanisms	Typical bloom levels
Inquiry/problem framing (open, guided, or case-based)	12	Full inquiry cycles, hands-on experiment design, case observation, hypothesis testing	Analyze → Evaluate → Create
Scaffolding and fading (teacher prompts, in-app hints, graduated tasks)	7	Prompt-and-probe questioning, escalating problem complexity, worked examples that disappear over time	Analyze → Evaluate
Collaborative argumentation and CPS	6	Small-group debate, evidence-based claims, peer feedback, matrix of collaborative problem-solving behaviors	Analyze → Evaluate → Create
Metacognition and reflection	5	Forced reflection pauses, self-monitoring checklists, flipped "read-reflect-recite-review" loops	Evaluate
Hands-on data generation (sensors, VR authoring, AR manipulation)	6	Smartphone probes, design-your-own VR tours, AR manipulation of circuits or sky models	Apply → Analyze → Create
Multimedia visualization (3-D, simulation, interactive e-books)	7	3-D digestive system, water-quality dashboards, Focusky slides with tiered quiz levels	Understand → Apply → Analyze
Assessment and feedback loops	5	Quizizz HOTS item banks, MoLearn analytics, custom tests that target C4-C6 only	Analyze → Evaluate → Create

*Percentages rounded; one study can appear in more than one category.

TABLE 7 Communication skills development approaches.

Strategy family	Share*	Mobile affordance that drives it	Typical communication moves†	Collaborative layer
Always-on messaging and chat (WhatsApp, Viber, in-app chat)	7	Persistent group spaces outside class	Written Q&A, peer feedback, teacher micro-scaffolds	Small-group or whole-class threads (Kousloglou et al., 2023, p. 6; Isnaeni et al., 2021)
Shared artifact creation (VR builds, posters, videos, cloud docs)	6	Authoring tools + cloud publishing	Multimodal explanation, visual literacy, audience-aware writing	Co-design and peer review (Chang et al., 2020, p. 921; Zhang et al., 2024, p. 289)
Inquiry talk around live data (sensors, SPARKvue screen-share, AR views)	5	Real-time display or shared AR scene	Oral explanation of findings, argumentation, critique	Face-to-face and remote comparisons (Kousloglou et al., 2023, p. 5; Demircioglu et al., 2023, p. 1,182)
Structured discussion forums / reports	6	Forum posts, report uploads, KWLH charts	Formal scientific writing, poster/ report genre	Peer commenting, teacher feedback (Maryuningsih et al., 2019; Temsah and Safa, 2021, p. 72)
Analytics-backed CPS apps (m-Orchestrate, MoLearn)	3	Chat room + analytics dashboard	Task regulation dialog, reflection	Heterogenous teams with role division (Song et al., 2022, p. 3, 4; Haryadi and Pujiastuti, 2022, p. 574)
AR/VR shared spaces	4	Co-located virtual objects, multiplayer AR	Joint manipulation talk, negotiated meaning	Team exploration and dispute resolution (Kuanbayeva et al., 2024, p. 154; Ziden et al., 2022, p. 2)
Interactive e-books/ multimedia slides	3	Tap-to-media, annotation	Question-asking, explanatory talk	Informal peer sharing (Hasbiyati et al., 2019, p. 6, 7; Putri and Aznam, 2019, p. 14)

*n = 21 studies that gave explicit evidence on communication; one study can sit in several rows. †Moves align with the P21 "4 Cs" taxonomy (expressing ideas, active listening, adaptive delivery, etc.).

2020). Likewise, Chinese medical undergraduates published posters and short videos on WeChat, finding that "New Media effectively expand the influence of students' popular science works" (Zhang et al., 2024). Preparing work for a real audience raises the communicative stakes and pushes students to refine visual layout, narrative flow and tone.

Equally powerful are shared visual anchors that focus discussion on evidence. In the Greek inquiry sequence, sensor data were mirrored to every tablet so classmates could debate sources of error and variable control (Kousloglou et al., 2023). Turkish eight-graders, peering through AR constellations, gathered "rich data to support students' arguments" (Demircioglu et al., 2023); the common image gave their arguments precision and kept talk tethered to observable phenomena.

Formal writing also benefits from digital structure. Lebanese elementary pupils filled KWLH charts and wrote short lab reports; teachers noticed that such tasks "reinforced their written communication skills" by demanding clear tables, graphs and sentences (Temsah and Safa, 2021). Chemistry majors using the

Mobile-NOS model prepared worksheets, then defended them aloud, an exercise that "stressing intellectual and communication skill development" (Khery et al., 2020) and pushed students to switch smoothly between written and spoken registers.

Some platforms go further, making the process of collaboration visible. The m-Orchestrate app logged every chat, note and upload against a matrix of collaborative problem-solving moves, helping teachers and peers see who was building common ground and who needed nudges (Song et al., 2022). Such analytics-driven feedback teaches students to regulate turns, distribute tasks and keep dialog purposeful.

Finally, extended-reality spaces intensify the need for clear negotiation. Kazakh students wiring virtual circuits in a shared AR "tab" admitted it felt odd at first, yet concluded that "collaboration helped me to develop my communication skills" (Kuanbayeva et al., 2024). The technology forced them to verbalize what they saw, resolve disputes and integrate ideas—a microcosm of the argumentative give-and-take scientists engage in daily.

Taken together, these studies show that mobile technologies work best for communication when they create continuous channels, anchor talk in tangible artifacts, and provide structured spaces—whether analytic dashboards or AR scenes—where learners must articulate, listen and revise in real time.

