2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #51100

Balancing Immersion and Usability: Student Perspectives on XR and 2D Interfaces for Parse Tree Constructions

Ayano Nozawa, Bucknell University

Ayano Nozawa is a junior student at Bucknell University, majoring in Education.

Lea Wittie, Bucknell University

Lea Wittie is an Associate Professor and co-chair in the department of Computer Science in the Engineering College at Bucknell University.

Sing Chun Lee, Bucknell University

Sing Chun is a Computer Science Faculty at Bucknell University. During his PhD Study on Geometry Processing and Augmented Reality, he discovered his passionate interest in Computer Science Education. Currently, Sing Chun is investigating how Mixed Reality (XR) might help with Computer Science Education through the lens of multimedia learning theory and cognitive load theory. He is also interested in promoting undergraduate research and making advanced computational concepts more accessible to undergraduate students. In his spare time, he likes to immerse himself in the XR world and discuss the future of XR, particularly for education, with others.

Balancing Immersion and Usability: Student Perspectives on XR and 2D Interfaces for Parse Tree Constructions

Ayano Nozawa
an013@bucknell.edu
Education
Bucknell University

Lea Wittie
lwittie@bucknell.edu
Computer Science
Bucknell University

Sing Chun Lee
singchun.lee@bucknell.edu
Computer Science
Bucknell University

Abstract

Parse tree construction plays an essential role in understanding programming languages and compilers. Yet, many students find it challenging to construct parse trees from formal language grammars. Traditional instruction often relies on static, hand-drawn tree exercises that offer limited feedback and can discourage sustained practice. To improve engagement and scalability, this project investigates interactive tools that allow students to construct and evaluate parse trees in immersive and non-immersive environments. We developed a browser-based drag-and-drop tool and an immersive version of the same activity, and we compared them against an existing text-input editor baseline. Together, these interfaces allow students to experiment with parse tree construction through direct manipulation rather than solely symbolic reasoning.

Our initial classroom trial suggested that the immersive interface increased perceived engagement and motivation, while also surfacing usability, comfort, and visual clarity challenges that may affect sustained use. Building on these observations, the current study uses qualitative methods to examine how students experience and interpret the trade-offs between immersion and usability during parse tree construction. Data were collected through classroom observations and semi-structured interviews. Using thematic analysis, observation fieldnotes and interview transcripts were coded to identify recurring patterns in how learners describe attention and distraction, embodied interaction and effort, legibility and navigation, and the perceived necessity of immersive interfaces relative to non-immersive alternatives.

Special attention is given to how students describe balancing novelty and comfort, engagement and fatigue, and clarity and distraction while completing parse tree construction across modalities. The paper reports qualitative themes that explain when immersive interfaces are experienced as beneficial (*e.g.*, as an attentional enclosure that supports sustained focus and persistence) and when they are experienced as burdensome (*e.g.*, when legibility, navigation, or comfort constraints shift effort toward interface management). These findings inform design and pedagogical recommendations that foreground clarity, comfort, and accessibility.

1 Introduction

Parse trees (also called syntax trees), which represent how an expression can be derived from a grammar, are central to topics such as parsers, interpreters, and compilers. Although parse trees are a standard component of programming languages and compilers curricula, students often describe difficulty when moving from grammar rules to a correct hierarchical structure. Classroom practice commonly relies on hand-drawn trees. These approaches can support a conceptual introduction, but they may discourage sustained practice when feedback is delayed or when learners repeatedly make errors that obscure conceptual misunderstandings.

Interactive tools offer an alternative by providing structured construction, immediate validation, and opportunities for iterative corrections. In particular, drag-and-drop construction can externalize structure and reduce syntactic bookkeeping, potentially helping students focus on conceptual reasoning about grammar-derived structure. At the same time, immersive environments such as virtual reality (VR) can change how students attend to and interact with learning materials. Immersion may increase motivation and engagement, but it can also introduce practical barriers, including discomfort, legibility challenges, and interaction overhead. These trade-offs are especially noticeable for parse tree construction because the task is symbol-heavy, requires precision, and can become visually dense as trees grow larger. A prior classroom study compared interface modalities for the same parse-tree tasks and surfaced recurring tensions (*e.g.*, engagement vs. legibility/comfort) that motivate the qualitative focus of this paper[1].

This paper reports a qualitative investigation of students' experiences using three interfaces for parse tree construction: a structured text-input editor, a browser-based drag-and-drop tool, and an immersive version of the drag-and-drop tool. The text-input editor served as an existing, course-aligned baseline representation, while the browser-based and immersive interfaces were the project-developed environments for direct manipulation.

Rather than treating immersion as uniformly beneficial or harmful, this study focuses on how students describe the conditions and mechanisms that shape their experience, such as attention management, embodied interaction, legibility at scale, and perceived necessity. The contribution is an explanatory account of learner experience grounded in classroom observation and interviews, intended to support instructors in selecting, deploying, and scaffolding interactive tools for abstract syntax learning. The study is guided by the following research questions.

RQ1 How do students describe the benefits and drawbacks of immersive vs. non-immersive interfaces for parse tree construction in an authentic course setting?

RQ2 How do immersion and embodied interaction shape students' attention, physical effort, and strategies during parse tree construction?

RQ3 What conditions do students identify as necessary for sustained classroom use of immersive tools, including usability, comfort, accessibility, and perceived learning value?

