Considering Large Language Model Integration in Expressive Computer Science Learning Environments for Blind and Visually Impaired Learners Through Co-design

Shi Ding[1(✉)], Jason Brent Smith[2], and Brian Magerko[1]

[1] Georgia Institute of Technology, Atlanta, GA 30308, USA
sding84@gatech.edu
[2] Northwestern University, Evanston, IL 60208, USA

Abstract. Although large language models (LLMs) have transformed the way people can interact with information and receive personalized help in learning spaces, we have a limited understanding of how blind and visually impaired (BVI) learners use these LLMs-powered technologies. This paper explores the potential of LLMs to improve the BVI learners' experience in EarSketch, a music coding platform designed to foster computational thinking through creative expression. Through co-design sessions with BVI learners and students, we identify key challenges such as intuitive search, navigation, and feedback, and propose LLM-based solutions to address these barriers. We synthesize key design implications with broader relevance for expressive computer science learning environments.

Keywords: Human Centered Computing · Large Language Models · Learning System · Accessibility

1 Introduction

Expressive computer science (CS) learning environments (e.g. Scratch) promote computational thinking and creative expression by combining coding with music, gameplay, and storytelling [20]. EarSketch[1], for example, enables users to create music by coding in Python or JavaScript, leveraging a large sound library and a custom API to arrange and manipulate audio loops. It teaches music and coding concepts through a web-based curriculum [4]. However, the design of such platforms is highly dependent on visual elements, making them less accessible to diverse learners [3]. Figure 1 in the Appendix shows the EarSketch interface.

The recent rise of generative artificial intelligence (AI) has transformed educational experiences across domains. Among these advancements, large language

[1] https://earsketch.gatech.edu/.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. I. Cristea et al. (Eds.): AIED 2025, LNAI 15880, pp. 472–480, 2025.
https://doi.org/10.1007/978-3-031-98459-4_36

models (LLMs) have attracted attention due to their potential to support personalized and adaptive learning [22]. Their integration into an expressive CS learning environment marks a meaningful opportunity to enhance learning interactions and promote AI literacy [7]. However, this shift also raises critical questions about how and to what extent LLMs can meaningfully enhance creativity and learning outcomes for diverse populations within interactive learning settings.

Our research aims to investigate how LLMs can improve learning experiences in an expressive CS learning environment through a case study of EarSketch with the following research question: **What constraints do current blind or visually impaired (BVI) learners encounter when integrating LLMs into an expressive CS learning environment?** To address this question, we conducted co-design studies with teachers and BVI students to identify key design implications for using LLMs in accessible expressive CS learning contexts. Building on these findings, we hope to develop a universal design framework to advance accessible AI in expressive CS learning environments, aligning with the diverse needs of all learners and with universal design learning principles [19].

2 Related Work

Recent research explores how AI can improve programming learning tasks by enhancing learning experiences in tools such as intelligent tutoring [10] and automated grading [24]. AI tools have shown promise in improving self-efficacy in students [18,31]. However, many existing AI systems focus primarily on code accuracy, often overlook deeper learning needs such as conceptual understanding, differentiated and accessible support for diverse learners. There is a critical need to examine how AI-human collaboration can enhance the efficiency and overall programming learning experiences in diverse teaching ecosystems [12]. Our research aims to encourage learners to think collaboratively with AI, rather than using AI solely as a solution provider.

As an expressive CS learning platform, EarSketch allows AI assistants to enhance learners' coding skills. Previous studies evaluated an AI agent in EarSketch, measuring its ability to support users as co-creative learning companions by advancing knowledge of coding concepts [25], using authored dialogue trees developed by analyzing co-creative dialogue between human users [6]. In our research, we aim to explore how LLMs can offer an AI agent higher degrees of computational creativity support than a menu-based system by generating contextual and adaptive dialogue based on learners' profiles(e.g., BVI vs. sighted learners) within EarSketch.

LLMs integrated into information retrieval (IR) systems have been used to redefine traditional search stacks [28] and improve exercise retrieval for language learning [30]. As LLMs evolve, they offer potential improvements in query understanding, retrieval, ranking, and question-answering components of IR systems, paving the way for more intelligent and context-aware search experiences [33]. Our study investigates unexplored contextual and multimodal interactions with LLMs to upgrade the design of IR systems.

