Extended Reality in Computer Science Education: An Umbrella Review of Reviews

Muhammad Ahsan[1], **Maria Fardous**[2], **Laura Coura**[1], **Saul Delabrida**[1]

[1] Postgraduate Program in Computer Science, Federal University of Ouro Preto, Brazil

[2]Department of Computing, Abasyn University, Islamabad, Pakistan

muhammad.ahsan@aluno.ufop.edu.br, maryaakhan72@gmail.com

laura.coura@aluna.ufop.edu.br, saul.delabrida@ufop.edu.br

Abstract. *The Extended Reality (XR) technologies of Augmented, Virtual, and Mixed Reality are increasingly proposed for computer science education (CSE), yet the existing review and survey literature is fragmented across technologies, topics, and methods. This umbrella review synthesizes evidence from reviews and surveys on XR in computer science and related computing education, focusing on devices, platforms, software, educational levels, learning outcomes, and reported challenges. We searched the Scopus database from 2013 to 2025 for review-type or systematic studies on XR in computing education. We included five reviews covering 224 primary studies published between 2009 and 2023. While XR consistently supports motivation, engagement, and conceptual understanding of complex computing concepts, technical limitations, accessibility constraints, and weak instructional design remain persistent challenges. We identify trends, gaps, and research priorities, offering actionable implications for computing educators and researchers.*

Keywords: XR; Systematic Review; Computer Science Education; Umbrella Review; PRISMA.

1. INTRODUCTION

Extended Reality (XR) technologies, encompassing Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), and WebXR, are increasingly adopted to support teaching and learning in computer science education (CSE) [Pirker et al. 2021]. These immersive environments allow students to visualize abstract concepts, interact with complex systems, and practice skills in safe, controlled settings, which is particularly valuable for topics such as algorithms, data structures, programming, and cybersecurity. In educational contexts, XR is relevant because it enables learners to interact using games with cross-curricular learning [MacCallum 2022]. For CSE specifically, XR supports game-based approach for teaching algorithms and data structures [Su et al. 2021], situated practice of programming and debugging in meaningful scenarios, and immersive simulations of cybersecurity or networking environments that are difficult to replicate in traditional classrooms [Alnajim et al. 2023]. Several quasi-experimental and experimental studies provide evidence of higher post-test scores [Tsai and Lai 2022], greater attention and motivation [Ekman et al. 2024], and improved practical skills [Lampropoulos et al. 2023, Buchner and Kerres 2021] when XR is well designed. Evidence also emphasizes that a sense of presence including spatial, social, cognitive, and

co-presence dimensions significantly influences learning, underlining the importance of instructional design, interaction strategies, and tailored environments [Wei et al. 2025].

Existing literature shows that XR has been applied across a broad range of computer science topics, from introductory programming and data structures to algorithmic problem solving and computational thinking, typically through task- or game-based designs aligned with curricular objectives. Prior studies report the use of AR for logic and introductory programming [Tsai and Lai 2022, Mina et al. 2021], mobile AR applications for Python and data-structure practice [Branco and Pombo 2023, Narman et al. 2020], and VR simulations or games for algorithm visualisation and problem solving [Villegas-Ch et al. 2024, Dewi et al. 2024]. Beyond technical learning, AR- and gamification-based approaches have also been associated with increased engagement, motivation, and collaborative skills [Lampropoulos et al. 2023, Buchner and Kerres 2021]. However, existing reviews and primary studies often focus on single technologies, specific educational levels, or narrow methodological perspectives, resulting in a fragmented evidence base that limits consolidated understanding of effective XR designs and research gaps in CSE. As XR in CSE is already represented by several systematic reviews and mapping studies, an umbrella review is appropriate to integrate review-level evidence, compare findings across reviews, and identify consistent gaps that are not visible in a single systematic review. This umbrella review synthesizes review-level evidence on XR in CSE.

XR in CSE has expanded rapidly in scope and scale, yet educators and researchers still lack a consolidated, higher-level view that connects technologies, applications, outcomes, and methodological quality across existing reviews. This umbrella review is relevant because it synthesizes dispersed review-level evidence into a coherent picture that directly addresses the four research questions as follows:

1. **RQ1 – Technological landscape:** What XR devices, platforms, and software tools are reported across reviews of XR in CSE?
2. **RQ2 – Educational applications and outcomes:** Which CSE topics and educational levels are targeted, and what learning outcomes and evidence are reported?
3. **RQ3 – Global and regional patterns:** How do XR-CSE review findings and emphases vary across regions, with particular attention to Brazil/Latin America versus the rest of the world?
4. **RQ4 – Methodological quality and gaps:** What is the methodological quality of existing XR-CSE reviews, and what gaps and future research priorities do they identify?