2 Related Work

2.1 Interactive Support for Learning Grammars, Parsing, and Trees in CS Education

Formal languages and parsing topics are widely recognized as conceptually demanding, in part because they require students to coordinate symbolic rules with hierarchical structure. A long-standing approach for reducing this barrier is to provide interactive visualizations and editors that externalize structure and offer immediate feedback during practice. One well-known example is JFLAP, which supports interactive exploration of formal language concepts and includes grammar and parsing-related functionality; a multi-institution study reported that many students perceived

JFLAP as making concepts easier to learn and increasing engagement in formal language courses[2]. Another line of work focuses on helping students understand parsing procedures and intermediate representations. For instance, PAVT visualizes multiple parsing algorithms and provides step-by-step views intended to support learning in compiler construction contexts[3]. These algorithm visualizations and graphical tree editors allow students to construct parse trees directly, potentially reducing the burden and enabling more frequent practice.

Within this landscape, recent development efforts have explored autogradable, hands-on parse tree construction environments that provide structured interaction and automated validation to support practice at scale in classroom settings[4]. Our work builds on this direction by comparing multiple interaction modalities for the same learning goal and by centering qualitative evidence to explain the mechanisms behind perceived benefits and barriers. In closely related work, we compared multiple interface modalities for classroom parse tree construction and documented trade-offs between perceived engagement and practical constraints such as comfort and legibility[1]. The present study extends that work by using observations and interviews to explain when and why those trade-offs emerge in students' accounts.

2.2 Interface Modality: Text Input vs. Direct Manipulation

A recurring theme in computer science education tools is the contrast between text-based input and direct manipulation interfaces. Although parse tree construction is not identical to programming, the modality question is analogous: text-oriented representations can be compact and precise but can increase error-proneness and attention to surface syntax, whereas direct manipulation can offload notation management and highlight structure. In programming education, comparisons of isomorphic block-based and text-based environments provide evidence that modality can affect what learners attend to and the amount of extraneous effort required to make progress. For example, Weintrop and Wilensky report differential performance and interaction patterns across isomorphic block-based and text-based conditions, emphasizing that interface form can shape learning experience and outcomes even when underlying concepts are held constant[5]. Subsequent work examining longer-term transitions suggests that initial modality effects may diminish over time but can still influence learners' early experiences and attitudes[6]. These findings motivate careful evaluation of modality in symbol-heavy tasks such as parse tree construction, where students must coordinate grammar rules with a rapidly growing hierarchical representation.

2.3 Immersive Learning in VR: Engagement Benefits and Practical Trade-offs

Immersive virtual reality has been widely investigated as a way to increase engagement, presence, and interest, but the literature also cautions that these benefits can be accompanied by additional cognitive and practical costs. In a widely cited experiment, Makransky and colleagues found that adding immersive VR to a science lab simulation increased presence but led to poorer learning outcomes and higher cognitive load indicators, suggesting that immersion can introduce distraction or processing demands if not carefully designed[7]. Similarly, Parong and Mayer reported that immersive VR may increase motivation while also imposing higher demands that can hinder performance relative to non-immersive presentations[8]. To organize these mixed findings, Makransky and Petersen proposed the Cognitive Affective Model of Immersive Learning (CAMIL), which explains how immersion and interactivity shape learning through psychological affordances (presence and agency) and mediating factors (motivation, embodiment, cognitive load, and self-regulation)[9]. CAMIL is relevant for classroom deployments because it highlights that learning outcomes depend on how immersion and interaction design influence both affective and cognitive processes.

Within computer science (CS) education specifically, systematic reviews describe growing interest in VR for CS topics and emphasize both opportunities and constraints. Pirker and Dengel synthesize VR applications in CS education and note recurring challenges around usability, hardware constraints, and aligning interactions with learning objectives[10]. These observations align with practical issues that often emerge in classroom VR use, such as comfort, motion sensitivity, legibility of text and symbols, and the need to provide accessible alternatives.

Related immersive work also includes augmented reality approaches aimed at making abstract program structures visible in-situ. For example, AST[AR] explores using augmented reality and abstract syntax trees as a medium for teaching data structures to novice programmers, illustrating a broader interest in spatializing otherwise invisible computational representations[11]. Taken together, this literature motivates our focus on student experience: immersive interfaces may support engagement and attention, but their learning value in symbol-dense tasks depends on usability, legibility, and comfort conditions.

2.4 Qualitative Methods for Understanding Experience with Educational Technology

Because interface modality influences not only performance but also attention, affect, and perceived learning value, qualitative methods are essential for explaining how students experience tool-mediated learning in context. Thematic analysis is a widely used approach for identifying patterned meaning in open-ended responses and interviews; Braun and Clarke describe it as a flexible method for analyzing qualitative data while emphasizing methodological clarity in how themes are constructed and interpreted[12]. Within CS education, thematic analysis has been used to characterize learners' experiences with instructional interventions. For example, Celepkolu and Boyer analyzed student reflections on pair programming and reported themes spanning cognitive, affective, and social dimensions, demonstrating how qualitative evidence can reveal mechanisms that complement outcome measures[13]. For interview-based work, reporting standards can strengthen transparency and credibility. The consolidated criteria for reporting qualitative research (COREQ) checklist provides criteria for reporting qualitative studies based on interviews and focus groups, including research team reflexivity, study design, and analytic reporting[14]. In this paper, qualitative evidence is used not merely to catalog preferences but to explain trade-offs between immersion, usability, and conceptual clarity during parse tree construction.

3 Theoretical Framework

This study uses a trade-off lens to interpret students' accounts of learning with both immersive and non-immersive tools. The lens treats perceived learning value as shaped by multiple, sometimes competing, dimensions: perceived usefulness for understanding, perceived interaction effort, legibility and navigation constraints, and physical comfort and accessibility. To connect these perceived trade-offs to learning-relevant mechanisms, we also draw on CAMIL as a sensitizing framework. CAMIL highlights how immersion and interactivity operate through psychological affordances and mediating factors, which help interpret why a given interface may feel beneficial or demanding for symbol-dense tasks like parse tree constructions. Rather than assuming immersion is inherently beneficial, the lens foregrounds tensions that often surface in classroom deployments: engagement and motivation vs. usability and legibility; attentional support vs. embodied effort and fatigue; and novelty and enjoyment vs. perceived necessity and efficiency. These tensions guide our qualitative analysis and interpretation of students' reported experiences across the three interfaces described below.

