Bridging AI and Human Feedback: Hybrid Intelligence in Embodied Math Education

Giulia Cosentino[1], Jacqueline Anton[2,3], Kshitij Sharma[1], Mirko Gelsomini[1,4], Michail Giannakos[1] and Dor Abrahamson[2,5]

[1]Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
[2]Berkeley School of Education, University of California Berkeley, Berkeley, USA
[3]Department of Special Education, San Francisco State University, San Francisco, USA
[4]Department of Innovative Technologies, Scuola Universitaria Professionale della Svizzera Italiana, Lugano, Wwitzerland
[5]College of Education, Korea University, Seoul, South Korea

Abstract

This paper explores the role of generative AI (GenAI) in providing adaptive summative feedback within an embodied learning environment for children's mathematics education. Using a body-scale digital number line, children engaged in learning integer operations through physical interaction. The study employed a between-group design: one group received feedback from a human instructor, while the other received AI-generated feedback. A mixed-method approach combined multimodal data (system logs, motion sensors) with qualitative observations of student interactions. The results showed no significant differences in task performance but revealed key differences in engagement: the teacher feedback encouraged multimodal, reflective responses involving gestures and body movements, while the AI feedback promoted streamlined, task-specific strategies with reduced cognitive load. These findings demonstrate the complementary strengths of human and AI feedback, underscoring the potential of hybrid intelligence systems to enhance adaptive learning environments.

Keywords

Hybrid intelligence, Generative AI, Teacher-AI collaboration, Summative feedback, Embodied learning

1. Introduction

Embodied learning, rooted in embodied cognition theory, links cognitive processes with physical interactions in the environment [1, 2]. It is particularly effective in education, where engaging physically with learning content enhances conceptual understanding, especially in abstract subjects like mathematics [3, 4]. Multisensory environments (MSEs) amplify these benefits by integrating visual, auditory, and kinesthetic stimuli, fostering immersive and interactive learning experiences [5]. Research demonstrates that MSEs significantly improve engagement and learning outcomes [6, 7]. However, the integration of advanced technologies such as Generative AI (GenAI) into these frameworks remains underexplored.

GenAI offers the potential to enhance MSEs by providing personalized feedback, addressing cognitive and attentional needs, and managing challenges such as cognitive overload in real-time [8, 9, 10, 11]. This study investigates the impact of GenAI-generated summative feedback on learning mathematics through embodied interaction with a digital number line (NL). A between-group design was used: the control group received feedback from a human instructor, while the experimental group received feedback from a large language model (LLM) informed by students' movement data. Qualitative observations combined with eye-tracking and motion data logs provided insight into learning behaviors and cognitive engagement. This research addresses the following research question:

RQ: How do teacher and GenAI feedback differ in shaping cognitive engagement, task efficiency, and multimodal interaction in embodied learning environments, and what is the potential for hybrid integration (with a focus on expanding teachers' skills and not replacing them)?

By examining the benefits and limitations of GenAI feedback, this study contributes to understanding its role in education, highlighting implications for educators, researchers, and developers aiming to design hybrid intelligence systems that enhance learning experiences [12]. In our work, we promote a

Woodstock'22: Symposium on the irreproducible science, June 07–11, 2022, Woodstock, NY

human-centered hybrid intelligence approach by investigating ways of combining teachers' expertise with AI capabilities, and ensuring responsible and impactful teaching and learning.

2. Background

Embodied cognition theory emphasizes the link between bodily experiences and cognitive processes, highlighting the role of physical actions such as gestures and movements in facilitating learning [13, 14, 15]. In mathematics education, embodied learning approaches help bridge abstract concepts with practical understanding, significantly improving comprehension and retention [16, 6, 17]. MSEs further enhance these benefits by integrating visual, auditory, and kinesthetic stimuli, providing dynamic and immersive ways for students to engage with educational content [18, 19, 20]. The integration of embodied learning strategies with GenAI technologies presents new opportunities for creating adaptive learning experiences. GenAI systems can process multimodal data to generate personalized feedback in real-time, tailoring responses to students' physical interactions and learning progress [21, 22]. This synergy enables educational environments where AI complements embodied learning by providing structured, task-specific guidance, enhancing engagement and cognitive understanding [23, 24].

