International Society of the Learning Sciences

Embodied Geometry Fingerprint: An Example of How to Make Implicit Explicit in Multimodality Research on Embodied Learning

Alex Luu, University of Georgia, alexluu@uga.edu
Theodore J. Kopcha, University of Georgia, tjkopcha@uga.edu
Ceren Ocak, Georgia Southern University, cocak@georgiasouthern.edu

Abstract: Multimodal analysis of embodied learning through video data can be overwhelming for both researchers and practitioners. It requires attending to the spatial and temporal components of verbal and non-verbal interactions simultaneously, while also identifying moments that reveal learning as an embodied phenomenon. For practitioners, it is often unclear how the researchers arrive at these conclusions and whether those conclusions are optimal for the practitioners' context. This paper presents the development of an embodied geometry fingerprint (EGF), which is an interactive visualization of embodied learning interactions. The EGF enhances the exploration of embodied learning, making the process of identifying salient moments of interaction more explicit. This has the potential to broaden research-practice partnerships, offering educators and other stakeholders a tool to more effectively express and brainstorm ideas, refine and expand data exploration trajectories, as well as foster clarity and transparency of communication and decision-making across the partnership team.

Introduction

Embodied cognition postulates that human cognition is "deeply rooted in the body's interactions with the world" (Wilson, 2002, p. 625). Subsequently, it is commonly accepted that evidence of one's embodied understanding of subject matter unfolds in both their verbal (spoken and written discourses) and non-verbal (e.g. gestures, body movements) communication (Nathan, 2022). The manner in which these modes of interaction come together reflect how a learner is creating meaning at a given point in time (Jewitt, 2013; Bezemer, 2014). Being able to capture the intricate connections between a learner's verbal and non-verbal communication, video-based multimodal analysis has become a common choice for studying embodied learning. Specifically, the research in this area usually analyzes video recordings of naturally occurring embodied learning interactions using Interaction Analysis (Jordan & Henderson, 1995; Hall & Stevens, 2015) and Multimodal Transcription (Bezemer, 2014). While the analyses based on these frameworks are often reported in a systematic sequence of linear steps, there is a good part of the actual work of finding and making sense of empirical evidence that is "much more complex and messy than illustrated" (Saldaña, 2013, p. 12), with much of the process remaining implicit, not only between the research team and their audience but also among the research team members. This is particularly true early in the process, where the researcher must identify salient and important verbal and non-verbal interactions that reveal more about a phenomenon of interest (Jewitt, 2013). The act of identifying those interactions can be 'messy' in that it requires the researcher to first notice patterns of activity that may be of interest and then interpret whether those patterns are relevant to the phenomenon of interest as they emerge over time (Bezemer, 2014).

In this paper, we present the development of an *embodied geometry fingerprint* (EGF) as a real-life example of how to make the exploration nature of multimodality research on embodied learning more explicit. Like a human fingerprint, the EGF represents the unique and complex interactions between a learner's spoken words, their bodies, and the environment as they explain their mathematical reasoning. In embodied learning, that fingerprint is both a temporal and spatial phenomenon (Kopcha & Ocak, 2025; see also Bezemer, 2014). Some aspects of reasoning that can only be understood through the learner's interaction with the environment in a given moment while others must also be understood in the way that they emerge *over time.* Thus, a central practice surrounding the EGF is the creation of a comprehensive data visualization that brings the spatial and temporal components of embodied learning together in a simultaneous manner. At the exploratory stage of the research process, the EGF allows the salient features of embodied learning – both verbal and non-verbal – to emerge from the data to determine how the learner has ascribed meaning to those features.

For researchers, then, the EGF offers a powerful tool that can help identify patterns of interaction between the learner and the environment as they emerge over time. This creates a tangible common ground for explicitly brainstorming, formulating and discussing observable aspects of embodied learning rather than ending up in a *garden of forking paths* in which research decisions are implicitly made based on intuition and a narrow focus on select aspects of the data (Gelman & Loken, 2013). Regarding practitioners, having access to a visual overview of embodied learning offers more transparent and contextualized perspectives on how to make sense of and apply research findings. In a way, the EGF offers a tool that complements the traditional definition of research-practice

 © ISLS

partnership – "[l]ong-term, mutualistic collaborations between practitioners and researchers that are intentionally organized to investigate problems of practice and solutions for improving district outcomes" (Coburn et al., 2013, p. 2) – by helping practitioners gain better insights into research work and its implications, which can then encourage more actionable collaborations between practitioners and researchers.