This umbrella review contributes to the literature in three main ways. First, it consolidates results from five recent systematic reviews and systematic mapping studies, together covering 224 primary studies, into a single, coherent synthesis focused on Computer Science and programming education. Second, it maps XR technologies, educational applications, and outcome evidence across both K–12 and higher education, highlighting convergent findings and divergences between contexts. Third, it systematically identifies cross-cutting methodological limitations and research gaps identified in the past reviews including regional diversity, thereby providing a concrete future research agenda for XR in CSE.

2. METHODS

This study adopts an umbrella review approach, synthesizing evidence from systematic reviews and mapping studies rather than primary empirical research [Fernandez et al. 2025]. This umbrella review was carried out following the PRISMA[1] protocol, which ensures a structured approach throughout the entire process from data collection until reporting of results [Page et al. 2021]. For the purpose of data collection, we selected the Scopus[2] database, as it integrates content from various major sources in the field (e.g., ACM, IEEE Xplore, and Springer) and has been widely adopted as a comprehensive data source. The search query was: **(TITLE-ABS (”virtual reality”) OR TITLE-ABS (VR) OR TITLE-ABS (”augmented reality”) OR TITLE-ABS (AR) OR TITLE-ABS (”mixed reality”) OR TITLE-ABS (XR) OR TITLE-ABS (”immersive learning”) OR TITLE-ABS (WebXR)) AND (TITLE-ABS (”computer science education”) OR TITLE-ABS (”computer education”) OR TITLE-ABS (”programming education”) OR TITLE-ABS (”computing education”) OR TITLE-ABS (”software engineering education”)) AND (review OR ”systematic review” OR ”scoping review” OR ”meta-analysis” OR survey OR ”literature review”)**.

2.1. Eligibility Criteria

Eligibility criteria were defined a priori to identify review-level studies relevant to XR in computer science and programming education. The **inclusion criteria** are studies explicitly identified as review-type publications (systematic reviews, mapping reviews, scoping reviews, meta-analyses, surveys, or structured literature reviews) and that examined XR technologies (including VR, AR, MR, XR, WebXR, or metaverse platforms) as a primary focus or substantial component. Eligible studies needed to address computer science, computing, programming, or software engineering education, be published as peer-reviewed journal articles or full conference papers, and written in English between 2013 and 2025. Additionally, studies were required to provide adequate methodological transparency regarding their search procedures, selection criteria, and synthesis approach in order to be considered systematic in nature

Studies were part of the **exclusion criteria** if they were primary empirical research, non-review conference papers, thesis chapters, opinion or commentary pieces, book chapters, or errata. Non-English papers were also removed. Bibliometric studies were excluded because they primarily offer descriptive overviews of research activity rather than conceptual or systematic syntheses. Finally, three review studies that appeared relevant based on their titles and abstracts were excluded because full texts could not be obtained despite contacting the authors.

2.2. Search Procedure

The Scopus database was searched for title/abstract terms related to XR and computing education between 2013 and 2025 using a predefined query (see Section 2). The initial search returned 135 records. After removal of errata/book chapters (n=8), non-English items (n=2), and primary empirical studies (n=110), 15 candidate reviews remained. Following a full-text assessment and team discussion, we further excluded out-of-scope reviews (n=4), bibliometric analyses (n=3) that were not conceptual or systematic reviews,

[1] http://prisma-statement.org
[2] https://www.scopus.com/home.uri.

and 3 reviews that were inaccessible despite attempts to contact the authors. Ultimately, 5 reviews met all criteria and were included in this umbrella review. Figure 1 presents a flowchart that summarizes our procedures for identifying the final sample based on the PRISMA framework.