4 The Tools in the Classroom

The intervention consists of three interfaces that support the same learning task: constructing a parse tree for an expression given a context-free grammar. The text-input interface uses structured bracket notation to represent a tree structure. The browser-based drag-and-drop interface parses the grammar and provides draggable nodes that students place into a partially scaffolded tree structure. The immersive interface on a standalone headset enables interaction with hands or controllers and offers a larger interaction space than a typical laptop screen. The text-input editor served as an existing baseline representation used in the course, while the browser-based and immersive interfaces were developed in this project to support direct manipulation and automated validation (see Figure 1). Students used the tools during class activities aligned with instruction on grammar and parsing. The tools were introduced with a brief orientation and example tasks. In immersive sessions, students were offered opt-out alternatives when they experienced discomfort.

5 Methodology

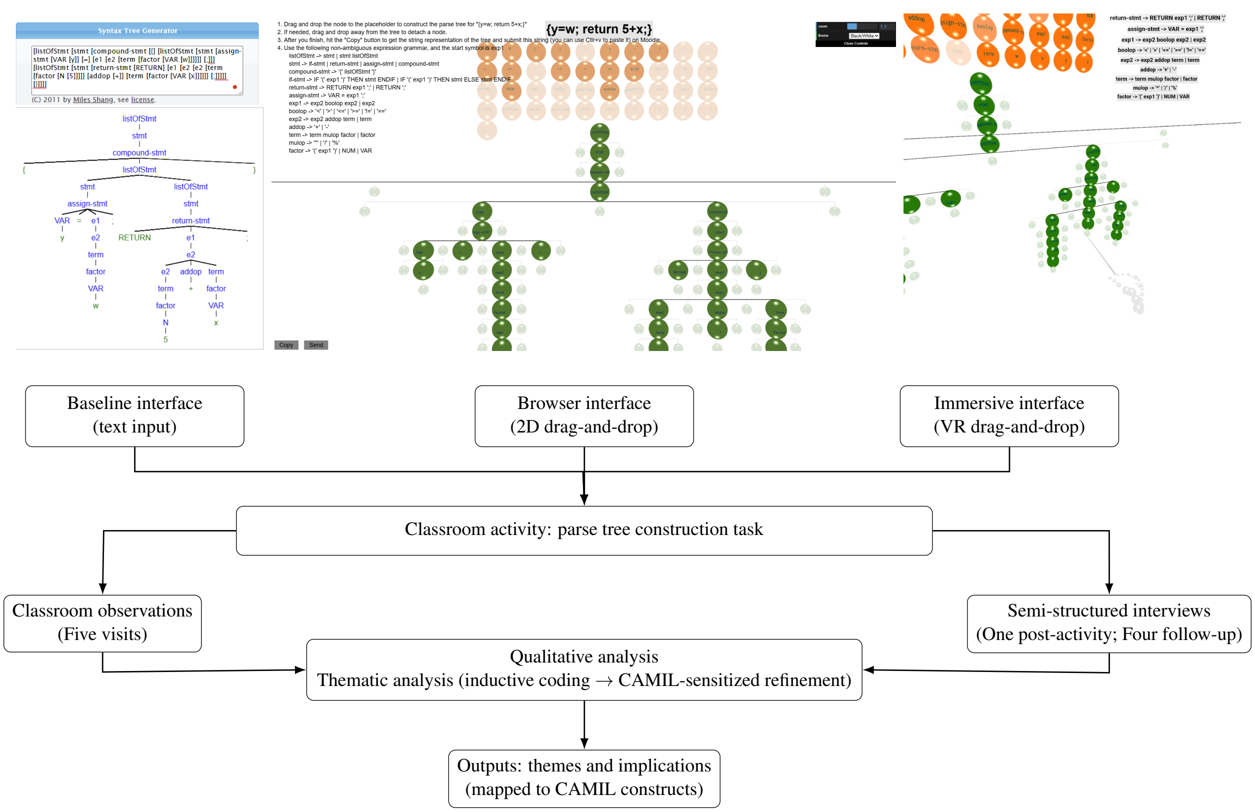

Figure 1: Overview of study context, interfaces, data collection, and analysis pipeline.

5.1 Qualitative Research Design

This paper reports a qualitative study designed to deepen understanding of students' experiences with immersive and non-immersive interfaces for parse tree construction. The design emphasizes explanation of learner experience in context, drawing on classroom observations and semi-structured interviews. The study seeks to identify conditions under which immersion supports or hinders students' engagement and perceived learning value. Figure 1 shows the three interfaces and summarizes the classroom context, interfaces, data sources, and analysis workflow.

5.2 Setting and Participants

The study took place in an upper-division programming languages course at a primarily undergraduate institution. Students were computer science majors who received course instruction on context-free grammars, parse trees, and related topics. The class activity involved constructing parse trees using three interfaces (baseline text input, 2D browser-based drag-and-drop, and an immersive VR drag-and-drop interface). The study protocol received Institutional Review Board (IRB) approval (IRB# 2526-033). Participation in interviews was voluntary and did not affect course grades. Participation was self-paced and students could opt out of immersive participation without penalty and complete the activity using alternative interfaces. Table 1 reports interface exposure time estimated from LMS activity logs.

Table 1: Tracked interface exposure for interview participants (self-paced activity).