Hybrid intelligence (HI) combines the strengths of AI and human instructors, fostering collaboration between the adaptive capabilities of AI and the contextual understanding of teachers [12, 25]. While GenAI excels at analyzing patterns and delivering consistent feedback, teachers promote critical reflection and metacognitive skills. Together, they create systems where AI handles routine feedback, allowing instructors to focus on fostering deeper learning [11]. Research into HI highlights its potential for advancing educational practices by aligning AI feedback with pedagogical goals and ensuring it complements human expertise [26]. Frameworks such as those proposed by Holstein et al. [11] outline how AI and human instructors can augment each other's strengths, emphasizing adaptability and scalability in hybrid models. However, challenges remain, including building trust in AI, addressing accessibility, and mitigating cognitive overload when integrating complex technologies.

This study builds on these foundations, exploring how GenAI, integrated with MSEs, can support embodied learning through hybrid intelligence, offering insights into the benefits and limitations of AI-human collaboration in education.

3. Methods

3.1. Design

This study utilized a portable platform called MOVES designed to enable MSEs, addressing limitations inherent in traditional learning environments. The educational framework involved students engaging with a body-scale NL to solve integer arithmetic problems. This NL served as an effective teaching tool, presenting integers in a spatially organized and ordinal layout, where negative integers mirrored their positive counterparts [3]. The SENSEi software [27] supported this system by coordinating dual projectors, one projecting on the wall and the other on the floor, to track and monitor students' movements, positions, and orientations during their interaction with the NL.

In Figure 1 an example of the NL activity. To solve the equation "-1 − 2," the student begins by standing on the -1 hash mark on the NL. As they step onto -1, the number beneath their feet turns blue, while the -1 displayed on the wall in front changes to green, accompanied by a sound. Next, the student turns to the left to face the subtraction direction. Upon turning, the subtraction sign on the wall also turns green, and another sound plays. Finally, the student takes two steps forward and raises their hands above their head to indicate that they have reached the solution. If correct, the entire problem, along with the solution on the wall, turns green, and a congratulatory sound plays. If incorrect, the solution remains unchanged, and no sound plays, as negative feedback is avoided to prevent demotivating or discouraging students [28].

This study integrated the GPT-4 language model into a Node.js web server to enhance the interactive

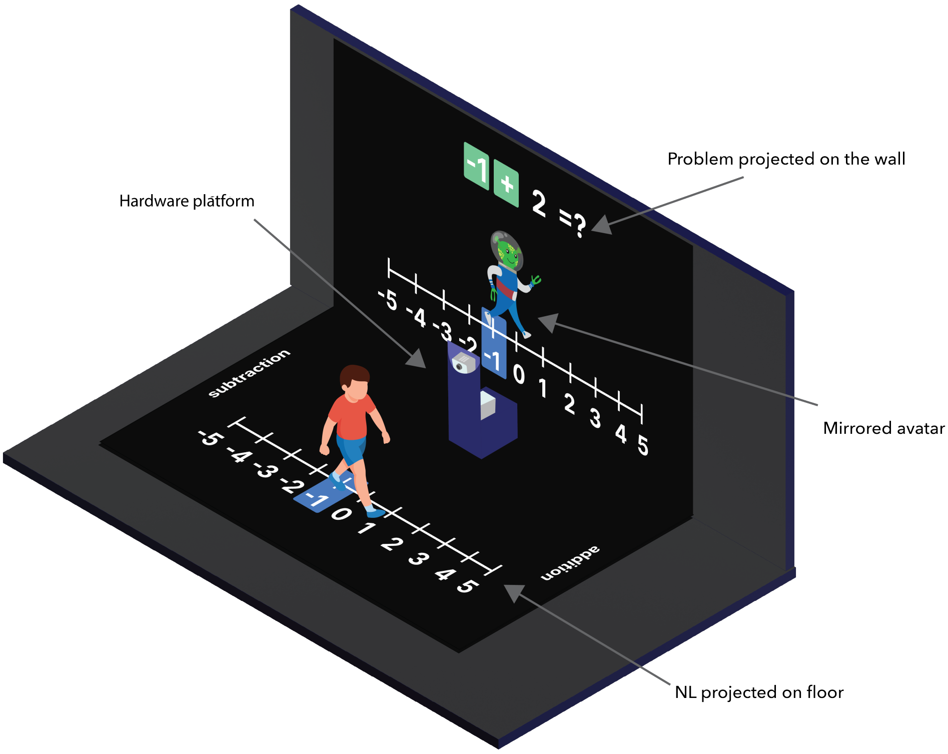

Figure 1: Students walk a floor-projected interactive number line, and a screen-based avatar mimics their whole-body movements