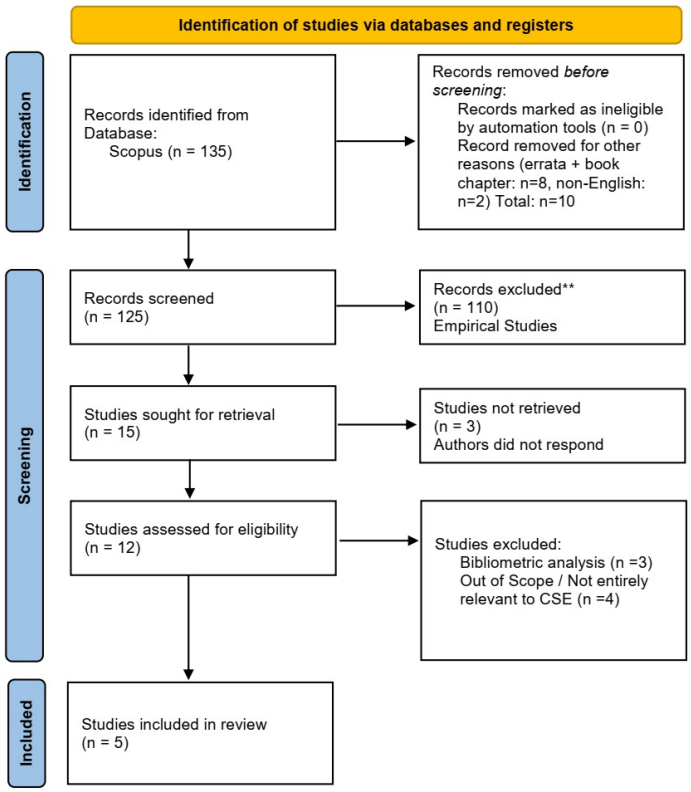


Figure 1. PRISMA flowchart for the umbrella review

From each included review we extracted: scope and aims, review type, databases searched, inclusion criteria used by the review authors, number of primary studies covered, reported XR devices/platforms/software (RQ1), targeted CS topics and educational levels (RQ2), reported learning outcomes and evidence (qualitative/quantitative), geographic coverage (RQ3), and the methodological strengths/limitations (RQ4) reported by the review authors. Extraction was performed independently by two authors to ensure reliability and validity, and disagreements were resolved through discussion.

Because the aim was to synthesize at reviews level, the umbrella review assessed methodological transparency and completeness rather than attempting to re-evaluate risk of bias for all underlying primary studies. We examined the included reviews' design and reporting quality using standard characteristics for review reporting (explicit research questions, database list and search dates, inclusion/exclusion criteria, study selection flow, critical appraisal of included primary studies, and transparency of synthesis methods).

3. RESULTS

The five included reviews (as denoted below from R1 to R5) together synthesize 224 primary studies including studies published between 2009 and 2023 (See Figure 2), covering

XR technologies in educational contexts relevant to computer science and programming. They comprise:

1. R1: A systematic review of educational metaverses across disciplines, drawing on 26 studies and using the Evaluating Game and Simulation-based Education (EGSE) framework [Ak et al. 2025].
2. R2: A systematic review of emerging technologies AI, learning analytics , and XR for assessment and feedback in higher education computing, analysing 38 studies [Sembey et al. 2024].
3. R3: A systematic mapping review of emerging technologies (AI, ML, IoT, AR, VR) in K–12 education, based on 107 records from multiple fields including computing education [Van Mechelen et al. 2023].
4. R4: A systematic review of virtual reality in computer science education, focusing on devices, contexts, and evaluation methods in 22 VR focused CS studies [Malik et al. 2023].
5. R5: A systematic review of augmented reality in programming education, analysing 31 AR-based interventions aimed at computational thinking and coding skills [Theodoropoulos and Lepouras 2021].

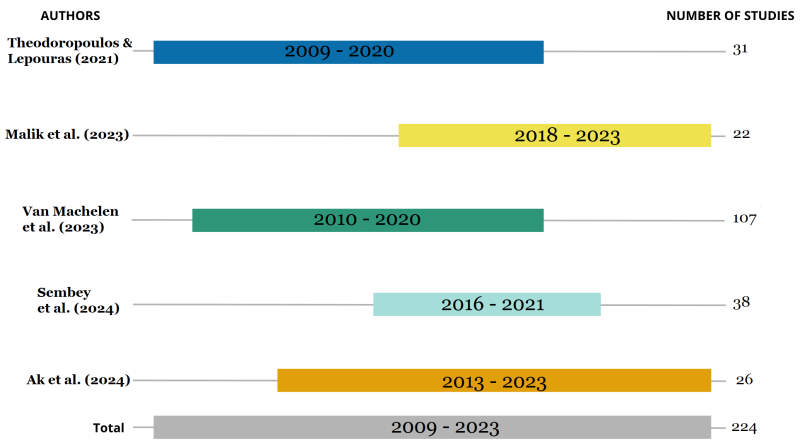


Figure 2. Publication development of articles in this umbrella review (n = 5)

In the following sections, we answer the defined RQs considering the results obtained from the analysis of these reviews.