In programming education, LLMs can generate improved error messages and offer multimodal feedback such as visuals and audio cues [1,26]. They can also create programming exercises and generate pseudo-code with line-by-line explanations [9,23]. Researchers have developed accessibility-focused interfaces integrating block-based, auditory, and tangible approaches [5]. However, challenges remain in ensuring environment-grounded outputs and fine-grained guidance [32]. Future studies should emphasize the need for strategies to help BVI learners understand code structure [8]. We aim to develop advanced scaffolding approaches that ensure ethically grounded support for BVI learners.

Previous studies have used innovative tools to create an inclusive learning environment, but often overlook sociotechnical constraints [29]. Active user involvement and viewing users as experts in the co-design process [17] has led to outcomes that better meet user needs and contexts [13,16]. Prior work has engaged BVI children and adults as co-design partners [14], but work with BVI teens has been limited [15]. To address this gap, we engage BVI teens in expressive CS learning environments as co-design partners.

3 Methodology

We conducted two co-design sessions where students and teachers interacted with EarSketch without LLM integration. Participants collaboratively identified challenges and suggested features related to accessibility, navigation, and learning outcomes. Insights from these discussions will directly inform the development of LLM-powered enhancements for EarSketch.

3.1 Participant Recruitment

This study was approved by the Institutional Review Board (IRB), and all participants provided informed consent before participation and compensation after the co-design sessions. We collaborated with five teachers ($n = 5$) and six students ($n = 6$) as participants at the California School for the Blind (CSB), a K-12 school specialized in BVI education. The teachers (T1-T5) were selected based on their diverse expertise and experience; they include T1, a female Curriculum Specialist; T2 and T3, two female Assistive Technology (AT) Specialists; T4, one male Assistive Technology (AT) Specialists; and T5, one female Teacher of the Visually Impaired (TVI). Among them, T3 had a visual impairment, and T5 was fully blind. Student participants (S1-S6, 5 male, 1 female) in this study have various levels of visual impairment and experience with assistive technology. Among the six student participants, whose ages ranged from 15 to 18, S1, S2, and S3 exclusively used screen readers, while S4 (female) and S5 exclusively used screen magnifiers, and S6 alternated between reader and magnifier.

3.2 Co-design Procedures

The co-design study with teachers included multiple remote asynchronous sessions followed by an in-person co-design session. Teacher participants began with

an EarSketch onboarding tutorial, created a music project as a homework task, and participated in individual 45-minute Zoom interviews to share their overall experiences. The teachers were then invited to discuss adapting EarSketch for a BVI-focused curriculum in a three-hour in-person co-design workshop, which included group discussions facilitated by a virtual Miro board[2], as well as Think-Pair-Share [2] and User Needs Analysis [11] activities.

The student co-design study spanned four in-person days. During the first three days, students participated in one-hour guided activities with EarSketch to practice experimenting with sound arrangements and creating music projects. On the final day, students participated in a co-design workshop, brainstorming and prototyping future EarSketch interfaces using tactile materials and speculative design activities. They compared EarSketch to other tools they use and shared ideas to create a more accessible and engaging learning environment. Table 1 in the Appendix shows the list of activities and data collected per group.

Both students and teachers discussed LLMs and their potential integration into EarSketch from various perspectives. These discussions highlighted LLMs as educational tools and drove us to incorporate AI-related inquiries into our future design considerations. We documented student artifacts through photographs, recorded teachers' comments as digital "sticky notes" and transcribed interview data and observation notes to inform future design improvements.

3.3 Data Analysis

We used a qualitative data analysis approach [21] to obtain comprehensive and in-depth insights due to the small sample size. We first collaboratively reviewed all collected data, including observational and interview notes from co-design sessions and interview transcripts. We used an existing codebook, initially developed after our teacher co-design session, structured into four main categories: workflow, navigation, feedback, and learning context. Ultimately, we synthesized our findings into several design implications for integrating LLMs into expressive CS learning environments based on feedback from both teachers and students.