Most studies in the corpus converge on five demographic patterns that shape how mobile-science interventions are conceived and interpreted (Table 4). First, they overwhelmingly gravitate toward secondary-school learners (ages 12–18). Projects from Indonesia, Turkey, Malaysia and Greece treat high-school classrooms as proving grounds for 21-century competencies, often citing their national curricula's explicit HOTS mandates; this focus reflects a belief that teenagers possess both the cognitive maturity and the smartphone ubiquity needed to exploit advanced apps and sensors. Second, a smaller but important cluster works with upper-primary pupils (about 10–12 years old), arguing that early exposure to tablets, VR or analytics-rich inquiry builds STEM identity before attitudes harden; here, one-to-one iPad programs in the UK or GPS-tagged fieldwork in Hong Kong illustrate how mobile tools can be scaffolded for younger hands. Third, universities appear chiefly through pre-service teacher cohorts, where digital-native undergraduates rehearse the very pedagogies they are expected to deploy in future classrooms; this dual-purpose sampling (content learning plus method modeling) is evident in Turkish and Indonesian teacher-education faculties that integrate AR labs or mobile inquiry apps. Fourth, the evidence base is geographically skewed toward Indonesia, which supplies more than half of the studies; while this offers rich insight into an Android-heavy, WhatsApp-saturated context, it also cautions against uncritical transfer to regions with different infrastructure or policy climates. Finally, authors repeatedly emphasize that cultural and institutional context mediates participation—whether it is Greek researchers selecting high-achieving science-club volunteers already "proficient with their smartphones," Chinese teams leveraging WeChat because it dominates daily communication, or a Kazakh lyceum chosen for its elite digital facilities. Together, these patterns remind reviewers that the "who" of a study—age, schooling track, socio-technical environment—deeply conditions both the feasibility of mobile tools and the meanings students make from them.

Across the 27 studies, six methodological patterns stand out (Table 1). First, quasi-experimental, pre-/post-test arrangements dominate: from Chang et al. (2020)'s VR trial that contrasted a "hands-on design VR system" with a guided version to Hariadi's blended-learning work that used matched control classes, most teams sought measurable before-and-after gains. Second, researchers often embed these experiments within larger development cycles (such as the 4-D model, Borg and Gall's 11-step sequence, or Luther's multimedia pipeline) so that prototyping, expert validation, and classroom piloting occur in a single continuum. Third, mixed or multi-method data regimes are now routine: cognitive tests and Likert surveys are almost always paired with open-ended artifacts (posters, VR projects), log files [process-mining in Song et al., 2022], observations or interview transcripts analyzed in NVivo or via thematic coding. Fourth, statistical treatment tends to be basic but consistent—n-gain scores, paired t-tests, ANCOVA, two-way ANOVA, regression—while qualitative material is handled through content or thematic analysis, Toulmin argument mapping or directed coding; only a handful of cases [e.g., (Song et al., 2022)] push into advanced analytics such as Markov-chain modeling. Fifth, sample size and timelines vary wildly: pilots with a single science club of 10 Greek ninth-graders run beside Indonesian field-tests with 1,070 pupils; yet the modal study still involves one or two intact classes (≈30–70 learners) over 4 to 8 weeks, enough for a teaching unit but short of long-term follow-up. Finally, purposeful or convenience sampling prevails and instruments are almost always psychometrically checked—authors report Cronbach's α between 0.68 and 0.90, inter-rater reliabilities around 0.9 and item-response diagnostics—reflecting a shared concern for internal validity even when external generalization remains limited. Together these patterns reveal a community that balances pragmatic classroom experimentation with growing methodological sophistication, yet still wrestles with small samples, short durations and context-bound designs.

The extreme differences in study duration and data collection approaches revealed in Table 1 significantly undermine the validity of cross-study comparisons and meta-inferences. Duration inconsistencies create a fundamental interpretation problem: studies claiming significant HOTS development after a minimum exposure duration (e.g., Kuanbayeva et al., 2024: 4 × 45 min sessions; Temsah and Safa, 2021: 3-week intervention) cannot be meaningfully compared to longer-term interventions (Blackmore and Rønningsbakk, 2023: 3 academic years; Kousloglou et al., 2023: 5-month timeline). This temporal heterogeneity suggests that while many "positive outcomes" may reflect novelty effects, they may not reflect sustained cognitive development, especially given that HOTS formation often requires long-term implementation and support.

Variation in evaluation timing: further complicates interpretation. The prevalence of post-test designs without follow-up measures (18/27 studies) means that retention of learning, an important indicator of true HOTS development, is largely unknown. Studies that administer immediate tests with single-lesson interventions (e.g., Khery et al., 2020: "single lesson"; Ziden et al., 2022: duration unspecified) may capture transient performance gains rather than long-term higher-order thinking skills. Conversely, the few studies with longer observation periods (Blackmore and Rønningsbakk, 2023: 40 h of observation over 3 years) show more nuanced and developmentally appropriate learning curves.

TABLE 8 Challenges and suggestions.

Thematic cluster	How often it appeared*	Representative studies (ID)	Typical remedies suggested by authors
1. Technical and infrastructure hurdles AR/VR tracking glitches, slow apps, connectivity, marker printing, sensor limitations	9 of 27 studies	5, 21, 23, 24, 26 (also 10, 11, 20, 22 for minor reports)	Pre-lesson device training and download sessions • Remote-view/in-app analysis tools (phyphox) • Up-front printing / sharing of AR markers • Upgrading or selecting more stable platforms
2. Teacher capacity and pedagogical design Limited TPACK, media-design skills, low school-policy support	11 studies	2, 10, 11, 12, 17, 25, 26 (plus 3, 16, 18, 23)	Professional-development workshops Ready-made lesson packets and apps with embedded scaffolds Blended models that keep some face-to-face time Policy alignment and administrative backing
3. Time and workload constraints Short lessons, crowded curricula, long prep for inquiry/ VR design	7 studies	5, 11, 13, 1, 21, 26, 20	Extend activities over several sessions or homework Streamline or trim number of tasks Provide ready-to-use templates and scaffolding
4. Learner-related factors Smartphone misuse, low motivation, unequal prior skills, ceiling effects	8 studies	14, 15, 7, 1, 5, 9, 24, 25	Strong supervision and clear task structure Motivation-sensitive pedagogy (e.g., PowerPoint for low-motivation learners) Embedding activities into assessment to raise stakes
5. Research-design limitations Small samples, single-case settings, limited generalizability	6 studies	1, 5, 19, 21, 22, 24	Plan replications with larger cohorts or multiple sites Combine quantitative and qualitative evidence
6. Assessment and instrument issues Misinterpreted questionnaire items, difficulty capturing certain skills (e.g., graph-drawing)	4 studies	1, 3, 25, 16	Pilot and refine items, add step-by-step scaffolds, triangulate with product analysis

The variability of sample size (ranging from n = 10 to n = 1,070) raises additional validity concerns, as smaller studies consistently report larger effect sizes. This is a pattern suggestive of selection bias or insufficient statistical power in larger and more rigorous applications. "The predominance of convenience sampling from "intact classrooms" (seen in 22/27 studies) further limits generalizability, as these samples may not represent typical technology adoption or pedagogical integration scenarios. In sum, these methodological differences suggest that the apparent consensus on the effectiveness of mobile learning may be an artificial result reflecting methodological heterogeneity rather than reflecting actual pedagogical impact (Table 8).

