

# AI-Enhanced Sociocultural, Multimodal Approaches for Equitable Collaborative Learning: Learning, Power, and Multimodality

Hakeoung Hannah Lee, University of Virginia, hannahlee@virginia.edu

**Abstract:** My research lies at the intersection of the learning sciences, learning analytics, and Science, Technology, Engineering and Mathematics education, and centers on making collaborative learning more equitable for K–12 learners, particularly multilingual, multicultural, and marginalized youth. Across three interconnected lines of inquiry, I position myself as a scholar who bridges the learning sciences and learning analytics, advancing theoretically grounded, situated, and equity-oriented approaches to illuminate the complexity of collaborative learning at multiple scales.

## Introduction

I am a learning scientist committed to advancing equitable, multimodal learning. This commitment has shaped my work across diverse contexts—from developing nationally awarded K–12 collaborative physical computing curricula, to teaching in a rural Ugandan village and developing sustainable curricula that integrate local folk tales, music, and art into lessons, to training as a student teacher in a culturally diverse Australian public school, to co-founding the nonprofit Society of Technology for Education and Learning Analytics, and to my current research in learning sciences and artificial intelligence (AI). After completing my Ph.D. in Science, Technology, Engineering and Mathematics (STEM) Education at the University of Texas at Austin, I joined the University of Virginia as a tenure-track Assistant Professor in the School of Education and Human Development.

My work focuses on developing and deploying AI-enhanced multimodal learning analytics approaches across diverse global contexts. Drawing on sociocultural perspectives of learning (Engeström, 1987; Lave & Wenger, 1991; Vygotsky, 1978; Wertsch, 1998) and translanguaging theory (García & Kleifgen, 2019; García & Wei, 2014), I examine multimodal participation dynamics—“the interplay of verbal, nonverbal, physical, and/or digital contributions that learners make as they engage with peers, (shared) tools, and objects” (Lee et al., 2026b, p. 3)—and seek to make these dynamics more visible in authentic, co-located collaborative STEM learning settings. Traditional educational frameworks have long privileged narrow, monolingual, and Western-centric forms of participation, such as verbal or written contributions, marginalizing learners whose linguistic and cultural practices differ from dominant norms or institutionally sanctioned expectations (Annamma et al., 2013; Flores & Rosa, 2015; McDermott & Hood, 1982). My research is thus driven by the imperative to broaden, through more inclusive perspectives, the boundaries of what is recognized and valued as meaningful participation. I anchor my work in the understanding that STEM classrooms should invite students to observe, question, reason, create, and collaborate (Cohen et al., 1999), while recognizing diverse ways of knowing and learning. Equally central is the conception of collaborative learning as a space where access to learning is shared, fostering collective authority and responsibility rather than reproducing teacher-centered authority within group settings (Barron, 2003; Cohen et al., 1999; Langer-Osuna et al., 2020; Lee et al., 2026b). These perspectives shape how I design research.

Across projects, I engage participatory and justice-oriented approaches, including research–practice partnerships and participatory design research (Bang & Vossoughi, 2016). These approaches foreground lived experiences, sustained relationships, and attention to power so that tools and studies emerge from the needs and aspirations of those most impacted. Methodologically, video recordings of authentic, co-located group work serve as my primary data source, as they capture the complex embodied, relational, and material dimensions of participation. I integrate qualitative, quantitative, and computational approaches in unique ways to study participation and collaboration at multiple scales using video data, as well as to examine the affordances and limitations of AI-supported approaches in analyzing these data.

## Research areas

*My first line of research* advances scalable, less intrusive, and ecologically valid multimodal learning analytics approaches for analyzing multimodal participation in co-located collaborative STEM learning environments using AI. Expanding my dissertation, I develop and deploy a theoretically grounded video analysis system that uses computer vision and neural networks to temporally track and visualize observable multimodal human activities (e.g., typing, talking, writing) to study collaborative group dynamics—an approach I refer to as *Activity Mapping*. This system and the associated visualization tool open new analytic possibilities by using AI to make nonverbal and tool-mediated interactions visible and to enrich subsequent quantitative and qualitative analysis. For example,