4 Design Implications

4.1 Implication 1: Accessing Media Assets

Several key themes emerged regarding universal search and navigation during the co-design session with students: S4 stated that she prefers to have universal search in EarSketch to quickly look up songs, or navigate across the site. Both S5 and S6 reiterate their needs for better focus and announcement of sections, especially when using screen readers. They noted that when the cursor is on the first line of the code editor, the screen reader provides too much information, making navigation overwhelming. Students found the content manager (see 1 in the Appendix) difficult to navigate, and with too much information to parse, they

[2] https://miro.com.

would like a shorter path to finding the sounds they want. S3 said, *"Universal navigation is all headings and is where you can select Sound Library, Projects, etc. Once selected, it will automatically open the relevant contents of the selected section into the main page."* S1 suggested that AI-powered search tools find a sound rather than have it provide suggestions like *"Here are similar sounds to the one you selected."* S5 shared the importance of *"sharing sounds, searching using AI for other people's projects (by region or sound type)."* S4 also emphasized the need for clearer search functionalities, including filters for SoundCloud tabs and regional sound variations, while expressing concerns about ensuring learners cannot directly copy and paste others' code. This feedback highlights the need for intuitive navigation and search capabilities to support learners in exploring the platform effectively. Enhancing these features through LLMs integration could significantly improve usability, accessibility, and engagement for all users, particularly those relying on assistive technologies.

To ease workflow and enhance task completion, the ideal integration of LLMs should allow users to access and retrieve structured information efficiently. In expressive CS learning environments like EarSketch, improving access to media resources such as songs and sound libraries can significantly enhance the learning experience. By making it easier to find relevant content, the platform can further empower learners' agency to navigate the site with more independence.

4.2 Implication 2: Switching Between Curriculum and Authoring Environment

Learners faced significant challenges when switching between different components of the EarSketch interface, such as the curriculum, API, and code editor. T4 emphasized the importance of structured guidance when switching tools and activities, stating, *"The chatbot could give me all the steps and keyboard commands for how to check my Outlook email using JAWS (a popular screen reader)."* Screen reader user S2 required extensive 1:1 assistance to write a call to `fitMedia()`, an EarSketch API function designed to place an audio sample on a track for playback. He struggled with syntax, capitalization, punctuation, and function parameters. He required character-by-character instructions (including capitalization), where to find commas and parentheses on the keyboard, and what each argument means. Without structured support, the process became about memorizing steps rather than understanding.

Integrating LLMs into EarSketch can help bridge this gap by providing step-by-step guidance and contextual reminders during transitions between tools and activities. For example, LLMs can remind BVI learners that `fitMedia()` requires having both a start and an end time during the transition so that the learners can explore the platform more independently. Additionally, integrating detailed instructions and keyboard shortcuts for screen readers (like JAWS) would enhance accessibility, enabling BVI learners to navigate EarSketch more independently. LLMs can promote a more intuitive learning experience and a deeper engagement with BVI learners through a more seamless transition between learning resources and coding tasks.

4.3 Implication 3: Provide Suggestions (instead of Complete Answers) to Empower Learner Agency

Teachers and students emphasized the importance of AI fostering independent problem-solving rather than providing direct answers. T4 explained *"If the AI just builds out the code for them, I don't know that would be worthwhile... But if it gave suggestions, that would be pretty cool. Students should be experimenting themselves."* T2 highlighted the value of guidance over completion: *"They could ask an accessibility question like, 'how do I navigate by block?' Or, 'I'm coming across this issue, How do I do that?"'* Student S4 echoed these concerns, sharing, *"I use AI in school, but I don't want to copy and paste and not learn anything."* These insights align with Vygotsky's Zone of Proximal Development [27], the distance between what learners can do independently and what they can do with appropriate assistance. LLMs should guide learners through scaffolding toward solutions rather than simply providing answers.

LLMs should prioritize generating suggestions over solutions to support meaningful and effective learning. By transforming code generation requests into problem-solving prompts and accessible cues such as spoken hints, AI can encourage exploration while maintaining necessary guardrails. This ensures BVI learners remain actively engaged through deeper understanding rather than passive reliance on AI-generated content.

5 Conclusion and Future Work

This study explores how LLMs can support BVI learners in an expressive CS learning environment by redesigning the EarSketch platform. Through co-design with teachers and students, our initial results highlight that integrating LLMs into EarSketch can enhance learning and exploration for learners who are BVI.

Our co-design sessions were conducted with a small sample size at a single BVI-focused institution, which may limit generalizability. Co-designing with a more diverse range of participants could help ensure the adaptability of our work. As a preliminary experiment, we developed a GPT-4-based chatbot integrating the EarSketch curriculum to assist with coding and music queries. Early testing showed improved workflow and feedback, but revealed interface disruptions and reliability issues. Future iterations will enhance accessibility, learner agency, and AI-driven navigation. We are exploring plugins to refine search, navigation, and feedback for a more engaging learning experience. Future evaluation will assess interactions between BVI learners and AI, feedback reliability, and overall usability; These considerations represent critical next steps in the future.