Participant	Tracked Time[a] (Text/2D/VR)	Reported Prior VR Experience
Calvin	15m40s / 27m13s / 13m52s	Very little experience in his childhood.
James	19m46s / 21m33s / 25m28s	No prior experience.
Jude	10m53s / 2m44s / 36m37s	Pretty experienced with his own VR headset and VR games.
Kylie	23m40s / 52m51s / 0s[b]	Little experience with VR and motion sickness experience.
Simon	7m56s / 25m36s / 1h	No prior experience.

Exposure and comparability across interfaces. All interview participants had access to the baseline text, 2D browser, and VR interfaces as part of the course activity; however, because participation was self-paced and students could opt out of immersive participation without penalty, time spent in each interface varied across participants (Table 1). We therefore treat cross-interface comparisons in interview accounts as interpretive contrasts grounded in each participant's personal exposure history rather than as controlled equivalence across modalities. To reduce over-interpretation from uneven exposure, we triangulated interview claims with classroom observations and focus our findings on how students described interface effects (*e.g.*, attention, precision, comfort) rather than on prevalence claims.

[a]Times are estimated from LMS activity logs and reflect time the activity was open/active, not necessarily continuous attention or task completion.

[b]Kylie opted out VR due to motion sickness and chose to use the 2D browser-based drag-and-drop.

5.3 Data Collection

Data collection occurred over one academic semester and included repeated classroom observations and semi-structured interviews. The researcher conducted five in-class observations of regularly scheduled course meetings. The purpose of the observations was to document how students engaged with parse-tree-related activities and to capture contextual factors that could shape students' experiences with different interfaces. During each observation, the researcher recorded fieldnotes describing students' interactions, notable questions or comments, and observable behaviors related to attention, effort, and comfort. These fieldnotes were used to contextualize interview accounts and to support triangulation.

Following the observation period, the researcher invited students from the same course to participate in a research interview focused on immersive activities. Recruitment was conducted by announcing the opportunity to the class and requesting volunteers. In addition, the researcher conducted one post-activity interview with a volunteer who had participated in an earlier phase of the project, to capture reflective comparisons across experiences. In total, the research team conducted four immersive-activity interviews and one post-activity interview. All interviews were audio-recorded with participants' permission.

5.4 Transcription and Data Preparation

Interview recordings were transcribed using AI-assisted transcription software (noScribe[15] and aTrain[16]), followed by a human review process to improve accuracy. During verification, the research team compared outputs from two AI transcription tools and spot-checked transcripts against the audio to correct misheard technical terms and to ensure that critical moments were captured accurately. Verified transcripts, together with observation fieldnotes, were then coded for analysis.

5.5 Researcher Positionality and Reflexivity

The primary qualitative analyst served as a classroom observer and interviewer. Because qualitative interpretation is shaped by researchers' roles and assumptions, the research team used reflexive memoing throughout data collection and analysis to document evolving interpretations, questions, and potential biases. The team also discussed emergent themes during analysis meetings to surface alternative explanations and to ensure that themes were grounded in the data rather than in prior expectations about immersive technology.

5.6 Data Analysis and Trustworthiness

Qualitative analysis followed a thematic analysis approach[12]. We used a hybrid strategy: (1) inductive open coding to capture recurring patterns without imposing predefined categories, followed by (2) theory-sensitized refinement using the Cognitive Affective Model of Immersive Learning (CAMIL)[9]. After candidate themes were developed inductively, we used CAMIL to make interpretations explicit by noting, for each theme, the most relevant immersive affordance(s) (*e.g.*, presence, agency, or constraints on control) and learning-relevant mediator(s) (*e.g.*, motivation/interest, cognitive load, self-regulation, self-efficacy).

Coding procedure. After each observation, a researcher reviewed and coded the fieldnote and iteratively updated the codebook through memoing and constant comparison. Interview transcripts were coded using the same evolving codebook. Codes were grouped into candidate themes and reviewed against the full dataset for coherence and distinctiveness.

Codebook verification. In total, two researchers were involved in analytic checking: the primary analyst coded all five interview transcripts and all five observation fieldnotes. A second researcher conducted a peer debrief on the codebook and candidate themes, challenging definitions and boundaries; resulting revisions were documented in memos and incorporated into refinement. Disagreements were resolved through discussion.

Theme refinement. Themes were iteratively refined by comparing instances across the dataset, seeking converging support from at least two sources (*e.g.*, interviews and observations), and documenting divergences and deviant cases. During this step, CAMIL served as a sensitizing lens to sharpen theme labels and boundaries by explicitly linking each theme to its most salient affordances and mediators[9].

Trustworthiness. We strengthened credibility via triangulation (interviews and observations), attention to deviant cases, and an audit trail (fieldnotes, codebook versions, coding notes, and analytic memos) that enabled the team to revisit how interpretations were formed. Reporting is informed by qualitative reporting guidance - COREQ[14].

6 Findings

Across interviews and observations, students described immersive parse tree construction as engaging and attention-focusing, while also identifying constraints related to legibility, comfort, and perceived necessity. These themes align with prior immersive learning research that emphasizes both motivational affordances (*e.g.*, presence and interest) and practical or cognitive costs (*e.g.*, usability barriers and distraction)[7,8,9,10]. Below, we present four themes that explain how students experienced these trade-offs during parse tree practice. Because interface exposure differed across interview participants (Table 1), the themes below emphasize the mechanisms students described (attention, precision, comfort) rather than implying uniform effects across all modalities for all students. We interpret these themes using the Cognitive Affective Model of Immersive Learning (CAMIL)[9] to make explicit how reported experiences relate to immersive affordances (*e.g.*, presence and agency) and learning-relevant mediators (*e.g.*, motivation/interest, cognitive load, and self-regulation).