The overview of challenges shows that technical friction—lagging apps, marker-scanning failures, unreliable Wi-Fi—remains the most visible stumbling block, yet researchers rarely abandon the tools; instead, they insulate learning with pragmatic fixes such as pre-class download sessions, printed QR codes or remote-display features. Equally prominent are human factors: teacher readiness and pedagogical design repeatedly limit impact, underscoring that hardware alone cannot guarantee higher-order learning without sustained professional development and turnkey lesson materials. Time pressure also cuts across contexts—from 40-min science periods in Taiwan to crowded Indonesian curricula—so the most successful studies redistribute workload through flipped-classroom homework or by trimming inquiry tasks. Learner motivation and habits further modulate outcomes: smartphones can amplify both productive engagement and off-task behavior, which means clear supervision, assessment links and intrinsically meaningful tasks (like competitions or real-world science-communication) are vital. Methodologically, many projects acknowledge "micro-level" samples or single-school settings, signaling a need for multi-site replications and mixed-method designs to strengthen generalizability. Finally, assessment itself is a subtle bottleneck; when questionnaire items are misread or skills like graph-drawing go unmeasured, reported gains lose clarity, prompting calls for iterative instrument refinement and triangulation with artifact analysis. Together, the evidence suggests that mobile-enhanced science learning flourishes only when technical reliability, teacher competence, realistic time allocation and carefully aligned assessments converge; neglecting any of these pillars breeds the very challenges cataloged across the literature.

Discussion

The evidence base confirms that mobile learning in science can raise higher-order thinking and communication. Key results are followings. Evidence is geographically skewed: Indonesia supplies 14 of the 27 studies, with Turkey, Germany, and Hong Kong forming a distant second tier; many regions remain under-represented. Custom science apps dominate the technology mix (41% of cases), followed by social-communication platforms (26%) and sensor-based probes (15%), while AR/VR projects are rising but still account for only about 15% of interventions. Integrations that stream app data directly into familiar messaging spaces such as WhatsApp, Viber, WeChat yield the largest learning jumps because students can debate evidence anytime. Publishing student products to real audiences (e.g., WeChat science posts with thousands of external views) extends communication gains beyond the classroom. Methodologically, quasi-experimental pre/post designs with validated instruments prevail, but samples are typically one or two intact classes over four–eight weeks, limiting generalizability and masking long-term effects. Common obstacles include technical glitches, limited bandwidth, and time-pressured curricula; studies that added pragmatic fixes (pre-class downloads, printed QR codes) and clear task structures mitigated these barriers.

The pronounced Indonesian dominance in our evidence base (52% of studies) creates a complex interpretive challenge that extends beyond simple geographic bias concerns. While the sheer volume of Indonesian research initially appears to strengthen evidence quality, deeper analysis reveals that reported effectiveness may reflect optimal cultural-technological alignment rather than universally applicable pedagogical principles. Indonesian students' pre-existing familiarity with WhatsApp group dynamics, combined with collectivist learning orientations and high teacher authority acceptance, creates uniquely favorable conditions for mobile collaborative learning that may not replicate in individualistic cultures with stronger teacher-student boundary maintenance (Setyaningrum et al., 2022). The theoretical frameworks underlying successful interventions that particularly Social Constructivist learning through persistent peer dialog and Activity Theory's mediated collaboration remain valid, but their technological instantiation appears heavily mediated by cultural communication norms and infrastructure constraints. This suggests that mobile learning effectiveness operates through theoretically universal mechanisms deployed via culturally specific technological configurations, requiring significant adaptation rather than direct transfer across educational contexts.

The evidence reveals that mobile technologies facilitate communication skill development through specific theoretical pathways that transcend simple "technology enhancement" narratives. Multimodal Communication Theory (Buck and VanLear, 2002) provides the most robust explanatory framework for understanding why mobile-integrated science lessons consistently outperform traditional approaches in developing both verbal and written scientific communication. Students creating AR-enhanced presentations (Demircioglu et al., 2023) or publishing science videos on WeChat (Zhang et al., 2024) engage multiple semiotic systems simultaneously such as visual, spatial, linguistic, and gestural developing digital rhetoric competencies essential for contemporary scientific discourse while strengthening traditional communication skills. The recurrent pattern of real-time data collection followed by social sharing and collaborative interpretation reflects Computer-Supported Collaborative Learning principles, where mobile platforms enable cognitive load distribution and collective knowledge building through persistent interaction histories (Kaliisa et al., 2025). However, the effectiveness of these mechanisms appears contingent on Social Presence Theory conditions (Weidlich et al., 2023) that students must perceive their mobile interactions as socially meaningful rather than technologically imposed, explaining why familiar platforms like WhatsApp consistently outperform purpose-built educational apps in communication skill development outcomes.

These findings mirror earlier syntheses. Afikah et al. (2022) and Ansori et al. (2024) both concluded that collaborative inquiry and project-based mobile lessons outperform conventional teaching on higher-order skills, and our updated set agrees. A meta-analysis of mobile-integrated education found an overall moderate mean effect size of 0.523, suggesting that mobile devices significantly enhance learning outcomes compared to traditional approaches across various educational contexts (Sung et al., 2016). Together these external reviews confirm that the trend in science is consistent with the broader literature.

Context continues to shape outcomes. More than half of the studies were run in Indonesia, where affordable Android phones and the near-universal use of WhatsApp lower entry barriers. Results in Turkey, Germany, and Hong Kong were similar in direction but sometimes smaller, suggesting that policy rules, bandwidth, and classroom routines constrain scale. The dominance of Indonesia may be due to the emphasis placed on HOTS in the Indonesian curriculum, which is why researchers give priority to it (Zana et al., 2024). A multi-sector review of success factors for mobile learning shows that portability and context awareness are key mediators. Teacher readiness is equally important (Hamzah et al., 2022).