Acknowledgments. This material is based on work supported by the National Science Foundation Awards Nos. 2300631 and 2300633. The views expressed are those of the authors and do not necessarily reflect those of the NSF. EarSketch is available online at https://earsketch.gatech.edu/.

A Appendix 1: EarSketch Interface

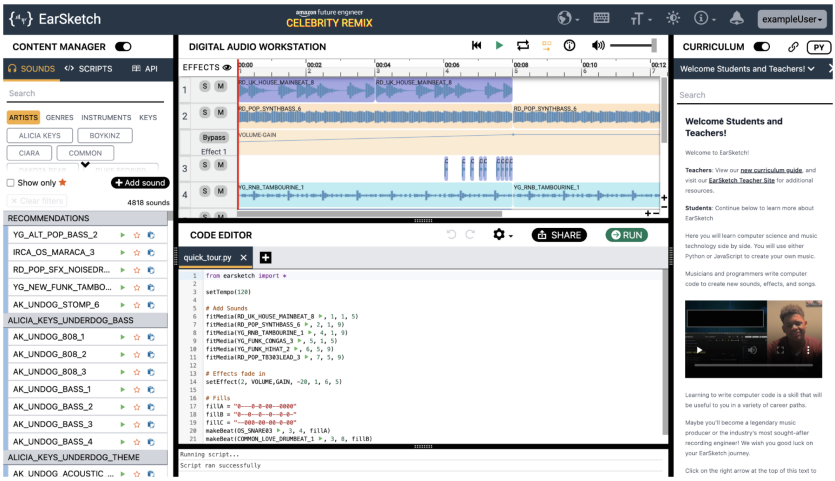

Fig. 1. The EarSketch web client interface with a) the Content Manager (left)–containing the Sound Browser, Script Browser, and API; b) a Digital Audio Workstation (top); c) the Code Editor (bottom); and d) the Curriculum (right).

B Appendix 2: Study Design

Table 1. Activities and data collected during the teacher and student co-design studies.

Group	Activity	Data Collected
Teacher	Homework/Interviews, Think-Pair-Share, Group Discussion, User Needs Analysis	Miro Board Artifacts, Conversational Logs
Student	ES Introduction, Think-Pair-Share, Prototype Design, Speculative Design	Observation Notes, Design Images/Descriptions

References

1. Ali, F., Choy, D., Divaharan, S., Tay, H.Y., Chen, W.: Supporting self-directed learning and self-assessment using teachergaia, a generative AI chatbot application: learning approaches and prompt engineering. Learn. Res. Pract. **9**(2), 135–147 (2023)

2. Chaffer, J., Nashton, S., Simmons, D.A.: Think-pair-share (TPS) in a virtual workshop setting (2021)
3. Ding, S., Smith, J.B., Garrett, S., Magerko, B.: Redesigning Earsketch for inclusive CS education: A participatory design approach. In: Proceedings of the 23rd Annual ACM Interaction Design and Children Conference, pp. 720–724 (2024)
4. Garrett, S., et al.: Improving the accessibility of the EarSketch web-based audio application for blind and visually impaired learners. In: Proceedings of the International Web Audio Conference. WAC 2024
5. Goolsby, B., Pawluk, D., Kim, H.W., Fusco, G.: A tangible block editor for the scratch programming language. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2021)
6. Griffith, A.E., et al.: Discovering co-creative dialogue states during collaborative learning. In: Roll, I., McNamara, D., Sosnovsky, S., Luckin, R., Dimitrova, V. (eds.) AIED 2021. LNCS (LNAI), vol. 12748, pp. 165–177. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78292-4_14
7. Grover, S.: Teaching AI to k-12 learners: lessons, issues, and guidance. In: Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1, pp. 422–428 (2024)
8. Hadwen-Bennett, A., Sentance, S., Morrison, C.: Making programming accessible to learners with visual impairments: a literature review. Int. J. Comput. Sci. Educ. Schools **2**(2), 3–13 (2018)
9. Kazemitabaar, M., et al.: CodeAid: evaluating a classroom deployment of an LLM-based programming assistant that balances student and educator needs. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–20 (2024)
10. Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.A.: Intelligent tutoring goes to school in the big city. Int. J. Artif. Intell. Educ. **8**, 30–43 (1997)
11. Lindgaard, G., et al.: User needs analysis and requirements engineering: theory and practice. Interact. Comput. **18**(1), 47–70 (2006)
12. Liu, J., Li, S.: Toward artificial intelligence-human paired programming: a review of the educational applications and research on artificial intelligence code-generation tools. J. Educ. Comput. Res. **62**, 07356331241240460 (2024)
13. Magnusson, C., Hedvall, P.O., Caltenco, H.: Co-designing together with persons with visual impairments. In: Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies, pp. 411–434 (2018)
14. McElligott, J., Van Leeuwen, L.: Designing sound tools and toys for blind and visually impaired children. In: Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community, pp. 65–72 (2004)
15. Metatla, O., Oldfield, A., Ahmed, T., Vafeas, A., Miglani, S.: Voice user interfaces in schools: co-designing for inclusion with visually-impaired and sighted pupils. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2019)
16. Muller, M.J., Wildman, D.M., White, E.A.: Taxonomy of PD practices: a brief practitioner's guide. Commun. ACM **36**(6), 26–28 (1993)
17. Penuel, W.R., Roschelle, J., Shechtman, N.: Designing formative assessment software with teachers: an analysis of the co-design process. Res. Pract. Technol. Enhanc. Learn. **2**(01), 51–74 (2007)
18. Phung, T., et al.: Generative AI for programming education: benchmarking ChatGPT, GPT-4, and human tutors. In: Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 2, pp. 41–42 (2023)