educational systems reliant on body movement and dynamic user input. The system uses real-time sensor data from students interacting with the number line, sending responses to the web server, which communicates with GPT-4 via OpenAI's API to generate personalized feedback. The Node.js server ensures efficient asynchronous communication, incorporates caching to reduce redundant API calls, and includes error-handling protocols for reliable operation. To ensure the effectiveness of GenAI feedback, the prompts were co-designed with the teacher, aligning AI responses with specific learning objectives and instructional strategies. This iterative process created a seamless blend of AI-driven feedback and human pedagogical guidance, delivering contextually relevant, constructive suggestions to support student engagement and learning. Below, we present an example of feedback generated by GPT-4 from system logs.

```
GenAI FEEDBACK EXAMPLES FROM LOGS:

EXAMPLE:
NLS:operationSelected|"1|+|-4"
NLS:LLM-UserMessage|"\nO1: +1\nOperator: + \nO2: -4 \nResult: -3\n
The student walked on the correct number: 1\nThe student correctly
rotated his body to the right\nThe student correctly walked backward\n
The student walked on the correct number: -3\nProvide feedback."
NLS:LLM-Response|"You correctly walked to 1, rotated right,
and walked back 4 steps to -3. Well done following the operation
rules with negative numbers on the number line!"
```

3.2. Context and procedures

In collaboration with a secondary school in Trondheim, the study lasted 2 weeks in May 2024. The teacher provided us with a list of participants for each day, and one of the researchers randomly called the participating students from the class list one by one; on average, each student's session lasted 40-50 minutes. The study aimed to assess the impact of adaptive summative feedback within the MOVES-NL educational tool. Participants were divided into two groups:

- Control Group (N=16): Received summative feedback from a teacher.
- Experimental Group (N=14): Received summative feedback through GenAI.

The study room was set up in a dedicated classroom inside the school to avoid external distractions. Each study session consisted of the phases:

1. Facilitator's introduction covering our identity, planned activities, and the data collection process (including camera recording and eye-tracking setup).
2. Introduction to the Walking NL: the facilitator showed students how to move to solve problems.
3. Walking NL: the student walks the NL in order to solve arithmetic problems while an avatar mirrors movements on the screen.
4. Summative feedback: the student gets feedback on their task based on the group setting.

During the sessions, children worked on various difficulty levels of math problems, with specific success criteria outlined for each level.

3.3. Participants

Our sample consisted of 34 students (18 females) between 11-13 years old, selected based on the curriculum timing when students begin learning about negative numbers and their applications in mathematical concepts. Prior to their participation, written informed consent was obtained from their legal guardians. All the ethical procedures were approved by the national human research ethics organization. Students' participation and data collection were conducted after approval from the Norwegian Agency for Shared Services in Education and Research (Sikt) and with Institutional Review Board approval (protocol ID 2022-10-15703), following all the regulations and recommendations for research with students.

3.4. Data Collection

We recorded students' interactions with MOVES-NL and employed sensing devices which allowed us to capture students' experience via multimodal data. The decision to use these data collection techniques was also influenced by the fact that they account for (to some extent) students' embodied learning and their importance in multisensory systems (e.g., students externalize their actions with the use of their body/skeleton). The sensing devices and their respective multimodal data allowed us to closely monitor and understand how students experienced the received support, leveraging the key affordances of multimodal data (e.g., temporality and direct access to indicators of students' cognitive and affective processes [29]). Students' activity sessions were recorded using two mobile cameras and one additional sensor device: gaze data from eye-tracking glasses. We collected children's gaze data using Tobii eye-tracking glasses at a 50 Hz sampling rate and one-point calibration. Due to errors in data collection, we had to discard four students.

3.5. Preliminary Results

The results reveal distinct patterns in how students interacted with feedback in the teacher and GenAI conditions, informed by both quantitative metrics (system logs, eye tracking, and motion sensors) and qualitative observations (video analysis). While task performance did not differ significantly between the two groups ($t(31.63) = 0.73, p > 0.05$), notable differences emerged in cognitive engagement and interaction styles. Students in the teacher feedback condition exhibited higher cognitive load ($t(26.12) = 2.89, p < 0.01$) and employed global processing strategies, as evidenced by their higher Information Processing Index (IPI) scores ($t(27.56) = 5.16, p < 0.001$). These students frequently re-enacted embodied solutions to problems, using gestures and full-body movements to justify their answers, reflecting the multimodal nature of human interaction.