The integration of mobile learning within broader e-learning ecosystems reflects a fundamental shift in how higher-order cognitive processes are developed in digitally-mediated educational contexts. Recent phenomenological studies reveal that e-learning environments in higher education provide unique affordances for fostering complex thinking skills through collaborative knowledge construction and reflective practice (Khadka et al., 2025). This digital transformation particularly benefits technical-engineering disciplines, where faculty members increasingly recognize mobile technologies' potential for enhancing student engagement and learning outcomes through authentic, context-aware applications (Mohammadi et al., 2020).

Mobile-science studies show a clear tilt toward inquiry-centered, problem-oriented teaching. This pedagogical orientation aligns with broader evidence demonstrating that STEM education effectiveness depends critically on active learning approaches that engage students in authentic problem-solving and collaborative inquiry (Wahono et al., 2020). Project-based learning models, particularly when enhanced through mobile technologies, have shown particular promise in developing pre-service teachers' higher-order thinking capabilities, suggesting that pedagogical innovation in teacher education can create ripple effects throughout educational systems (Haryadi and Pujiastuti, 2022). About a third of the corpus adopts full inquiry cycles in which learners pose questions, gather real-time data with sensors or augmented reality, and defend explanations; these designs consistently sit at the top of the learning-gain table. Problem-based lessons, guided instruction with fading hints, collaborative workspaces, and blended or flipped formats appear less often yet draw on the same idea: mobile tools make data and dialog available everywhere, so classrooms can shift from content delivery to shared investigation (Bidarra and Rusman, 2017). Success depends on coupling purpose-built science apps with everyday chat or cloud platforms; when every reading and graph streams into a WhatsApp or Viber group, discussion stays alive after class and learning gains deepen (Isnaeni et al., 2021; Kousloglou et al., 2023; Liu and Zhang, 2022; Zhang et al., 2024).

Graduated scaffolding emerges as the second pillar of effective implementation. Projects like Indonesia's IPMLM start with teacher modeling and in-app hints, then taper support as competence grows, mirroring Vygotsky's zone of proximal development (Dasilva et al., 2019). Blended models such as MoLearn push exposition into homework videos so that class time can focus on analysis or VR design; students report that this anytime-anywhere flow lets them complete inquiry tasks that once seemed impossible within a single period (Hariadi et al., 2022). Even in resource-tight contexts, pragmatic fixes—pre-lesson downloads, printed AR markers, remote-view functions—protect lesson flow when bandwidth falters, showing that implementation quality turns on both pedagogy and logistics (Ziden et al., 2022).

Higher-order thinking grows most when those pedagogical moves are paired with explicit HOTS strategies (Sun et al., 2022; Supeno et al., 2019). Research consistently demonstrates that HOTS development in science and STEM education contexts requires systematic pedagogical approaches that integrate technology with evidence-based teaching methods (Haryadi and Pujiastuti, 2022; Wahono et al., 2020), while innovative assessment strategies such as interactive questioning techniques can significantly enhance students' analytical capabilities in biology and related sciences (Wahyuni et al., 2022). Inquiry or problem framing dominates, obliging learners to identify variables, evidence, and justify claims—activities that map directly onto analyze, evaluate, and create in the revised Bloom taxonomy (Anderson and Krathwohl, 2001). Layered scaffolds keep cognitive load productive: escalating problem complexity, worked examples that fade, and teacher prompts released just in time all steer mental effort toward meaning rather than confusion. Collaborative argumentation adds social pressure to reason well; Toulmin coding from AR sky-observation lessons and discourse matrices from primary apps both show richer claim-and-rebuttal chains when chat or shared artifacts channel talk around common evidence (Demircioglu et al., 2023).

Metacognitive reflection reinforces these gains. Forced pauses, self-checklists, and flipped "read-reflect-recite-review" loops lead students to monitor their thinking; after a 12-week sensor sequence in Greece, 9 out of 10 students reported halting to reflect, triple the baseline rate. Authentic data generation through smartphone probes or VR authoring tightens the observation-to-inference loop, compelling learners to interpret noisy signals and iterate hypotheses on the fly—behaviors linked to large jumps in creativity and critical-thinking scores (Kousloglou et al., 2023). Finally, analytics-rich assessments like Quizizz HOTS banks and model-aligned rubrics lock progress into place by giving immediate feedback on analyze/evaluate/create tasks rather than end-point recall (Isnaeni et al., 2021).

Together these findings suggest that mobile learning lifts higher-order thinking when inquiry-based, socially mediated pedagogy meets layered scaffolding and data-rich tasks (Vilela et al., 2025). Implementation succeeds not because of devices alone but because teachers orchestrate apps, chat, and prompts into an environment where every action—collecting data, arguing in a thread, pausing to reflect—feeds directly into HOTS development (Ansori et al., 2024; Hasbiyati et al., 2019). Future work should test these combined designs over full semesters and across diverse regions to see whether the observed gains endure once novelty fades and curricula tighten.

Methodological quality is improving but still limited. Most studies use quasi-experimental pre-post designs with validated instruments and triangulated interviews, which gives moderate internal validity. Sample sizes stay small at one or two intact classes and intervention windows remain short at 4 to 8 weeks, so long-term retention and transfer are unknown. Only a few studies (Blackmore and Rønningsbakk, 2023; Demircioglu et al., 2023) apply advanced analytics such as process mining or discourse coding.

The pronounced methodological heterogeneity documented in our analysis reveals a fundamental maturation challenge facing mobile learning research in science education. The field appears caught between proof-of-concept enthusiasm and rigorous effectiveness evaluation, with most studies positioned in an exploratory phase that prioritizes demonstrating feasibility over establishing efficacy. This developmental stage, while understandable given mobile technology's relative novelty in educational contexts, creates a false consensus problem where accumulating positive findings may reflect methodological convenience rather than genuine pedagogical breakthrough (Sung et al., 2016). The inverse relationship between intervention duration and reported effect sizes particularly suggests that researchers may be inadvertently capturing technology novelty effects that students' temporary engagement with unfamiliar tools rather than the sustained cognitive restructuring that authentic HOTS development requires (Anderson and Krathwohl, 2001).