Below, we describe the process of generating an EGF, providing specific decision-making details that help make the implicit aspects of analyzing multimodal video data more explicit. We conclude with multifaceted empirical evidence of the EGF's methodological contributions.

Method: The development of an embodied geometry fingerprint (EGF)

The EGF emerged as part of analyzing video data in which several 5th-grade children explained their reasoning about the symmetry of two 3-D shapes, a square-based pyramid and an equilateral triangular prism (Kopcha et al., 2025). As part of our analysis, we struggled to identify key moments of embodied reasoning while also understanding how those moments emerged over time. This led us to explore new ways to visualize the interactions that we were observing on the video. This paper presents our process of developing our first EGF, which was associated with one participant. We selected this participant because her communication was expressive and her non-verbal interactions clearly connected with her verbal communication.

To explicitly capture the interactions within this participant's 11-minute video session, we decided to systematically document our exploratory annotations in a machine-readable format so that we can use it to create a comprehensive exploratory visualization (i.e. the EGF) in an automatic manner. Reasoning about the nature of the visualization, we wanted it to be able to show the time-aligned member relevance (i.e. the aspects of participants which are relevant to a particular interactional event) and consequentiality (i.e. the interrelationship between interactional events) over the course of the interview for sequential analysis (Hall & Stevens, 2015). These intentions informed us to design our exploratory study as a sequence of tasks:

(Multimodal annotation)

- select the annotation tool that is optimal for our data and study purposes
- review the available annotation schemes that are relevant to the phenomena under study
- create annotations (including components of our annotation scheme) and optimize them on the fly

(Multimodal visualization)

- visualize time-aligned embodied interactions

Multimodal annotation

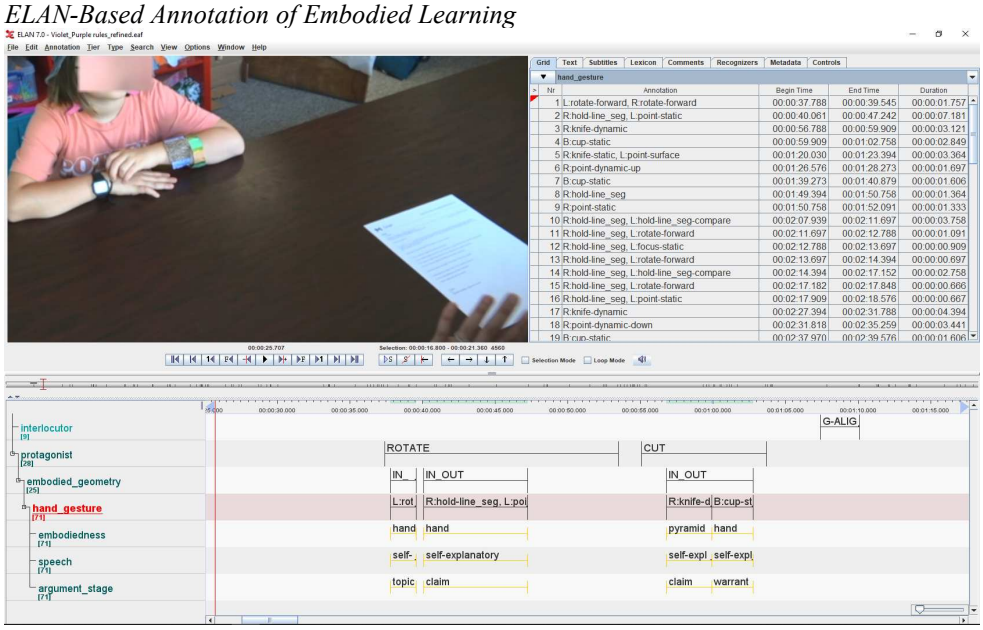
We selected an annotation tool such that our process could be easily reproduced and customized by other people while also allowing our annotation results to be freely and conveniently used as input for further data analyses or annotation tasks. This resulted in our selection of ELAN (ELAN (Version 7.0), 2025), a free, actively maintained, and cross-platform open-source annotation tool for audio and video recordings (Glüer, 2018), also recommended by Bezemer (2014). Data-wise, the selected tool allowed for multi-layered time-aligned annotations of video data as we want to locate and characterize multimodal aspects of embodied reasoning instances emerging during the interview session. Figure 1 shows the ELAN interface of our annotation work.

To be more informed when creating our own annotations, we studied the Multimodal Multidimensional (M3D) gesture labeling system (Rohrer et al., 2025), which focuses on both the *form,* i.e. the physical components of the interactions, as well as the *meaning,* which identifies the semantic and pragmatic functions of a gesture or interaction, and decided to use the system's handshape labels as much as possible.