3.1. RQ1 – Technological Landscape: XR Devices, Platforms, and Tools

The review on VR in CSE reports that VR implementations rely mainly on head-mounted displays, with Oculus-class consumer headsets most commonly used to deliver immersive, game-like instructional modules and serious games for topics such as programming, data structures, algorithms, cybersecurity, cloud computing, and IoT [Malik et al. 2023]. Evaluations typically combine user-experience and usability questionnaires with performance-based measures (e.g., tests or task completion). The AR programming review (R5) synthesises classroom and child–computer interaction studies in which AR is predominantly delivered via mobile phones and tablets, complemented by tangible AR setups that augment physical blocks to represent programming constructs

and control flow [Theodoropoulos and Lepouras 2021]. Most systems are bespoke educational applications (e.g., custom-built software solutions), often marker-based and occasionally using head-mounted AR, designed to make abstract programming concepts visible and manipulable for novice learners [Theodoropoulos and Lepouras 2021]. The metaverse review (R1) depicts a broader XR landscape, where 3D virtual worlds (e.g., Second Life, OpenSim), 2D metaverse-like platforms (e.g., Gather Town), and hybrid XR-integrated environments are used for simulations, role-play, collaboration, and virtual field trips across disciplines [Ak et al. 2025]. Access ranges from desktop and laptop computers to fully immersive HMD-based configurations, with a trend towards more configurable platforms tightly coupled to learning management systems[Ak et al. 2025].

In higher-education computing assessment and feedback, XR appears as VR training environments and AR applications (including HoloLens) embedded in assessment workflows to provide performance and metacognitive feedback alongside data collection [Sembey et al. 2024]. Although less prominent than AI and learning analytics, these cases illustrate how XR can be integrated into feedback cycles in computing courses.[Sembey et al. 2024]. The K–12 emerging-technologies mapping (R3) treats AR and VR as one category of "technologies and tools for learning", indicating that both off-the-shelf and custom XR systems are used alongside AI, ML, and IoT to support computing and computational literacy, without detailing specific devices [Van Mechelen et al. 2023].

Summary: Across the five reviews, consumer VR HMDs and mobile/tablet-based AR are the most frequently reported hardware, with Unity as the dominant development environment and AR frameworks such as ARCore/ARKit and Vuforia commonly used for mobile AR.[Ak et al. 2025, Sembey et al. 2024, Van Mechelen et al. 2023, Malik et al. 2023, Theodoropoulos and Lepouras 2021]. Web-based XR (e.g., WebXR/WebGL) is noted as an emerging but less common solution, while software implementations are predominantly bespoke visualisations, block-based programming environments, and simulation systems rather than reusable, standardised platforms, limiting replication and scalability [Ak et al. 2025, Van Mechelen et al. 2023, Malik et al. 2023, Theodoropoulos and Lepouras 2021].

3.2. RQ2 – Educational Applications and Outcomes

Across the five reviews, XR is applied to a wide range of CSE topics and learning objectives. In K–12 contexts, AR and VR are primarily used to support:

1. Fundamental programming concepts such as variables, conditionals, loops, and functions.
2. Introduction to data structures like lists, arrays, stacks, and queues.
3. Algorithmic thinking, including searching and sorting algorithms.
4. Computational thinking skills such as decomposition, pattern recognition, and debugging.

In higher education, VR and XR systems support more advanced CS topics:

1. Cybersecurity and network security, including firewalls, intrusion detection and prevention, honeypots, and attack simulations.
2. Cloud computing and mobile-edge computing, where virtual environments simulate distributed infrastructures and services.

3. IoT and tactile internet scenarios, where networked devices and sensor data are visualised and manipulated in virtual environments.
4. Software and systems courses using simulations and games to teach design, testing, and system behaviour.
5. The VR review (R4) provides examples of bubble-sort visualisations, AI/ML-enhanced VR labs, and network security training games, illustrating the breadth of CS topics addressed with immersive technologies.

A word cloud was generated based on the frequency of CS topics across the five reviews as shown in Figure 3. The AR-focused review (R5) [Theodoropoulos and Lepouras 2021] reports applications in which learners use tangible blocks or AR overlays to construct and test algorithms, supporting programming and computational thinking through immediate visual feedback. Moreever, from (R3) in K–12 settings, AR and VR are also used to promote broader computational literacy and critical engagement with emerging technologies, often in conjunction with AI, ML, and IoT concepts [Van Mechelen et al. 2023].



Figure 3. Word cloud of the CSE topics using XR technologies

Across all five reviews, XR is linked to cognitive gains, though effect sizes and study quality vary. AR tends to improve programming and computational thinking, while VR supports topics such as data structures and cybersecurity [Malik et al. 2023]. Compared with non-immersive tools, XR more consistently improves engagement and motivation than learning outcomes, and assessment-focused studies report real-time feedback but limited evidence of long-term retention or transfer [Sembey et al. 2024].