one of my recent studies (Lee et al., 2026b), which integrates the Activity Mapping approach with Epistemic Network Analysis (Shaffer, 2017) and qualitative video-based interaction analysis (Erickson, 2012; Jordan & Henderson, 1995), found that groups with dominant gender and language backgrounds exhibited distinct participation patterns, illuminating the varied ways in which roles, tasks, and ideas are negotiated and contributions are made. Importantly, my work is not oriented toward classifying, predicting, or labeling learners using AI. Rather, I design multimodal learning analytics approaches as augmentative tools that surface multimodal participation dynamics, enabling educators and researchers to notice the diverse, complex, and multimodal ways students participate. These approaches provide new analytical lenses that go beyond the limits of human observation or would otherwise require an impractical amount of time, supporting (not replacing) human critical interpretation. To learn more about this line of work, see Lee et al. (2026a, 2026b) and Lee (2025a). Ongoing work continues to expand the behavioral repertoire the system can detect, strengthen model performance, and prototype versions that support real-time teacher noticing in authentic classrooms (e.g., Lee et al., under review).

The *second area* examines how power and learning are co-constructed (Philip & Gupta, 2020) in the context of collaborative student groups, focusing on key mediational artifacts. I study power relations, including influence, access, and authority, such as who controls access, how behavior is shaped, and how these artifacts reflect and/or mediate broader societal structures and (in)equities. My recent research (Lee, 2025b, under review-b) investigates how access to a shared keyboard mediates power dynamics in a collaborative STEM learning group, focusing on how Latinx middle school students negotiate participation through social and intellectual authority (Langer-Osuna et al., 2020). Integrating the aforementioned AI-supported Activity Mapping approach with qualitative video-based interaction analysis (Erickson, 2012; Jordan & Henderson, 1995), quantitative analysis revealed shifting patterns of keyboard access, while qualitative findings exposed the nuanced ways these patterns reflected or obscured underlying authority dynamics (Lee, under review-b). Findings illustrate how the keyboard functioned not merely as a tool for input, but as a multifaceted mediational artifact serving various functions; how tool access does not inherently equate to authority; and how authority is not only asserted but must be socially ratified through peer uptake. These insights inform the design of more inclusive collaborative learning environments that honor sociocultural complexity. Having previously focused on Latinx multilingual learners, I am now expanding this work to more diverse multicultural and multilingual classrooms through new partnerships.

The *third area* centers on participatory design and sociocultural inquiry into learning analytics and AI-supported collaborative classrooms. Across partner hubs in Virginia, New Mexico, South Korea, and Brazil, educators and community leaders have voiced concerns that AI systems for education are often not designed with—or for—the communities most impacted. To challenge the Western, monolingual, and individualistic assumptions embedded in these systems, I plan to lead the Global Responsible AI in Education Collaborative (pending funding decision). This transnational, transdisciplinary initiative brings together researchers, practitioners, community partners, youth, AI developers, and policy experts from international hubs to co-develop globally grounded research agendas and responsible AI guidelines for teachers through cross-cultural collaboration, youth-led inquiry, and participatory design. The project ultimately seeks to reimagine what counts as knowledge, evidence, and participation in AI-supported education (Bang & Vossoughi, 2016), and to redefine “responsible AI” through equity-centered, multilingual, multimodal, and cross-cultural perspectives (Lee, under review-a). This work is currently underway as a pilot in one hub (funded by the University of Virginia IDEAs grant), where I lead a research–practice partnership with South Korean public schools to co-design learning analytics– and AI-supported classrooms that explore how such tools can be responsibly developed to advance equitable K–12 STEM learning across global contexts.

## Contributions to the learning sciences

Across these interconnected lines of work, I contribute to the field by (1) advancing theoretically grounded, AI-supported multimodal learning analytics approaches that honor sociocultural complexity; (2) deepening understanding of learner power dynamics, participation, and tool mediation in K–12 collaborative learning; and (3) shaping globally responsive, equity-centered frameworks for responsible AI in education. As a researcher trained in the learning sciences and committed to global, multilingual, multimodal, and community-centered forms of innovation, my work seeks to make methodological, theoretical, and practical contributions that expand how the field conceptualizes and studies participation and collaboration, while also providing a technical pathway for studying these dynamics at scale in real-world classrooms. Grounding my studies in converging insights from sociocultural theories of learning, critical and decolonial education research, translanguaging, multimodal learning analytics, and participatory design research (Bang & Vossoughi, 2016; Engeström, 1987; García & Kleifgen, 2019; García & Wei, 2014; Lave & Wenger, 1991; Vygotsky, 1978; Worsley et al., 2021), I aim to continue focusing on the critical yet underexplored intersection of AI, multilingualism, multimodality, and equity (Lee, 2025a, under review-a; Lee et al., 2026b).