For educational practitioners, this heterogeneity creates a translation dilemma: which study conditions most closely approximate their implementation context, and how should duration and assessment differences inform their expectations? The predominance of short-term, small-sample studies means that teachers attempting semester-long implementations lack evidence-based guidance for managing the post-novelty phases when initial enthusiasm wanes. Moreover, the assessment timing concentration around immediate post-tests provides little insight into whether mobile-supported HOTS gains transfer to standardized examinations, long-term projects, or novel problem-solving contexts—the ultimate measures of educational value. This evidence gap is particularly problematic given that mobile learning interventions often require substantial resource investment, teacher training, and curriculum restructuring that can only be justified through demonstrated durability of learning outcomes. Future research must therefore prioritize methodological standardization and extended evaluation windows to provide practitioners with actionable evidence rather than promising but potentially ephemeral proof-of-concept demonstrations.

Future research should move beyond pilot mode. Semester-long trials woven into normal timetables will allow us to see whether the gains last. Multi-site collaborations can test scalability outside the current hotspots. Researchers also need to compare augmented and virtual reality against simpler sensor-based tasks to see whether the extra complexity pays off. Richer analytics that combine log traces, discourse analysis, and delayed follow-ups will clarify how mobile interactions translate into durable higher-order thinking and authentic scientific communication.

Conclusion

This scoping review examined 27 empirical studies that used mobile tools in science lessons between 2016 and 2024, charting where, when and how the technology was applied. Most of the evidence originates in Indonesia, which supplies 14 of the 27 studies, while Turkey, Germany and Hong Kong contribute only two or three each; publication activity accelerates after 2018, mirroring the post-pandemic surge in interest. Custom science apps are the most common technology (41%), followed by chat or social-media platforms (26%), on-board sensors and external probes (15%) and a smaller but growing share of augmented- and virtual-reality projects (about 11 and 4% respectively).

Inquiry-based lessons dominate the pedagogical landscape, with problem-based, collaborative, blended and self-paced designs filling out the mix. Across these variants three design moves recur: real-time data collection, instant sharing through chat or cloud slides and scaffolded prompts that steer reflection. When purpose-built science apps are tightly coupled with familiar messaging spaces, students gather authentic evidence, debate interpretations and publish explanations to real audiences—an implementation pattern that consistently strengthens higher-order thinking and communication.

Limits temper these positives. Many interventions involve just one or two intact classes and run for fewer than 8 weeks; very few track retention or transfer beyond the immediate unit. The geographic skew toward Indonesia and the scarce use of advanced analytics mean that results may not generalize to regions with different infrastructure or policy climates.

In sum, mobile learning raises analysis, evaluation, creativity and scientific expression when inquiry-rich tasks, social sharing and graduated scaffolds act together. Future research should extend trials over full semesters, diversify study sites, invest in teacher professional development and gather richer process data so that the field can see how mobile interactions translate into lasting, transferable learning gains.

The evidence assembled for this review is robust enough to reveal patterns, yet several boundaries limit the strength and reach of its conclusions. Most studies were conducted in Indonesia, leaving large swaths of Africa, South America and much of Europe unrepresented; this geographic skew means that findings may not travel unchanged to settings with different infrastructure, policy, or cultural expectations. Samples were usually one or two intact classes, interventions seldom exceeded 8 weeks, and advanced analytics were rare, so long-term retention, transfer and nuanced process mechanisms remain largely unknown. Technical friction—lagging apps, sensor glitches and patchy connectivity—regularly disrupted lessons, while teacher capacity, crowded timetables and uneven learner motivation further constrained impact; in a third of the corpus, assessment tools themselves failed to capture the targeted skills accurately.

Several practical and research-oriented steps could address these constraints. First, future work should plan semester- or year-long interventions that run in everyday timetables across multiple schools and regions; such designs would test durability, scalability and cultural fit while generating richer data for process analytics. Second, replicating promising models outside Indonesia and pairing them with context-sensitive adjustments would reduce geographic bias and reveal how infrastructure or policy mediates outcomes. Third, pragmatic fixes already noted by many authors—pre-lesson device training, remote-view functions, printed AR markers and stronger Wi-Fi—should become standard preparation so that technical issues no longer steal cognitive time from inquiry. Finally, sustained professional development and turnkey lesson packets can lift teacher TPACK and lighten design workload, while iterative instrument refinement and triangulation with product analysis will sharpen measurement of higher-order thinking and communication. Taken together, these recommendations aim to widen the study base, deepen methodological rigor and ensure that mobile science learning delivers durable, transferable gains for diverse learners.

Mobile learning in science education stands at a critical juncture where technological affordances increasingly align with pedagogical understanding of how students develop complex thinking and communication competencies. The evidence assembled in this review, despite its methodological limitations and geographic concentration, points toward a fundamental shift in how scientific inquiry and discourse can be supported through ubiquitous, socially-connected technologies. The consistent emergence of collaborative, inquiry-based approaches across diverse cultural contexts suggests that mobile technologies may serve as universal catalysts for active learning when thoughtfully integrated with established pedagogical frameworks. However, realizing this transformative potential requires moving beyond the current phase of proof-of-concept demonstrations toward systematically addressing the implementation gaps identified in this review: developing culturally-sensitive adaptation protocols, establishing robust communication assessment frameworks, and conducting extended-duration trials that capture authentic learning progression rather than novelty effects. The stakes extend far beyond classroom innovation—as global challenges increasingly require scientifically literate citizens capable of evidence-based reasoning and effective science communication, the educational community cannot afford to let methodological shortcuts undermine what may be one of the most promising pathways for democratizing access to high-quality science education.

The convergence of global smartphone penetration, educational policy emphasis on 21st-century skills, and growing recognition of science communication's societal importance creates an unprecedented opportunity for mobile learning to address persistent inequities in science education quality. The Indonesian research concentration documented in this review, while limiting immediate generalizability, demonstrates how resource-constrained contexts can leverage mobile technologies to achieve learning outcomes historically associated with well-funded laboratory environments. This suggests that thoughtfully implemented mobile science learning could help bridge the global digital divide in education—but only if the international research community commits to the rigorous, culturally-responsive methodology needed to distinguish genuine pedagogical innovation from technological novelty. The path forward requires unprecedented collaboration between researchers, practitioners, and policymakers to establish evidence standards that honor both methodological rigor and cultural diversity, ensuring that mobile learning research serves its ultimate purpose: enhancing every student's capacity to think critically about scientific phenomena and communicate effectively about evidence-based solutions to humanity's most pressing challenges. The evidence base reviewed here provides a foundation, but the transformative potential of mobile learning in science education will only be realized through sustained commitment to research excellence, cultural humility, and unwavering focus on equitable student outcomes across all educational contexts.