Summary: Overall, the review-level evidence suggests that XR technologies can effectively support a wide range of CS topics at different educational levels, with especially strong indications of positive effects on engagement and motivation. There is also consistent, though more variable, evidence of cognitive benefits in programming and advanced CS topics. However, these benefits are moderated by design quality, assessment practices, and contextual factors such as resources, teacher expertise, and student characteristics.

3.3. RQ3 - Global and Regional Research Patterns

The five included reviews collectively draw on studies conducted across multiple regions, but most do not provide detailed country-by-country breakdowns. The systematic map-

ping of emerging technologies in K–12 education (R3) from [Van Mechelen et al. 2023] records the first author's country of affiliation for each study and notes that contributions are unevenly distributed, with higher representation from Europe, North America, and parts of Asia than from other regions. Specific counts for Brazil or Latin America are not reported in the available summary. The VR in CS review (R4) from [Malik et al. 2023] and the AR in programming review (R5) from [Theodoropoulos and Lepouras 2021] describe their scope as global but do not systematically report geographic distributions of primary studies or highlight Latin American contributions. Likewise, the metaverse review (R1) from [Ak et al. 2025] and the assessment-and-feedback review (R2) from [Sembey et al. 2024] reference studies from multiple regions, including Europe and Asia, but do not provide structured regional comparisons or specific analysis of Brazil or Latin America. Taken together, these reviews indicate that XR research in education and especially in computing education is concentrated in a limited set of countries and institutions, with notable visibility of Taiwan, Spain, the United States, and some European and Asian universities. However, they do not allow a precise quantification of global patterns or regional disparities for XR in CSE, and they provide little explicit information about Latin American contexts. This constitutes a clear gap that this umbrella review can highlight for future research. Because most included reviews did not systematically disaggregate findings by country or region, conclusions regarding global distribution patterns should be interpreted cautiously.

Summary: The included reviews indicate that primary research is concentrated in North America, Europe, and East Asia; Latin America (including Brazil) and other regions remain underrepresented. Reviews rarely disaggregated results to permit systematic regional comparisons, and none of the five reviews provides in-depth coverage of Brazil/Latin America specifically. The field therefore lacks robust, regionally balanced evidence, and future reviews should include targeted searches in regional databases and non-English literature to better represent global practice.

3.4. RQ4 - Methodological Quality and Research Gaps

Methodological rigor was assessed using six criteria: (i) explicit research questions; (ii) transparent search strategies; (iii) clearly defined inclusion and exclusion criteria; (iv) multi-stage screening procedures; (v) structured coding or synthesis frameworks; and (vi) reporting of methodological limitations or quality-appraisal procedures. Evidence strength was assessed using four criteria: (i) consistency of findings across included studies; (ii) breadth and diversity of the evidence base; (iii) comparative or longitudinal evidence; and (iv) the extent to which conclusions were supported by empirical evaluation rather than descriptive reporting alone. Methodological rigor and evidence strength were assessed using a researcher-developed comparative rubric. The rubric was constructed as: 1 = low, 2 = moderate, and 3 = high. A score of 1 indicated that several relevant criteria were absent or insufficiently reported, a score of 2 indicated that most core criteria were addressed but important limitations remained; and a score of 3 indicated that the relevant criteria were addressed clearly and consistently. Two authors independently assessed the reviews and resolved disagreements through consensus discussion.

The methodological profile of the included reviews is mixed. All five reviews demonstrate reasonable procedural transparency in search and selection, but none applies a fully explicit and validated risk-of-bias framework to the underlying primary

studies. This limits the certainty with which their conclusions can be interpreted. In umbrella reviews, methodological rigor should not be reduced to whether a review followed a search protocol; it should also include the clarity of quality appraisal, the transparency of coding decisions, and the reproducibility of evidence synthesis. The R4 [Malik et al. 2023] and R5 [Theodoropoulos and Lepouras 2021] follow standard systematic literature review procedures, including clearly defined research questions, transparent screening criteria, and structured synthesis of devices, contexts, and outcomes. The R1 [Ak et al. 2025] uses the EGSE framework to organize evidence along four dimensions: pedagogical considerations, learner specification, mode of representation, and context. The R2 [Sembey et al. 2024] and R5 [Theodoropoulos and Lepouras 2021] applied Kitchenham's guidelines for systematic reviews, and the R3 [Van Mechelen et al. 2023] adheres to PRISMA-ScR and uses double coding and plenum consensus procedures. However, none of the reviews employs a formal, validated tool for assessing risk of bias in primary studies, and their reporting of study quality is typically narrative and high-level. Table 1 presents a comparative assessment of the five included reviews in terms of methodological rigor and evidence strength.