Theme 1: Immersion as an Attentional Enclosure

Students frequently described the immersive environment as "immersive" and connected that immersion to sustained attention during the activity. Observations showed that headset use reduced opportunities for off-task behavior, and interview accounts framed the immersive experience as a bounded workspace that supported concentration. This aligns with immersive learning theories that treat presence and agency as psychological affordances that can increase motivation and regulate attention[9]. One interviewee, James, explicitly described the immersive mode as academically helpful:

> "I liked it a lot. On an actual academic level, it was probably my favorite way to do VR or to do the parse trees assignment. I thought it was really immersive and helpful to be able to look at it through the VR headset." —*James*

Another interviewee, Calvin, similarly framed VR as a "training space", suggesting that the enclosed environment contributed to a task-focused mindset:

> "Just working on a laptop, I think it was very basic, very standard, but with the VR, it seemed like a much more immersive experience and it was much more pleasurable to do activities, because it seemed like you were in a training space." —*Calvin*

These accounts suggest that engagement was not merely novelty; students articulated immersion as a condition that altered how they attended to the task. This theme provides a plausible mechanism for why immersive interfaces may increase persistence during practice, consistent with prior work that links immersive affordances to affective and attentional outcomes. In CAMIL terms, students' descriptions align with presence supporting interest/motivation and self-regulation during practice.

Theme 2: Embodied Interaction as Both Engagement and Effort

Observation fieldnotes indicated that VR use involved more head, neck, and hand movement than laptop-based work, reflecting embodied navigation and object manipulation. Students described this embodied interaction as increasing interactivity and enjoyment, while also implying added physical effort. This dual characterization parallels prior CS education reviews that emphasize VR's potential for engagement while noting interaction overhead and ergonomic constraints[10]. Simon described the VR activity as engaging:

> "I thought the VR [version] was definitely more fun and it was my first time using VR, so that was also kind of a cool experience. I thought it was a lot more fun, because it felt more interactive." —*Simon*

At the same time, students' comments and observed behaviors suggest that embodied interaction can lead to fatigue and increased coordination demands. This pattern is consistent with immersive learning findings that increased presence may be accompanied by additional cognitive or physical load, depending on interaction design[7,8]. Students varied in how they weighed these demands against motivational benefits, indicating that embodied engagement is not uniformly positive and may depend on activity duration, task density, and individual comfort. Students also described strategy shifts, such as pausing to reorient nodes before placement and repeatedly repositioning their viewpoint to confirm parent-child relationships, indicating that the interaction strategy was shaped by the interface's embodied constraints. This reflects CAMIL's prediction that increased agency/embodiment can raise engagement while also increasing cognitive load through interaction demands.

Theme 3: Legibility and Navigation as Constraints for Symbol-Dense Tasks

A recurring barrier concerned legibility and navigation when trees became larger or when labels were difficult to read. This finding is consistent with the broader VR education literature that identifies readability and interaction precision as persistent challenges for tasks requiring fine-grained symbolic work[10]. Jude, an experienced VR user, described the difficulty of seeing node text and managing rotations and occlusion in a dense tree:

> "When you're looking at the tree, larger trees can extend, like, really far down. And it's hard to, it can be hard to see, like, what each node is. When one of the nodes was far down in the tree, it was kind of hard to see the text. which was kind of complicated by the fact that the nodes rotated as you, like, held them and stayed rotated, which meant the text could rotate kind of behind another node, or just behind something altogether." —*Jude*

Jude also described the zoom feature as insufficient for improving readability:

> "I couldn't really figure out how to zoom in. There was the option on the browser, but from what I could tell, it kind of just moved the tree closer and farther." —*Jude*

These accounts highlight a modality-specific constraint: parse tree construction is a symbol-dense task, and small readability issues can compound into workflow breakdowns. This theme also clarifies why the motivational benefits of immersion may not translate into perceived learning support when legibility and navigation impede sustained and accurate interaction. CAMIL suggests these breakdowns (legibility/navigation) can reduce effective agency and may increase extraneous cognitive load, undermining learners' confidence during this symbol-dense task.

Theme 4: Comfort, Access, and the Question of Necessity

Students' willingness to use VR depended on comfort and perceived necessity. Kylie opted out due to concern about motion sickness and described VR as inconvenient relative to the browser-based tool. Her account reflects common concerns in immersive learning research about accessibility, comfort, and the importance of providing alternatives[10]:

> "I think it's just more inconvenient to have to go and get a headset wear a headset figure out how to use it like if you can just do something in your browser that's a lot more efficient." —*Kylie*

> "It's way easier, more efficient to just do [the assignment] on your laptop." —*Kylie*

Jude similarly questioned whether VR added value when the activity was experienced as essentially "a 2D view" in a headset:

> "I do feel like [the assignment] basically just being a 2D view [in] the VR was kind of limiting. More simple browser graphics could help a lot with the visibility." —*Jude*

This theme aligns with prior findings that immersion can increase presence but may not improve learning if the interface introduces friction or distraction relative to a simpler presentation[7,8]. Importantly, student accounts treat "necessity" as part of perceived learning value: when VR does not offer a clear functional advantage for the task, students weigh setup effort and comfort costs more heavily. From the CAMIL lens, comfort/access constraints are "control factors" that can suppress motivation/interest and reduce persistence via reduced self-regulation.

7 Discussion

This qualitative study explains how students experience the trade-offs between immersive and non-immersive interfaces for a symbol-dense CS task. Rather than treating immersion as uniformly beneficial or harmful, students described conditional benefits and conditional costs.

RQ1: Benefits and Drawbacks of Immersive vs. Non-Immersive Interfaces

Taken together, the findings suggest that students evaluate immersive and non-immersive interfaces using two intertwined criteria: whether the interface supports sustained attention and motivation, and whether it preserves the perceptual and interaction precision required by symbol-dense work. Students' repeated use of "immersive" and their descriptions of being more "in" the task point to a benefit that is primarily affective and attentional: the headset can function as a bounded workspace that encourages persistence. At the same time, the drawbacks students emphasized were not generic complaints about VR but were tightly coupled to the representational demands of parse trees. Because tree construction requires reading many small labels, checking parent-child relationships, and revising structure iteratively, even modest breakdowns in legibility or view stability can shift the activity from reasoning about grammar-induced structure to managing the interface.