Author contributions

EG: Writing – review & editing, Writing – original draft. OS: Writing – review & editing, Writing – original draft. MZ: Writing – review & editing, Writing – original draft. NS: Writing – original draft, Writing – review & editing. VZ: Writing – review & editing, Writing – original draft. ND: Writing – review & editing, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Affandy, H., Sunarno, W., Suryana, R., and Harjana, J. (2024). Integrating creative pedagogy into problem-based learning: the effects on higher order thinking skills in science education. *Think. Skills Creat.* 53:1575. doi: 10.1016/j.tsc.2024.101575

Afikah, A., Astuti, S. R. D., Suyanta, S., Jumadi, J., and Rohaeti, E. (2022). Mobile learning in science education to improve higher-order thinking skills (HOTS) and communication skills: a systematic review. *Int. J. Adv. Comput. Sci. Appl.* 13:782. doi: 10.14569/IJACSA.2022.0130782

Agustihana, S., and Suparno, R. (2018). Effectiveness of physics Mobile learning media to improve higher order thinking skills of students in thermodynamics. *J. Phys. Conf. Ser.* 1097:2031. doi: 10.1088/1742-6596/1097/1/012031

Ahmad, M., Mansor, N. R., Sung, C. M., Rashid, R. A., Abdullah, N. A. C., Zakaria, R., et al. (2020). Mobile Technology in Enhancing Students' higher order thinking skill. *J. Phys. Conf. Ser.* 1529:2057. doi: 10.1088/1742-6596/1529/4/042057

Anderson, L. W., and Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives: Complete edition. New York: Addison Wesley Longman, Inc.

Ansori, I., Arianto, F., and Khotimah, K. (2024). The role of Mobile learning in facilitating science learning to enhance higher order thinking skills (HOTS): a systematic review. *Indon J Instruct Media Model* 6, 79–88. doi: 10.32585/ijimm.v6i2.5580

Arabacioglu, S., and Unver, A. O. (2016). Supporting inquiry based laboratory practices with mobile learning to enhance students' process skills in science education. *J. Balt. Sci. Educ.* 15, 216–231. doi: 10.33225/jbse/16.15.216

Arksey, H., and O'Malley, L. (2005). Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol.* 8, 19–32. doi: 10.1080/1364557032000119616

Bidarra, J., and Rusman, E. (2017). Towards a pedagogical model for science education: bridging educational contexts through a blended learning approach. *Open Learn.* 32, 6–20. doi: 10.1080/02680513.2016.1265442

Blackmore, K., and Rønningsbakk, L. (2023). Let us explain everything: pupils' perspectives of the affordances of mobile technology during primary science inquiry. *Frontiers in Education* 8:8459. doi: 10.3389/feduc.2023.1168459

Buck, R., and VanLear, C. A. (2002). Verbal and nonverbal communication: distinguishing symbolic, spontaneous, and Pseudo-spontaneous nonverbal behavior. *J. Commun.* 52, 522–541. doi: 10.1111/j.1460-2466.2002.tb02560.x

Cahyana, U., Mukhtar, V. A., and Yusmaniar, A. (2019). "The effect of mobile learning and motivation for students' high order thinking skills (HOTS) in electrolyte and nonelectrolyte solutions learning" in Empowering science and mathematics for global competitiveness. eds. Y. Rahmawati and P. Taylor (London: CRC Press), 270–276.

Chang, S. C., Hsu, T. C., Chen, Y. N., and Jong, M. S. y. (2020). The effects of spherical video-based virtual reality implementation on students' natural science learning effectiveness. *Interact. Learn. Environ.* 28, 915–929. doi: 10.1080/10494820.2018.1548490

Criollo-C, S., Guerrero-Arias, A., Jaramillo-Alcázar, Á., and Luján-Mora, S. (2021). Mobile learning technologies for education: benefits and pending issues. *Appl. Sci.* 11:4111. doi: 10.3390/app11094111

Damopolii, I., and Kurniadi, B. (2019). Training students metacognitive skill using mobile learning. *J. Phys. Conf. Ser.* 1317:2185. doi: 10.1088/1742-6596/1317/1/012185

Dasilva, B. E., Ardiyati, T. K., Suparno, K., Sukardiyono, M., Eveline, E., Utami, T., et al. (2019). Development of android-based interactive physics mobile learning media (IPMLM) with scaffolding learning approach to improve HOTS of high school students. *J. Educ. Gifted Young Sci.* 7, 659–681. doi: 10.17478/jegys.610377

Demircioglu, T., Karakus, M., and Ucar, S. (2023). Developing students' critical thinking skills and argumentation abilities through augmented reality–based argumentation activities in science classes. *Sci. Educ.* 32, 1165–1195. doi: 10.1007/s11191-022-00369-5

Dwikoranto, D., Setiani, R., and Tresnaningsih, S.. (2021). Effectiveness of Mobile Learning to Increase Collaborative Ability and Science Process Skills of Students. In: *International Joint Conference on Science and Engineering 2021 (IJCSE 2021)*, 441–446.