Table 1. Evidence Strength and Methodological Rigor of the Included Reviews

Study ID	Evidence Strength (1–3)	Methodological Rigor (1–3)
R1	2	1
R2	3	3
R3	2	3
R4	2	2
R5	3	2

Summary: RQ4 indicates that the included reviews vary in their methodological profile and evidential contribution. R2 showed the strongest combined profile, with high methodological rigor and evidence strength. R3 also achieved high procedural rigor through PRISMA-ScR reporting, double coding, and consensus procedures, although its evidence strength remained moderate because it was primarily a broad descriptive mapping review. R5 showed high evidence strength and moderate methodological rigor: it employed systematic review procedures and a Kitchenham-informed quality assessment, but its review reporting and synthesis procedures were assessed as less comprehensive than those of R2 and R3. R4 occupied an intermediate position, whereas R1 showed lower methodological rigor because of its broader and more exploratory design. Overall, the field demonstrates increasing use of systematic review practices, but still varies in the transparency, completeness, and comparability of quality appraisal and evidence synthesis.

4. DISCUSSION

The main cross-cutting similarities, advantages, and limitations identified across the five reviews are synthesized in Table 2, which provides a structured summary of the discussion lenses used to compare and interpret evidence across selected studies.

Table 2. Summary of the Discussion's Lenses among Five Reviews on XR in CSE

Discussion Lens	Sub-dimension	Synthesised Evidence Across Reviews
Similarities and differences	Shared characteristics	All five reviews (R1–R5) examine XR technologies in computing education, formulate explicit research questions, and cover both K–12 and higher education contexts.
	Divergent foci	R1 emphasizes metaverse and cross-disciplinary XR, R2 focuses on assessment and feedback, R3 on emerging technologies in K–12, R4 on VR in CSE, and R5 on AR for programming and computational thinking.
Advantages of XR in CSE	Affective outcomes	XR is consistently associated with higher engagement, motivation, and satisfaction across the reviewed contexts (R1–R5).
	Cognitive outcomes	XR can support conceptual understanding, problem-solving, and computational thinking when aligned with clear objectives, scaffolding, and feedback, particularly in R2, R4, and R5.
Limitations and research gaps	Methodological limitations	Recurring issues include small samples, short interventions, limited longitudinal evidence, heterogeneous outcome measures, and weak statistical reporting, which constrain causal inference (R1–R5).
	Contextual and equity-related gaps	The reviews underreport teacher development, infrastructure, accessibility, standardised assessment, and regional diversity, while formal risk-of-bias appraisal is largely absent (R1–R5).
Interpretive lens	Sociotechnical view	XR is best understood as a sociotechnical educational assemblage shaped by pedagogy, access, institutional support, data practices, and learner context rather than as a standalone delivery tool. This helps explain why similar XR interventions yield different outcomes and why engagement does not always translate into deeper learning (R1–R5).
Critical and sociotechnical perspectives	Ethical and critical concerns	The evidence base remains strongly innovation-centered, with limited attention to privacy, surveillance, psychological strain, design bias, vulnerable learners, and unequal access. These omissions are analytically significant and indicate a need for more critical XR research in CSE (R1–R5).

The five reviews included in this umbrella review span publication windows from 2009 to 2023, with most primary studies published after 2013, reflecting the rapid expansion of XR research in education over the past decade. Overall, the synthesis indicates that XR technologies including metaverse environments, VR, and AR have been widely explored as instructional tools in computer science and programming education across both K–12 and higher education contexts. However, the evidence base remains fragmented, methodologically heterogeneous, and unevenly distributed across regions, educational levels, and application domains, limiting the strength of generalisable conclusions regarding effectiveness (R1–R5).

4.1. Similarities and Differences Across Reviews

Despite differences in scope and emphasis, all five reviews share several common characteristics. Each review explicitly formulates research questions, focuses on XR or closely related immersive technologies, and reports growing interest in applying these technologies to computer science or computing-related education across multiple educational levels (R1–R5). At the same time, notable differences emerge in terms of technological focus (e.g., metaverse environments, VR-only, AR for programming), educational context (K–12 versus higher education), and thematic emphasis (learning outcomes, assessment, emerging technologies, or disciplinary focus). These variations lead to differences in included primary studies, methodological approaches, and reported outcomes, which complicates direct comparison and cross-review synthesis (R1–R5).