As our team was interested in speech gesture co-dependency in mathematical reasoning, we focused on the video fragments that demonstrated how the learner expressed her geometrical reasoning via both speech and hand gestures. We first created an ELAN annotation tier called *protagonist* to identify those fragments. As we were also interested in the interactional aspects of the interview, we added the video fragments expressing various types of salience in social conversation (Luru, 2022) such as laughter or confusion to the tier *protagonist*, and created another annotation tier called *interlocutor* to capture the video fragments during which the interviewer showed her gesture alignment with the learners. To capture the interaction between speech and gestures in the learner's embodied reasoning, we created two annotation tiers: *embodied_geometry*, which answers the question whether gestures are the consequence of the learner's thought expressed in her speech or vice versa, and *speech*, which answers the question if the speech content is self-explanatory or dependent on the corresponding gestures. Annotating the gesture forms (the tier *hand_gesture*), we distinguished the forms of the right hand and the left hand. Each form was described by an action name (e.g. *cup* or *knife*), possibly followed by its movement property (i.e. *static* or *dynamic*). To elaborate the embodied aspect of the gesture, we added the annotation tier *embodiedness* to note if the hand interacts with physical 3D shape within the considered gesture. Finally, we

created the annotation tier *argument_stage* to label the nature of mathematical reasoning demonstrated during the embodied fragments. The stages were based on Toulmin's stages (see Pedemont & Balacheff, 2016) and included an initial stage of identifying information, making claims or arguments based on that information, and backing up those claims with additional arguments or information.

Figure 1
ELAN-Based Annotation of Embodied Learning



Multimodal visualization

After finishing our exploratory annotation, we exported our annotated labels as well as their corresponding tiers and timestamps into a *csv* file. This file can then be transformed using spreadsheet software (e.g. Microsoft Excel or Google Sheets) and/or scripting/programming languages (e.g. Python or R) to create the data that are suitable for our intended visualization. We particularly used both Google Sheets and Python scripts to prepare our multimodal annotation data for visualization. We visualized our data using the Plotly Open Source Graphing Library for Python (https://github.com/plotly/plotly.py). The charts generated by Plotly are interactive and, therefore, very convenient for inspection. We selected Paul Tol's discrete rainbow color scheme (consisting of 23 colors) (*Paul Tol's Notes*, n.d.), which was appropriate for accommodating the biggest set of different color-coded categories in our study (18 categories) while also being accessible to people who are colorblind (Nichols, n.d.).

Figure 2
Plotly-Based EGF



To create an EGF that captured the member relevance and consequentiality over the course of the interview as mentioned above, we visualized each time-aligned annotation unit as a horizontal bar in a Gantt chart (see Figure 2), whose horizontal axis denoted a timeline (in minutes and seconds). All annotation units belonging to one annotation tier were put at the same height in the chart, signified by the tier's label. Different types of annotation units within a tier were color-coded differently, as listed in the chart's legend. The chart is interactive,

 © ISLS

supporting active exploration of the data. For example, when hovering over a horizontal bar, users can access detailed information including the annotation name and value, as well as its start and end times. In addition, the range slider beneath the Gantt chart allows users to conveniently examine the annotated data by panning and zooming in on the X-axis while maintaining an overview of the chart. Finally, the configuration options at the top-right corner of the figure provide additional interactive affordances such as *Box Select*, *Draw line*, *Download plot as a PNG*, etc.

Findings and future work: Affordances emerging from interacting with the EGF

Versatility

Our team has used the EGF as a grounding object for expressing and brainstorming ideas, and for refining and expanding data exploration trajectories (e.g. by adding new elements to the visualization). Being able to collectively refer to tangible aspects of the EGF helped us add clarity and transparency to our communication and decision-making, and devise concise and intuitive ways to present our discussion ideas for future reference.

Informativeness

The EGF allows us to identify emergent patterns of interactions at both micro- and macro-levels, displaying a detailed view of the components within those interactions and the relationships between those components. For example, at a macro-level (see Figure 2), there is clear variability in the learner's use of the left-hand. Figure 3 displays a micro-analysis of this pattern in which the left hand is consistently in shape (i.e. cup) and aligned with the portion of mathematical reasoning in which a conclusion is drawn based on evidence (i.e. warrant). Other patterns were noticeable at the macro-level, e.g. a decrease in the use of left-hand gestures over time or a rise in comparing one object to another toward the end of the session. We note these patterns not to continue presenting the results of our multimodal analysis but rather to illustrate the powerful role that an EGF can play in that analysis. The EGF allowed us to characterize the dynamic interactions that were associated with an embodied event, both over time and within a specific moment, helping us make sense of the embodied aspects of the learning experience while also identifying patterns relevant to our research questions.