## References

- Annamma, S. A., Boelé, A. L., Moore, B. A., & Klingner, J. (2013). Challenging the ideology of normal in schools. *International Journal of Inclusive Education*, 17(12), 1278–1294.
- Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and Instruction*, 34(3), 173–193.
- Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12(3), 307–359.
- Cohen, E. G., Lotan, R. A., Scarloss, B. A., & Arellano, A. R. (1999). Complex instruction: Equity in cooperative learning classrooms. *Theory into Practice*, 38(2), 80–86.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit.
- Erickson, F. (2012). Definition and analysis of data from videotape: Some research procedures and their rationales. In G. Camilli, J. Green, J. L. Green, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 177–191). Routledge.
- Flores, N., & Rosa, J. (2015). Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard educational review*, 85(2), 149–171.
- García, O., & Kleifgen, J. A. (2019). Translanguaging and literacies. *Reading Research Quarterly*, 55(4), 553–571.
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39–103.
- Langer-Osuna, J., Munson, J., Gargroetzi, E., Williams, I., & Chavez, R. (2020). “So what are we working on?”: How student authority relations shift during collaborative mathematics activity. *Educational Studies in Mathematics*, 104, 333–349.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Lee, H. H. (2025a). *AI-enhanced multimodal activity mapping: Understanding group dynamics through verbal, nonverbal, and object-based interactions in collaborative learning* (Doctoral dissertation). The University of Texas at Austin.
- Lee, H. H. (2025b). AI-enhanced sociocultural analysis of keyboard access in CSCL: Proxy, gatekeeper, or mediator of intellectual influence? In *General Proceedings of the ISLS Annual Meeting 2025*. International Society of the Learning Sciences.
- Lee, H. H. (under review-a). Responsible AI in education: Multilingual, multimodal, and relational futures. *Teachers College Record*.
- Lee, H. H. (under review-b). Table dynamics: AI-enhanced analysis of keyboard access and power dynamics in a collaborative STEM learning group. *Journal of Learning Sciences*.
- Lee, H. H., Lujan, M., Jatla, V., Egala, U., Pattichis, M. S., & Celedón-Pattichis, S. (under review). Activity Mapping: A sociocultural approach to developing an AI-based video analysis system for tracking and visualizing multimodal participation in collaborative learning. *Journal of Learning Analytics*.
- Lee, H. H., Lujan, M., Jatla, V., Egala, U., Pattichis, M., & Celedon-Pattichis, S. (2026a). Integrating AI and interaction analysis: A mixed-methods approach to tracing object-based participation in CSCL. In *Proceedings of the 19th International Conference on Computer-Supported Collaborative Learning (CSCL 2026)*. International Society of the Learning Sciences.
- Lee, H. H., Sung, H., Celedón-Pattichis, S., & Pattichis, M. S. (2026b). Toward an inclusive understanding of collaborative learning using MMLA: Exploring multimodal participation dynamics by gender and linguistic diversity. *Learning and Instruction*, 102, 102317.
- McDermott, R. P., & Hood, L. (1982). Institutional psychology and the ethnography of schooling. In P. Gilmore and A. Glatthorn (Eds.) *Children in and out of school: Ethnography and education*. Washington, DC: Center for Applied Linguistics, pp. 232–249.
- Philip, T. M., & Gupta, A. (2020). Emerging perspectives on the co-construction of power and learning in the learning sciences, mathematics education, and science education. *Review of Research in Education*, 44(1), 195–217.
- Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wertsch, J. V. (1998). *Mind as action*. Oxford University Press.
- Worsley, M., Martinez-Maldonado, R., & D'Angelo, C. (2021). A New Era in Multimodal Learning Analytics: Twelve Core Commitments to Ground and Grow MMLA. *Journal of Learning Analytics*, 8(3), 10–27.