4.2. Advantages of XR in CSE

Across all five reviews, there is convergent evidence that XR technologies offer clear affective and motivational benefits. Increased learner engagement, motivation, and satisfaction are consistently reported in metaverse-based learning scenarios (R1), higher education assessment and feedback contexts (R2), K-12 emerging-technology settings (R3), VR-supported computer science courses (R4), and AR-based programming and computational thinking activities (R5). Beyond affective outcomes, several reviews also report cognitive and skills-related benefits when XR is meaningfully integrated into instructional designs. Specifically, improvements in conceptual understanding, problem-solving, and computational thinking are more likely when XR interventions are aligned with clear learning objectives and supported by scaffolding, feedback, and coherent pedagogical strategies, rather than being deployed as stand-alone or novelty-driven tools (R2, R4, R5).

4.3. Limitations and Research Gaps

All five reviews consistently highlight substantial methodological and contextual limitations within the existing XR literature. Common weaknesses include small sample sizes, short intervention durations, limited use of longitudinal designs, heterogeneous outcome measures, and insufficient statistical reporting, which collectively restrict strong causal inference and robust estimation of effect sizes for XR in CSE (R1-R5). This study is limited by its reliance on Scopus as the sole database and by the fact that the evidence base itself is review-level rather than primary-study level. Future updates should include additional databases and, where feasible, broader language coverage to reduce the risk of missing relevant reviews.

In addition, several contextual and equity-related gaps are identified, including limited attention to teacher expertise and professional development, institutional support, accessibility and inclusive design, standardized assessment instruments, and formal risk-of-bias appraisal. The reviews also point to uneven geographical coverage and limited region-specific analyses, suggesting that current evidence does not fully capture global educational contexts. These gaps indicate a need for shared conceptual frameworks, stronger methodological standards, and geographically diverse research agendas.

4.4. Interpretive lens

To strengthen the interpretative dimension of this umbrella review, the study interprets XR in CSE as a socio-technical educational assemblage rather than only a delivery tool. From

this perspective, learning outcomes are shaped not only by the immersive interface itself, but also by pedagogical design, access conditions, institutional support, data practices, and the learner's situated context. From this perspective, issues such as accessibility, data collection, platform dependency, cognitive overload, algorithmic bias, and unequal technological access become central analytical concerns when evaluating XR adoption in educational settings. This lens helps explain why similar XR technologies can produce different effects across studies and why engagement gains do not always translate into deeper conceptual learning.

4.5. Critical and Sociotechnical Perspectives on XR in CSE

The current evidence base is notable for what it emphasizes and what it leaves out. Across the included reviews, XR is mainly framed as a tool for improving engagement, motivation, and conceptual understanding, while critical concerns receive little attention. A more balanced account should address accessibility barriers, privacy and data-governance concerns, psychological strain, design bias, surveillance risks, and unequal access for vulnerable learners. The relative silence on these issues is itself analytically important, because it suggests that the field remains predominantly innovation-centered rather than critically and ethically oriented.

5. CONCLUSION AND FUTURE RESEARCH

This umbrella review synthesizes evidence from five systematic reviews and mapping studies that together cover 224 primary studies on XR in educational contexts relevant to computer science and programming. The results indicate that VR, AR, and related XR technologies can enhance engagement, motivation, and conceptual understanding in CSE when they are carefully integrated into pedagogically sound learning designs. At the same time, the field faces significant methodological, regional, and equity-related gaps, including a lack of standardized assessments, limited long-term evidence, explicit quality appraisal of primary evidence, standardized outcome measures, under representation of certain regions such as Latin America, and insufficient attention to accessibility. We recommend: (1) investing in accessible platforms (web/mobile) and shared toolkits; (2) conducting larger, controlled, and longitudinal studies with validated outcome measures; (3) attending to equity, accessibility, and low-cost deployment contexts (with Brazil/Latin America); and (4) strengthening reporting standards to improve reproducibility.

Future XR research in CSE should prioritize longitudinal and well-controlled experimental designs that move beyond exploratory implementations. The adoption of standardized assessment frameworks for learning outcomes, engagement, and usability would significantly improve comparability across studies. Additionally, greater attention is needed on pedagogical integration, including the role of instructional design, teacher mediation, and scaffolding strategies in XR-supported learning environments. Research should also address underexplored dimensions such as accessibility, ethical considerations, and equity, particularly in regions that remain underrepresented in the current literature. Strengthening methodological rigor in these areas will be essential for establishing XR as a sustainable and evidence-based educational approach.

ACKNOWLEDGEMENTS

This work was carried out with the support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Funding Code 001. FAPEMIG Grant

code APQ-06466-24. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Grant code 30115/2025-7.