In contrast, students in the GenAI feedback condition experienced lower cognitive load, focusing on task-specific verification strategies that prioritized correct responses rather than reasoning behind the incorrect ones. Their interactions were less physically expressive, with fewer gestures observed during feedback responses, as students concentrated on detecting whether the system's evaluation was accurate. Time spent on Areas of Interest (AOIs) further highlights these differences: students in the teacher condition focused on feedback and task text ($t(22.64) = 5.06, p < 0.001$), while those in the GenAI condition engaged more with the correct option and the number line ($t(25.03) = 2.68, p < 0.05$).

These findings suggest that the teacher feedback encouraged a more reflective, multimodal engagement with problems, while the GenAI feedback streamlined the learning process, fostering efficiency but with reduced physical expressiveness. Together, these insights underline the complementary strengths of human and AI feedback in educational contexts, offering opportunities to balance critical reflection with task-focused efficiency.

4. Discussion and Conclusion

The preliminary results highlight distinct benefits and limitations of teacher and GenAI feedback within an embodied learning context. Teacher feedback encouraged multimodal interaction and critical reflection, as evidenced by students' frequent use of gestures and body movements to justify their answers. This aligns with higher cognitive load and global processing strategies, suggesting that teacher feedback prompts deeper engagement and fosters metacognitive skills. However, this also comes at the cost of increased mental effort, which may impact efficiency during problem-solving tasks. Conversely, GenAI feedback reduced cognitive load and supported efficient, task-focused learning. Students in the GenAI condition concentrated on verifying the accuracy of feedback and interacting with the number line, indicating streamlined engagement with less emphasis on multimodal expression. While this efficiency can enhance learning for routine tasks, it may limit opportunities for critical reflection and the development of higher-order thinking skills.

These findings suggest that hybrid intelligence systems combining teacher and AI feedback could leverage the strengths of both approaches. Teachers can provide reflective, multimodal engagement to deepen understanding, while GenAI can offer consistent, task-specific feedback and reduce cognitive demands, and enhance efficiency. Such integration could create adaptive learning environments that cater to diverse student needs and learning objectives. Future research should explore ways to balance these complementary strengths, focusing on designing hybrid systems that integrate AI feedback seamlessly with human instruction. This includes addressing the observed limitations of GenAI, such as reduced physical expressiveness and student skepticism. Additionally, longitudinal studies could investigate the long-term effects of hybrid feedback on learning outcomes and student engagement.

In conclusion, the study underscores the potential of hybrid feedback systems in embodied learning environments. By combining the teacher's contextual understanding and multimodal engagement with the AI's efficiency and adaptability, hybrid approaches can enhance the educational experience, promoting both critical reflection and task-focused learning.

Acknowledgments

We would like to thank all the children and teachers who participated in this study. Without their enthusiasm and dedication, this work would not have been possible. Moreover this study was possible thanks to the Peder Sather Grant programme.

References

[1] L. Shapiro, Embodied cognition, in: The Oxford Handbook of Cognitive Science, Oxford University Press, 2019, pp. 118–136.

[2] M. Sheets-Johnstone, The primacy of movement, The primacy of movement (2011) 1–606.

[3] S. Varma, D. L. Schwartz, The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts, Cognition 121 (2011) 363–385. doi:10.1016/j.cognition.2011.08.005.

[4] L. Bofferding, A. J. Hoffman, Learning negative integer concepts: Benefits of playing linear board games., International Group for the Psychology of Mathematics Education (2014). https://api.semanticscholar.org/CorpusID:231261619.

[5] G. Cosentino, M. Gelsomini, K. Sharma, M. Giannakos, Interaction modalities and children's learning in multisensory environments: Challenges and trade-offs, in: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference, 2023, pp. 397–410.

[6] P. S. Moyer-PackenhaM, J. F. ShuMway, E. Bullock, S. I. Tucker, K. L. Anderson-Pence, A. Westenskow, J. Boyer-Thurgood, C. Maahs-Fladung, J. Symanzik, S. Mahamane, et al., Young children's learning performance and efficiency when using virtual manipulative mathematics ipad apps, Journal of Computers in Mathematics and Science Teaching 34 (2015) 41–69.