This framing helps clarify why immersive learning results can appear mixed in prior research. Immersion may increase presence and motivation, yet it can also introduce usability and perceptual costs that disrupt task flow[7,8]. In our context, the costs are amplified by the task's reliance on fine-grained text and dense spatial layouts: when labels rotate, occlude one another, or become difficult to inspect as trees grow, students interpret the immersive mode as "limiting" even when they still find it engaging. Conversely, the non-immersive interfaces are experienced as more dependable for clarity and precision, supporting sustained work when efficiency and readability dominate. This aligns with CS education literature, which emphasizes that the classroom value of VR depends less on immersion as a feature and more on how interaction design and hardware constraints align with the learning objective[10].

The contribution of this qualitative comparison is therefore not a claim that one modality is better, but an explanation of the trade-off students articulate: immersive interfaces are most valued when the attentional and motivational benefits are noticeable, and the interface preserves readability and controllable navigation; non-immersive interfaces are preferred when precision, speed, and stable inspection are central. This mechanism-oriented account supports conditional, task-sensitive adoption decisions rather than modality-driven ones. This suggests that modality comparisons in XR should be evaluated as a fit between representational demands (symbol density and precision reading) and interaction affordances, rather than as a general preference for immersion.

RQ2: Attention, Effort, and Strategy in Embodied Interaction

Across interviews and observations, immersion appeared to alter not only where students directed attention, but also how they organized their work. Students' descriptions of the headset as a bounded space suggest that XR can reduce exposure to competing stimuli, thereby supporting sustained engagement. This aligns with immersive learning models in which presence and agency are treated as psychological affordances that shape motivation and self-regulation[9]. In our context, these attentional benefits were experienced as a practical advantage during practice: students described being more "in" the task and less interrupted by external distractions.

At the same time, fieldnotes and interview accounts suggest that the mechanism that supports attention also changes the work itself. Embodied interaction introduces additional perceptual-motor demands, including coordinating head and hand movement, maintaining a stable view as the tree grows, and manipulating objects in space while monitoring label orientation and placement. These demands can shift students' strategies from primarily conceptual checking (matching grammar rules to hierarchical structure) toward tactical interface management (repositioning, re-orienting, and re-inspecting nodes). Students varied in how they weighed this shift. Some interpreted embodied interaction as increased interactivity that made the activity more engaging, while others experienced the additional coordination as friction that competed with reasoning. This tension helps clarify why prior VR studies sometimes report increased involvement alongside mixed or conditional learning-related experiences: immersion can heighten engagement while simultaneously adding extraneous demands that depend on interaction design and task density[7,8]. For symbol-dense tasks like parse tree construction, the balance appears especially sensitive to view stability and the precision of manipulation and inspection.

RQ3: Conditions for Sustainable Classroom Use

Students' accounts indicate that sustained classroom use depends on conditions that determine whether immersion remains supportive beyond initial engagement. A first condition is perceptual reliability: legibility and stable navigation must hold as trees scale. When students could not consistently read labels, keep nodes oriented, or maintain a workable view of large trees, they questioned whether the immersive modality was practical for continued use. A second condition is comfort and access. Some students opted out due to motion sickness concerns or discomfort, and others emphasized the efficiency and convenience of completing the activity on a laptop. In these accounts, accessibility is not an afterthought but a central determinant of whether XR can function as a classroom tool rather than a special experience.

A third condition is perceived necessity. Several students evaluated immersive use through a cost-benefit lens, weighing setup friction and comfort costs against what the modality uniquely contributed to their practice. When the activity felt functionally similar across modalities, students tended to prioritize clarity and efficiency; when immersion supported attention and persistence in a way that felt valuable, students were more willing to tolerate overhead. These conditions collectively support a multimodal adoption strategy: immersive use is most defensible when deployed in bounded sessions that leverage attentional enclosure and engagement, while non-immersive alternatives remain essential for accessibility, precision reading, and extended work[10,9]. Framed this way, "sustainable use" is not a single property of the technology but an alignment among task demands, interaction design, classroom logistics, and equitable participation pathways. In students'accounts, these conditions were often expressed as balances (novelty versus comfort, engagement versus fatigue, and clarity versus distraction) which helps explain when XR becomes supportive and when it becomes burdensome.

8 Design and Pedagogical Implications

This study's recommendations follow directly from students' descriptions of trade-offs between attentional benefits of immersion and practical constraints in legibility, comfort, and perceived necessity. The goal is not to position immersive interfaces as universally superior, but to clarify the design and instructional conditions under which immersive practice can support learning without introducing barriers.

From a design perspective, legibility and navigation should be treated as first-order learning supports for parse tree construction in immersive environments. Because parse trees are symbol-dense and require precise reading of labels, the interface should preserve text readability as trees scale and as nodes extend beyond the user's initial field of view. Students' accounts of occlusion and rotated labels suggest that the tool should minimize unintended rotations, reduce overlap, and provide stable viewpoints that allow learners to inspect large structures without losing context. View controls should support reliable zoom and repositioning that meaningfully improves readability rather than merely changing distance. In addition, because students often evaluate a tool by how efficiently it allows them to correct mistakes, immersive interaction should include low-friction undo, delete, and reattachment mechanisms, and should provide feedback that remains visible while keeping the workspace uncluttered. Taken together, these changes prioritize clarity and error recovery, reducing the risk that interface management competes with conceptual reasoning.