Statement on the use of Artificial Intelligence

It is explicitly stated that Generative Artificial Intelligence (AI) tools were used exclusively in the stages of grammatical review, stylistic improvement of content and refinement of the textual cohesion of excerpts from this article. It is noteworthy that such computational tools were not listed nor do they meet the requirements for authorship or co-authorship of the work. Human authors assume full and complete responsibility for all intellectual and scientific content presented, guaranteeing the originality of the text and the absence of plagiarism in all sections of the article.

References

Ak, O., Şen Akbulut, M., Soydan, S., and Yildirim, E. (2025). Designing educational metaverses: a literature review to determine guidelines. *Interactive Learning Environments*, 33(2):925–943.

Alnajim, A. M., Habib, S., Islam, M., AlRawashdeh, H. S., and Wasim, M. (2023). Exploring cybersecurity education and training techniques: a comprehensive review of traditional, virtual reality, and augmented reality approaches. *Symmetry*, 15(12):2175.

Branco, J. and Pombo, N. (2023). Augmented reality for programming teaching: an exploratory study. In *2023 IEEE Global Engineering Education Conference (EDUCON)*, pages 1–7. IEEE.

Buchner, J. and Kerres, M. (2021). Students as designers of augmented reality: Impact on learning and motivation in computer science. *Multimodal Technologies and Interaction*, 5(8):41.

Dewi, I. P., Mursyida, L., Effendi, H., Giatman, M., Hanafi, H. F., Ali, S. K., et al. (2024). Virtual reality in algorithm programming course: practicality and implications for college students. *JOIV: International Journal on Informatics Visualization*, 8(3-2):1720–1727.

Ekman, J., Solsona, J., and Quintero, L. (2024). Codeseum: Learning introductory programming concepts through virtual reality puzzles. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences*, pages 192–200.

Fernandez, R. S., Sharifnia, A. M., and Khalil, H. (2025). Umbrella reviews: a methodological guide. *European journal of cardiovascular nursing*, 24(6):996–1002.

Lampropoulos, G., Keramopoulos, E., Diamantaras, K., and Evangelidis, G. (2023). Integrating augmented reality, gamification, and serious games in computer science education. *Education Sciences*, 13(6):618.

MacCallum, K. (2022). The integration of extended reality for student-developed games to support cross-curricular learning. *Frontiers in Virtual Reality*, 3:888689.

Malik, H. A., Ferdinand, R., Aurelius, B., Fajar, M., and Suri, P. A. (2023). Applications of virtual reality in computer sciences education: A systematic literature review. In *2023 International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS)*, pages 659–664. IEEE.

Mina, D., Salah, J., and Abdennadher, S. (2021). Arcode: Programming for youngsters through ar. In *International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning*, pages 65–74. Springer.

Narman, H. S., Berry, C., Canfield, A., Carpenter, L., Giese, J., Loftus, N., and Schrader, I. (2020). Augmented reality for teaching data structures in computer science. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)*, pages 1–7. IEEE.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372.

Pirker, J., Kopf, J., Kainz, A., Dengel, A., and Buchbauer, B. (2021). The potential of virtual reality for computer science education-engaging students through immersive visualizations. In *2021 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pages 297–302. IEEE.

Sembey, R., Hoda, R., and Grundy, J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software*, 211:111988.

Su, S., Zhang, E., Denny, P., and Giacaman, N. (2021). A game-based approach for teaching algorithms and data structures using visualizations. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, pages 1128–1134.

Theodoropoulos, A. and Lepouras, G. (2021). Augmented reality and programming education: A systematic review. *International Journal of Child-Computer Interaction*, 30:100335.

Tsai, C.-Y. and Lai, Y.-C. (2022). Design and validation of an augmented reality teaching system for primary logic programming education. *sensors*, 22(1):389.

Van Mechelen, M., Smith, R. C., Schaper, M.-M., Tamashiro, M., Bilstrup, K.-E., Lunding, M., Graves Petersen, M., and Sejer Iversen, O. (2023). Emerging technologies in k–12 education: A future hci research agenda. *ACM Transactions on Computer-Human Interaction*, 30(3):1–40.

Villegas-Ch, W., Govea, J., Godoy, L. N., and Mera-Navarrete, A. (2024). Virtual reality simulations for skills training: Improving learning through immersive experiences in educational environments. *IEEE Access*.

Wei, Z., Liao, J., Lee, L.-H., Qu, H., and Xu, X. (2025). Towards enhanced learning through presence: A systematic review of presence in virtual reality across tasks and disciplines. *arXiv preprint arXiv:2504.13845*.