[7] D. Mutlu-Bayraktar, P. Ozel, F. Altindis, B. Yilmaz, Split-attention effects in multimedia learning environments: eye-tracking and eeg analysis, Multimedia Tools and Applications 81 (2022) 8259–8282.

[8] Z. Akata, D. Balliet, M. De Rijke, V. Dignum, F. Dignum, G. Eiben, et al., A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence, Computer 53 (2020) 18–28. doi:10.1109/MC.2020.2996587.

[9] F. Ke, R. Liu, Z. Sokolikj, I. Dahlstrom-Hakki, M. Israel, Using eye-tracking in education: review of empirical research and technology, Educational technology research and development (2024) 1–36.

[10] J. Roschelle, J. Lester, J. Fusco, Ai and the future of learning: Expert panel report., Digital Promise (2020).

[11] K. Holstein, V. Aleven, N. Rummel, A conceptual framework for human-ai hybrid adaptivity in education 21 (2020) 240–254. doi:10.1007/978-3-030-52237-7_20.

[12] J. Roschelle, J. Lester, J. Fusco, Ai and the future of learning: Expert panel report (2020). doi:10.51388/20.500.12265/106.

[13] P. Dourish, Where the action is, MIT press Cambridge, 2001.

[14] J. Piaget, Genetic epistemology, Columbia University Press, 1970.

[15] D. Abrahamson, A. Bakker, Making sense of movement in embodied design for mathematics learning, Cognitive research: principles and implications 1 (2016) 1–13.

[16] M. M. Riconscente, Results from a controlled study of the ipad fractions game motion math, Games and Culture 8 (2013) 186–214.

[17] R. Barrocas, J. Bahnmueller, S. Roesch, M. Lachmair, K. Moeller, Design and empirical evaluation of a multitouch interaction game-like app for fostering early embodied math learning, International Journal of Human-Computer Studies 175 (2023) 103030.

[18] G. Cosentino, M. Giannakos, Multisensory interaction and analytics to enhance smart learning environments: A systematic literature review, IEEE Transactions on Learning Technologies (2023).

[19] N. Sinclair, E. Heyd-Metzuyanim, Learning number with touchcounts: The role of emotions and the body in mathematical communication, Technology, Knowledge and Learning 19 (2014) 81–99.

[20] M. Chen, Small leans into big steps: A mixed-reality environment to support embodied, ensembled mathematics learning, in: Interaction Design and Children, 2022, pp. 564–568.

[21] W. Yang, X. Hu, I. H. Yeter, J. Su, Y. Yang, J. C.-K. Lee, Artificial intelligence education for young children: A case study of technology-enhanced embodied learning, Journal of Computer Assisted Learning 40 (2024) 465–477.

[22] S. Druga, S. T. Vu, E. Likhith, T. Qiu, Inclusive ai literacy for kids around the world, in: Proceedings

of FabLearn 2019, 2019, pp. 104–111.

[23] Y. Huang, Y. Cheng, L. Chen, J. Y. Hsu, Human-ai co-learning for data-driven AI, CoRR abs/1910.12544 (2019). http://arxiv.org/abs/1910.12544.

[24] S. Järvelä, J. Malmberg, H. Järvelä, Generation z and beyond: Co-evolution of human capabilities and intelligent technologies in the twenty-first century (2022) 1–13. doi:10.1007/978-981-19-6887-7_115.

[25] S. Järvelä, A. Nguyen, A. Hadwin, Human and artificial intelligence collaboration for socially shared regulation in learning, British Journal of Educational Technology (2023). doi:10.1111/bjet.13325.

[26] G. Cosentino, J. Anton, K. Sharma, M. Gelsomini, M. Giannakos, D. Abrahamson, Hybrid teaching intelligence: Lessons learned from an embodied mathematics learning experience, British Journal of Educational Technology (2024).

[27] M. Gelsomini, Sensei - experience in all senses. multisensory and multimodal research framework, 2023. https://sensei.space/.

[28] A. C. Van Duijvenvoorde, K. Zanolie, S. A. Rombouts, M. E. Raijmakers, E. A. Crone, Evaluating the negative or valuing the positive? neural mechanisms supporting feedback-based learning across development, Journal of Neuroscience 28 (2008) 9495–9503. doi:10.1523/JNEUROSCI.1485-08.2008.

[29] M. Cukurova, M. Giannakos, R. Martinez-Maldonado, The promise and challenges of multimodal learning analytics, Wiley Online Library (2020). doi:10.1111/bjet.13015.