Elkot, M. A., Youssif, E., Hago Elmahdi, O. E., AbdAlgane, M., and Ali, R. (2025). Generative conversational AI: active practices for fostering students with mild intellectual disabilities to improve English communication skills. *Contemp. Educ. Technol.* 17:ep549. doi: 10.30935/cedtech/15688

Gikas, J., and Grant, M. M. (2013). Mobile computing devices in higher education: student perspectives on learning with cellphones, smartphones & social media. *Internet High. Educ.* 19, 18–26. doi: 10.1016/j.iheduc.2013.06.002

Gumbheer, C. P., Khedo, K. K., and Bungaleea, A. (2022). Personalized and adaptive context-aware Mobile learning: review, challenges and future directions. *Educ. Inf. Technol.* 27, 7491–7517. doi: 10.1007/s10639-022-10942-8

Hamidi, H., and Chavoshi, A. (2018). Analysis of the essential factors for the adoption of mobile learning in higher education: a case study of students of the University of Technology. *Telematics Inform.* 35, 1053–1070. doi: 10.1016/j.tele.2017.09.016

Hamzah, H., Hamzah, M. I., and Zulkifli, H. (2022). Systematic literature review on the elements of metacognition-based higher order thinking skills (HOTS) teaching and learning modules. *Sustainability* 14:813. doi: 10.3390/su14020813

Hariadi, B., Jatmiko, B., Sunarto, M. J. D., Prahani, B. K., Sagirani, T., Amelia, T., et al. (2022). Higher order thinking skills based learning outcomes improvement with blended web mobile learning model. *Int. J. Instr.* 15, 565–578. doi: 10.29333/iji.2022.15231a

Haryadi, R., and Pujiastuti, H. (2022). Enhancing pre-service physics teachers' higher-order thinking skills through STEM-PjBL model. *Int J STEM Educ Sustain* 2, 156–171. doi: 10.53889/ijses.v2i2.38

Hasbiyati, H., Sudiarti, D., and Hikamah, S. R. (2019). The effectiveness of using smartphone-based e-book in increasing students' learning outcomes in science learning. *IOP Conf. Ser. Earth Environ. Sci.* 243:2071. doi: 10.1088/1755-1315/243/1/012071

Hochberg, K., Kuhn, J., and Müller, A. (2018). Using smartphones as experimental tools—effects on interest, curiosity, and learning in physics education. *J. Sci. Educ. Technol.* 27, 385–403. doi: 10.1007/s10956-018-9731-7

Inel-Ekici, D., and Ekici, M. (2022). Mobile inquiry and inquiry-based science learning in higher education: advantages, challenges, and attitudes. *Asia Pac. Educ. Rev.* 23, 427–444. doi: 10.1007/s12564-021-09729-2

Isnaeni, W., Rudyatmi, E., Ridlo, S., Ingesti, S., and Adiani, L. R. (2021). Improving students' communication skills and critical thinking ability with ICT-oriented problem-based learning and the assessment instruments with HOTS criteria on the immune system material. *J. Phys. Conf. Ser.* 1918:2048. doi: 10.1088/1742-6596/1918/5/052048

Kaliisa, R., López-Pernas, S., Misiejuk, K., Damşa, C., Sobocinski, M., Järvelä, S., et al. (2025). A topical review of research in computer-supported collaborative learning: questions and possibilities. *Comput. Educ.* 228:105246. doi: 10.1016/j.compedu.2025.105246

Kapici, H. O., Akcay, H., and Koca, E. E. (2022). Comparison of the quality of written scientific arguments in different laboratory environments. *Int. J. Sci. Math. Educ.* 20, 69–88. doi: 10.1007/s10763-020-10147-w

Khadka, J., Dahal, N., Acharya, U., Puri, G., Subedi, N., and Hasan, M. K. (2025). Higher-order thinking skills in e-learning contexts in higher education: a phenomenological study. *Front Educ* 10:541. doi: 10.3389/feduc.2025.1555541

Khery, Y., Masjudin, A., Masjudin, M., Muzaki, A., Nufida, B., Lesnawati, Y., et al. (2020). Mobile-nature of science model of learning for supporting student performance on general chemistry classroom. *Int J Interact Mobile Technol* 14, 122–137. doi: 10.3991/IJIM.V14I12.15591

Kousloglou, M., Petridou, E., Molohidis, A., and Hatzikraniotis, E. (2023). Assessing students' awareness of 4Cs skills after Mobile-technology-supported inquiry-based learning. *Sustainability* 15:725. doi: 10.3390/su15086725

Kuanbayeva, B., Shazhdekeyeva, N., Zhusupkaliyeva, G., Mukhtarkyzy, K., and Abildinova, G. (2024). Investigating the role of augmented reality in supporting collaborative learning in science education: a case study. *Int. J. Eng. Pedagog.* 14, 149–161. doi: 10.3991/ijep.v14i1.42391

Kumar, A., Krishnamurthi, R., Bhatia, S., Kaushik, K., Ahuja, N. J., Nayyar, A., et al. (2021). Blended learning tools and practices: a comprehensive analysis. *IEEE Access* 9, 85151–85197. doi: 10.1109/ACCESS.2021.3085844

Levac, D., Colquhoun, H., and O'brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Sci.* 5:69. doi: 10.1186/1748-5908-5-69

Liu, J., Liu, Z., Wang, C., Li, X., and Xu, Y. (2024). Key factors and mechanisms affecting higher-order thinking skills of primary and secondary school students in the smart classroom environment. *Curr. Psychol.* 43, 9651–9664. doi: 10.1007/s12144-023-05136-5

Liu, D., and Zhang, H. (2022). Improving students' higher order thinking skills and achievement using WeChat based flipped classroom in higher education. *Educ. Inf. Technol.* 27, 7281–7302. doi: 10.1007/s10639-022-10922-y

Lu, K., Yang, H. H., Shi, Y., and Wang, X. (2021). Examining the key influencing factors on college students' higher-order thinking skills in the smart classroom environment. *Int. J. Educ. Technol. High. Educ.* 18:7. doi: 10.1186/s41239-020-00238-7

Ly, L. Q., and Kearney, M. (2023). Mobile learning in university science education: a systematic literature review. *Ir. Educ. Stud.* 43, 1287–1305. doi: 10.1080/03323315.2023.2256689

Made Tegeh, I., Bagus, I., Sundria, N., Komang Sudarma, I., Gde, I., and Sudatha, W. (2020). Mobile-learning development for supporting scientific approaches in chemical learning using flipped classroom strategy. *J Educ Res Eval* 4, 143–151. doi: 10.23887/jere.v4i2.24684

Maryuningsih, Y., Hidayat, T., Riandi, R., and Rustaman, N. Y. (2019). Critical thinking skills of prospective biology teacher on the chromosomal basic of inheritance learning through online discussion forums. *J. Phys. Conf. Ser.* 1157:2090. doi: 10.1088/1742-6596/1157/2/022090

Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educ. Psychol. Rev.* 36:9842. doi: 10.1007/s10648-023-09842-1

Mohammadi, M., Sarvestani, M. S., and Nouroozi, S. (2020). Mobile phone use in education and learning by faculty members of technical-engineering groups: concurrent mixed methods design. *Front Educ* 5:16. doi: 10.3389/feduc.2020.00016

Montgomery, T. D., Buchbinder, J. R., Gawalt, E. S., Iuliucci, R. J., Koch, A. S., Kotsikorou, E., et al. (2022). The scientific method as a scaffold to enhance communication skills in chemistry. *J. Chem. Educ.* 99, 2338–2350. doi: 10.1021/acs.jchemed.2c00113

Naveed, Q. N., Choudhary, H., Ahmad, N., Alqahtani, J., and Qahmash, A. I. (2023). Mobile learning in higher education: a systematic literature review. *Sustainability* 15:3566. doi: 10.3390/su151813566

Putranta, H., Supahar, N., Setiyatna, H., Choiriyah, S., Dwandaru, W. S. B., Widodo, S., et al. (2021). The effect of smartphone usage intensity on high school students' higher order thinking skills in physics learning. *J. Turk. Sci. Educ.* 18, 421–438. doi: 10.36681/tused.2021.82

Putri, A. S., and Aznam, N. (2019). The effectiveness of science learning media using Focusky software on junior high school students' higher order thinking skills. *Sci Educ* 8:12. doi: 10.24235/sc.educatia.v8i1.3886

Seif, A. A. (2023). Use of logic for improving the higher-order thinking skills of student teachers. *Eur J Interact Multimedia Educ* 4:e02304. doi: 10.30935/ejimed/13393

Sergeeva, O. V., Zheltukhina, M. R., Bikbulatova, G. I., Sokolova, E. G., Digtyar, O. Y., Prokopyev, A. I., et al. (2023). Examination of the relationship between information and communication technology competencies and communication skills. *Contemp. Educ. Technol.* 15:ep483. doi: 10.30935/cedtech/13819

Setyaningrum, R., Wijaya, A. R., and Subagyo, S. (2022). The characteristics of society in Indonesian based on the Hofstede cultural dimensions: measuring the five Indonesian island. *Int J Cult Art Stud* 6, 60–74. doi: 10.32734/ijcas.v6i1.8720

Shao, M., and Liu, X. (2021). Impact of the flipped classroom on students' learning performance via Meta-analysis. *Open J. Soc. Sci.* 9, 82–109. doi: 10.4236/jss.2021.99007

Song, Y., Cao, J., Yang, Y., and Looi, C. K. (2022). Mapping primary students' mobile collaborative inquiry-based learning behaviours in science collaborative problem solving via learning analytics. *Int. J. Educ. Res.* 114:101992. doi: 10.1016/j.ijer.2022.101992

Staacks, S., Hütz, S., Heinke, H., and Stampfer, C. (2018). Advanced tools for smartphone-based experiments: phyphox. *Phys. Educ.* 53:045009. doi: 10.1088/1361-6552/aac05e

Sulisworo, D., Mutammimah, M., and Parwiti, P. (2018). "The impact of cooperative blended learning on higher order thinking skills using a mobile application" in Character education for 21st century global citizens. eds. E. Retnowati, A. Ghufron, Marzuki, Kasiyan, A. Pierawan and Ashadi (London: Routledge), 291–297.

Sun, H., Xie, Y., and Lavonen, J. (2022). Exploring the structure of students' scientific higher order thinking in science education. *Think. Skills Creat.* 43:100999. doi: 10.1016/j.tsc.2022.100999

Sung, Y. T., Chang, K. E., and Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: a meta-analysis and research synthesis. *Comput. Educ.* 94, 252–275. doi: 10.1016/j.compedu.2015.11.008

Supeno, A., Astutik, S., Bektiarso, S., Lesmono, A., and Nuraini, L. (2019). What can students show about higher order thinking skills in physics learning? *IOP Conf Series Earth Environ Sci* 243:2127. doi: 10.1088/1755-1315/243/1/012127

Temsah, L., and Safa, N. (2021). New approaches to simulation-based science instruction to enhance reasoning and communication skills in Lebanese elementary education. *Middle East J Res Educ Soc Sci* 2, 56–79. doi: 10.47631/mejress.v2i1.180

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann. Intern. Med.* 169, 467–473. doi: 10.7326/M18-0850

Vilela, M., Morais, C., and Paiva, J. C. (2025). Inquiry-based science education in high chemistry: enhancing Oral and written communication skills through authentic and problem-based learning activities. *Educ Sci* 15:334. doi: 10.3390/educsci15030334

Wahono, B., Lin, P. L., and Chang, C. Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *Int. J. STEM Educ.* 7:1. doi: 10.1186/s40594-020-00236-1

Wahyuni, I., Mukrimah, M., Hanifah, F. N., Ariesca, J. D., and Windayani, W. (2022). Question card game to improve senior high school students' higher-order thinking skills in biology learning. *Int J Biol Educ Towards Sustain Dev* 2, 63–70. doi: 10.52889/ijbetsd.v2i2.147

Weidlich, J., Göksün, D. O., and Kreijns, K. (2023). Extending social presence theory: social presence divergence and interaction integration in online distance learning. *J. Comput. High. Educ.* 35, 391–412. doi: 10.1007/s12528-022-09325-2

Zana, F. M., Sa'dijah, C., Susiswo, A., and Zulnaidi, H. (2024). Curriculum and teacher assessment practices in mathematics learning: alignment with higher order thinking skills in Indonesian secondary schools. *J Maths Educ* 15, 1311–1334. doi: 10.22342/jme.v15i4.pp1311-1334

Zhang, Y., Wang, J., Yan, Y., Xu, J., Li, H., Zhang, T., et al. (2024). Enhancing medical students' science communication skills: from the perspective of new media. *Adv. Physiol. Educ.* 48, 288–294. doi: 10.1152/advan.00192.2023

Ziden, A. A., Ziden, A. A. A., and Ifedayo, A. E. (2022). Effectiveness of augmented reality (AR) on students' achievement and motivation in learning science. *Eurasia J. Math. Sci. Technol. Educ.* 18:em2097. doi: 10.29333/ejmste/11923