Figure 3

Micro-Analysis of Alignment Between Left-Hand Gestures and Argument



Enrichability

The EGF features a great level of enrichability to the analysis process. Along the Y-axis, we can add new tiers to capture more fine-grained aspects of multimodal interaction. For example, after our initial exploration of the EGF, we realized we need to take into account the social dynamic between learner and interviewer to better explain speech-gesture co-dependency and the relationship between teacher questions and learner responses. We can easily return to our annotation tool, ELAN, and add new aspects of multimodal interaction to our data. For example, we could annotate the type of question the teacher asks (e.g., probing, explaining) or even the gaze patterns (e.g., person-to-person, person-to-hand) to better capture the synergy of mathematical problem solving and social scaffolding that occurred during the recorded interview. This new data can be integrated into our data visualization pipeline to enrich the EGF, which will then enhance our understanding of embodied learning.

Future work

First, we will leverage the EGF and its derivatives to perform multimodal analysis on how embodied learning unfolds through the recorded interaction. Second, we will develop an open-source web application that allows users to create their own embodied fingerprints from *csv* files of their multimodal annotations. Finally, we will study to what extent the EGF can help practitioners make sense of and apply research findings.

 © ISLS

References

Bezemer, J. (2014). Multimodal Transcription: A Case Study. In S. Norris & C. D. Maier (Eds.), *Interactions, Images, and Texts: A Reader in Multimodality* (pp. 155–170). De Gruyter Mouton.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts [White Paper]. William T. Grant Foundation. https://eric.ed.gov/?id=ED568396

ELAN (Version 7.0). (2025). [Computer software]. Max Planck Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University. https://sites.stat.columbia.edu/gelman/research/unpublished/forking.pdf

Glüer, M. (2018). Software for Coding and Analyzing Interaction Processes. In E. Brauner, M. Boos, & M. Kolbe (Eds.), *The Cambridge Handbook of Group Interaction Analysis (*pp. 245–274). Cambridge University Press. https://doi.org/10.1017/9781316286302.014

Hall, R., & Stevens, R. (2015). Interaction Analysis Approaches to Knowledge in Use. In *Knowledge and Interaction.* Routledge.

Jewitt, C. (2013). Multimodal methods for researching digital technologies. *The SAGE handbook of digital technology research,* 250-265.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *Journal of the Learning Sciences, 4*(1), 39–103. https://doi.org/10.1207/s15327809jls0401_2

Kopcha, T., Diamond, J., & Kulp, K. (2025). Children's embodied reasoning about rotational symmetry: A multiple case study. *Mathematical Thinking and Learning, 0*(0), 1–26. https://doi.org/10.1080/10986065.2025.2579007

Kopcha, T., & Ocak, C. (2025). *Embodied Learning in the Schools: Designs That Link Brain, Body, and Technology in K-12 Classrooms*. Routledge.

Luu, A. (2022, August). *Edge Cases of Discourse Salience in American English Casual Dialogs: A New Window into the Co-Constructed Nature of Social Conversation*. Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue - Poster Abstracts. https://www.semdial.org/anthology/papers/Z/Z22/Z22-4031/

Nathan, M. J. (2021). *Foundations of Embodied Learning: A Paradigm for Education*. Routledge. https://doi.org/10.4324/9780429329098

Nichols, D. (n.d.). Coloring for Colorblindness. Retrieved September 28, 2025, from http://www.davidmathlogic.com/colorblind/

Paul Tol's Notes. (n.d.). Retrieved September 27, 2025, from https://sronpersonalpages.nl/~pault/

Pedemonte, B., & Balacheff, N. (2016). Establishing links between conceptions, argumentation and proof through the ck¢-enriched Toulmin model. *The Journal of Mathematical Behavior, 41*, 104–122. https://doi.org/10.1016/j.jmathb.2015.10.008

Rohrer, P. L., Tütüncübasi, U., Florit-Pons, J., Vilà-Giménez, I., Esteve-Gibert, N., Ren-Mitchell, A., Shattuck-Hufnagel, S., & Prieto, P. (2025). Multidimensional Labeling of Gesture in Communication: The M3D Proposal. *Corpus Pragmatics*. https://doi.org/10.1007/s41701-025-00197-2

Saldaña, J. (2013). The Coding Manual for Qualitative Researchers (Second edition). SAGE Publications.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. https://doi.org/10.3758/BF03196322