19. Rao, K., Garant-Jones, E., Dean, B.A., Eady, M.J.: Applying universal design for learning to work-integrated learning: designing for inclusion and equity. Int. J. Work-Integr. Learn. **25**(1), 67–81 (2024)
20. Resnick, M., et al.: Scratch: programming for all. Commun. ACM **52**(11), 60–67 (2009)
21. Saldaña, J.: The coding manual for qualitative researchers (2021)
22. Salemi, A., Mysore, S., Bendersky, M., Zamani, H.: LAMP: when large language models meet personalization. arXiv preprint arXiv:2304.11406 (2023)
23. Sarsa, S., Denny, P., Hellas, A., Leinonen, J.: Automatic generation of programming exercises and code explanations using large language models. In: Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1, pp. 27–43 (2022)
24. Schneider, J., Richner, R., Riser, M.: Towards trustworthy autograding of short, multi-lingual, multi-type answers. Int. J. Artif. Intell. Educ. **33**(1), 88–118 (2023)
25. Smith, J., Truesdell, E.J., Freeman, J., Magerko, B., Boyer, K.E., McKlin, T.: Modeling music and code knowledge to support a co-creative AI agent for education. In: Proceedings of the 20th International Society for Music Information Retrieval Conference, pp. 134–141 (2020)
26. Stamper, J., Xiao, R., Hou, X.: Enhancing LLM-based feedback: insights from intelligent tutoring systems and the learning sciences. In: Olney, A.M., Chounta, IA., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds.) International Conference on Artificial Intelligence in Education, pp. 32–43. Springer (2024). https://doi.org/10.1007/978-3-031-64315-6_3
27. Vygotsky, L.S.: Mind in Society: The Development of Higher Psychological Processes, vol. 86. Harvard university press (1978)
28. Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., Wei, F.: Large search model: redefining search stack in the era of LLMs. In: ACM SIGIR Forum, vol. 57, pp. 1–16. ACM New York, NY, USA (2024)
29. Windschitl, M., Sahl, K.: Tracing teachers' use of technology in a laptop computer school: the interplay of teacher beliefs, social dynamics, and institutional culture. Am. Educ. Res. J. **39**(1), 165–205 (2002)
30. Xu, A., Monroe, W., Bicknell, K.: Large language model augmented exercise retrieval for personalized language learning. In: Proceedings of the 14th Learning Analytics and Knowledge Conference, pp. 284–294 (2024)
31. Yilmaz, R., Yilmaz, F.: The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. Comput. Educ. Artif. Intell. **4**, 100147 (2023)
32. Zhao, Y., Zhang, Y., Xiang, R., Li, J., Li, H.: VIALM: a survey and benchmark of visually impaired assistance with large models. arXiv preprint arXiv:2402.01735 (2024)
33. Zhu, Y., et al.: Large language models for information retrieval: a survey. arXiv preprint arXiv:2308.07107 (2023)