From a pedagogical perspective, the findings support integrating immersive practice as a targeted modality rather than as the only pathway to completing parse tree work. Students who described immersion as helpful often framed it as supporting sustained attention; however, students also emphasized that comfort and legibility constraints can become burdensome over time. As a result, instructors may achieve the greatest benefit by using immersive sessions in short, structured intervals that capitalize on attentional enclosure while minimizing fatigue and reducing the likelihood that readability issues accumulate as trees grow. One practical approach is to use immersive practice for early-stage construction and engagement, while encouraging browser-based continuation for detailed inspection, longer edits, or extended practice. This sequencing frames the two modalities as complementary: immersion supports focus and initial momentum, while the browser supports precision and sustained work.

Accessibility requires more than offering an opt-out; it requires an equivalent pathway that does not disadvantage students who cannot or prefer not to use a headset. The study suggests that instructors should explicitly normalize non-immersive participation and provide an alternative interface that supports comparable course participation and practice opportunities, without penalties or social pressure. Comfort protocols should be communicated as part of routine course logistics, including guidance on taking breaks, discontinuing use if nausea or strain occurs, and accessing alternative participation immediately. In addition, because students raised questions about the necessity of VR when the task appeared functionally similar across modalities, instructors can improve adoption by making the instructional rationale explicit: immersion supports attention, persistence, and engagement during practice, while the browser remains available for clarity, efficiency, and accessibility. Establishing this rationale in advance aligns student expectations with the intended learning role of each interface and reduces the risk that VR is perceived as novelty or unnecessary friction.

Overall, the implications emphasize alignment among tool design, instructional scaffolding,

Table 2: Summary of themes, student-reported experiences, CAMIL constructs, interpretation through the CAMIL lens, and resulting design guidelines.

Theme 1	**Immersion as an Attentional Enclosure**
Reported Experience	VR felt like a bounded workspace that reduced distraction and supported concentration.
CAMIL constructs	Presence as attention support $\rightarrow$ higher self-regulation and motivation/interest
Interpretation	Increased presence may support attention and persistence by reducing competing stimuli during practice.
Design guideline	Design VR task and views to preserve "attentional enclosure" by keeping the interface stable and visually uncluttered.
Theme 2	**Embodied Interaction as Both Engagement and Effort**
Reported Experience	Direct manipulation felt engaging but could add physical/interaction effort.
CAMIL constructs	Agency (with embodied interaction) $\rightarrow$ higher motivation/interest but added cognitive load (interaction overhead)
Interpretation	Higher agency/embodied action may raise engagement while also increasing cognitive demands.
Design guideline	Reduce interaction overhead so agency feels effortless (*e.g.*, reliable selection/manipulation and low-cost correction) rather than physically or cognitively taxing.
Theme 3	**Legibility and Navigation as Constraints for Symbol-Dense Task**
Reported Experience	Rotated/occluded labels and navigation issues made dense trees hard to read/manage.
CAMIL constructs	Constrained Agency (interaction control) $\rightarrow$ higher cognitive load and lower self-efficacy
Interpretation	When legibility/navigation constraints affect effective user control, learners may experience higher extraneous load and reduced confidence during complex edits.
Design guideline	Prioritize legibility and navigation support so students can accurately inspect and revise the structures without "fighting" the display.
Theme 4	**Comfort, Access, and the Question of Necessity**
Reported Experience	Comfort/setup/access issues and perceived redundancy shaped whether VR felt "worth it".
CAMIL constructs	Control factors constraint presence/agency $\rightarrow$ lower motivation and persistence
Interpretation	Comfort/access constraints may suppress motivation/persistence and limit when immersive benefits translate into perceived learning value.
Design guideline	Treat comfort and accessibility as core requirements so the immersive option does not reduce perceived value or exclude participants.

and accessibility planning. When clarity and correction are supported in the interface, and when classroom integration emphasizes bounded, purposeful use with equivalent alternatives, immersive practice is more likely to enhance engagement without compromising clarity, comfort, or accessibility. Table 2 provides a consolidated mapping from themes to CAMIL constructs to design guidelines, making the implications actionable for tool designers and instructors.

9 Limitations

This study has several limitations that bound the interpretation of the findings. First, the work is situated in a single course context at one institution, so transferability to other settings should be considered with care. The task itself (parse tree construction) is symbol-dense and may place unusually strong demands on legibility and precision, thereby shaping the immersion experience.

Second, the interview sample is small, and participation was voluntary, which may introduce self-selection bias. Students' exposure to each interface varied based on activity timing and opt-in decisions (see Table 1), which may shape the strength and specificity of cross-modality comparisons in participants' accounts. The themes are therefore presented as explanatory patterns grounded in participants' accounts rather than as prevalence claims about the broader student population.

Third, novelty and prior experience with XR may have influenced some responses. For some students, using a headset was new and potentially motivating; for others, prior discomfort could shape willingness to participate. Our qualitative focus captures these perceptions as part of the phenomenon, but it limits claims about how experiences may evolve with repeated, long-term exposure.

Finally, observation fieldnotes and thematic analysis provide rich contextual insight but remain interpretive. We mitigate these limitations through triangulation across interviews and observations, attention to deviant cases, reflexive memoing, and maintaining an audit trail of analytic decisions. Even so, the findings should be read as qualitative evidence about students' experiences and the conditions that shape perceived value and sustained use across modalities, rather than as claims about the frequency of themes in the full population.

10 Conclusion

This paper presents a qualitative account of students' experience across three interfaces for parse tree construction in a programming languages course: a structured text-input editor, a browser-based drag-and-drop tool, and an immersive version of the drag-and-drop activity. Students described immersion as engaging and attention-supporting, yet they also identified constraints in legibility, comfort, and perceived necessity that shaped whether XR felt practically valuable for sustained classroom use. Across interviews and observations, the themes show that the value of immersion is conditional: when the interface preserves clarity for symbol-dense work and supports efficient correction, XR can function as an attentional enclosure that encourages persistence; when readability and navigation break down, the interaction overhead can displace attention from the learning task toward interface management.

By mapping these experiences to an interface trade-off lens, the study reframes XR adoption decisions as designable and teachable conditions rather than as a binary judgment about whether "VR works". The implications emphasize (1) design priorities for XR parse tree tools, especially legibility-at-scale, stable navigation, and low-friction error recovery; (2) pedagogical integration strategies that treat XR as bounded, purposeful practice alongside non-immersive alternatives; and (3) accessibility practices that normalize equivalent pathways for students who cannot or prefer not to use a headset. Taken together, the findings provide an explanatory basis for instructors and tool builders to incorporate immersive and non-immersive interfaces in ways that preserve clarity, comfort, and accessibility while leveraging the engagement benefits students associate with immersion. More broadly, the study suggests that evaluating XR for computer science education should attend to task-specific perceptual demands (such as symbol-dense tasks) and to the classroom conditions that determine whether immersion is experienced as beneficial or burdensome in practice.

Acknowledgment

We thank the students who volunteered their time to participate in interviews and share their learning experiences. We also thank the course instructors and teaching staff for supporting classroom data collection and facilitating study logistics.

References

[1] S. C. Lee, C. J. Soule, and L. Wittie, "Immersive vs. 2D interfaces for parse tree construction: An exploratory study of engagement, workload, and design trade-offs," in *Proceedings of the 12th International Conference of the Immersive Learning Research Network (iLRN)*, Jun 2026, to appear.

[2] S. H. Rodger, E. Wiebe, K. M. Lee, C. Morgan, K. Omar, and J. Su, "Increasing engagement in automata theory with JFLAP," *SIGCSE Bull.*, vol. 41, no. 1, p. 403–407, Mar 2009. [Online]. Available: https://doi.org/10.1145/1539024.1509011

[3] S. Sangal, S. Kataria, T. Tyagi, N. Gupta, Y. Kirtani, S. Agrawal, and P. Chakraborty, "PAVT: a tool to visualize and teach parsing algorithms," *Education and Information Technologies*, vol. 23, no. 6, p. 2737–2764, Nov 2018. [Online]. Available: https://doi.org/10.1007/s10639-018-9739-x

[4] C. J. Soule, L. Wittie, and S. C. Lee, "BOARD # 84: WIP: Auto-gradable hands-on parse tree learning tool in virtual reality," in *2025 ASEE Annual Conference & Exposition*. Montreal, Quebec, Canada: ASEE Conferences, Jun 2025. [Online]. Available: https://doi.org/10.18260/1-2--55901

[5] D. Weintrop and U. Wilensky, "Comparing block-based and text-based programming in high school computer science classrooms," *ACM Trans. Comput. Educ.*, vol. 18, no. 1, Oct 2017. [Online]. Available: https://doi.org/10.1145/3089799

[6] D. Weintrop and U. Wilensky, "Transitioning from introductory block-based and text-based environments to professional programming languages in high school computer science classrooms," *Comput. Educ.*, vol. 142, no. C, Dec. 2019. [Online]. Available: https://doi.org/10.1016/j.compedu.2019.103646

[7] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," *Learning and Instruction*, vol. 60, pp. 225–236, Dec 2019. [Online]. Available: https://doi.org/10.1016/j.learninstruc.2017.12.007

[8] J. Parong and R. E. Mayer, "Learning science in immersive virtual reality," *Journal of Educational Psychology*, vol. 110, no. 6, pp. 785–797, 2018. [Online]. Available: https://doi.org/10.1037/edu0000241

[9] G. Makransky and G. B. Petersen, “The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality,” *Educational Psychology Review*, vol. 33, no. 3, pp. 937–958, Jan 2021. [Online]. Available: https://doi.org/10.1007/s10648-020-09586-2

[10] J. Pirker, A. Dengel, M. Holly, and S. Safikhani, “Virtual reality in computer science education: A systematic review,” in *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST ’20. New York, NY, USA: Association for Computing Machinery, Nov 2020. [Online]. Available: https://doi.org/10.1145/3385956.3418947

[11] V. Agrahari and S. Chimalakonda, “AST[AR] – towards using augmented reality and abstract syntax trees for teaching data structures to novice programmers,” in *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, Jul 2020, pp. 311–315. [Online]. Available: https://doi.org/10.1109/ICALT49669.2020.00100

[12] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, Jul 2008. [Online]. Available: https://doi.org/10.1191/1478088706qp063oa

[13] M. Celepkolu and K. E. Boyer, “Thematic analysis of students’ reflections on pair programming in cs1,” in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE ’18. New York, NY, USA: Association for Computing Machinery, Feb 2018, p. 771–776. [Online]. Available: https://doi.org/10.1145/3159450.3159516

[14] A. Tong, P. Sainsbury, and J. Craig, “Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups,” *International Journal for Quality in Health Care*, vol. 19, no. 6, pp. 349–357, Sep 2007. [Online]. Available: https://doi.org/10.1093/intqhc/mzm042

[15] K. Dröge and contributors, “noScribe: Ai-powered audio transcription,” 2026, accessed 2026-01-17. [Online]. Available: https://github.com/kaixxx/noScribe

[16] A. Haberl, J. Fleiß, D. Kowald, and S. Thalmann, “Take the aTrain. introducing an interface for the accessible transcription of interviews,” *Journal of Behavioral and Experimental Finance*, vol. 41, p. 100891, Jan 2024. [Online]. Available: https://doi.org/10.1016/j.jbef.2024.100891