frontiers in Language Sciences

TYPE Perspective
PUBLISHED 11 March 2026
DOI 10.3389/flang.2026.1753250

Check for updates

OPEN ACCESS

EDITED BY
Omid Khatin-Zadeh,
University of Electronic Science and Technology of China, China

REVIEWED BY
Javier Diez-Palomar,
University of Barcelona, Spain
Roger De Abreu Silva,
Planalto Catarinense University, Brazil

*CORRESPONDENCE
Danyal Farsani
✉ Danyal.farsani@ntnu.no

RECEIVED 24 November 2025
ACCEPTED 14 February 2026
PUBLISHED 11 March 2026

CITATION
Autenrieth D, Autenrieth N and Farsani D (2026) From metrics to meaning: large language models and the computational turn in embodied educational research. *Front. Lang. Sci.* 5:1753250. doi: 10.3389/flang.2026.1753250

From metrics to meaning: large language models and the computational turn in embodied educational research

Daniel Autenrieth[1], Nina Autenrieth[2] and Danyal Farsani[3]*

[1]Independent Researcher, RWTH Aachen University, Aachen, Germany, [2]Center for Media Education, Institute of Educational Research, University of Education Schwäbisch Gmünd, Schwäbisch Gmünd, Germany, [3]Department of Teacher Education, Norwegian University of Science and Technology, Trondheim, Norway

Educational science has long grappled with a methodological tension: quantitative metrics offer scale but often lack depth, while qualitative inquiry offers depth but is difficult to scale. Mixed-methods approaches have sought to address this tension by combining both paradigms, yet practical constraints of time, labor, and analytical capacity have typically limited how fully researchers can integrate interpretive depth with large-scale analysis. At the same time, theories of embodied cognition emphasize that learning, language, and communication are grounded in sensorimotor, affective, and socio-cultural experience rather than in abstract, amodal symbols. The emergence of Large Language Models (LLMs) and multimodal Artificial Intelligence (AI) provides new opportunities to bridge these methodological and theoretical developments. In this article, we conceptualize "Computational Hermeneutics" as the interpretation of meaning at scale, informed by embodied perspectives on cognition, language, and communication. We outline three computational methodologies applicable to educational research: Semantic Similarity Rating (SSR), AI-based Qualitative Content Analysis (AI-QCA), and Computational Ethnography via multimodal video analysis. We show how these approaches can operationalize embodied, hermeneutic processes, such as interpreting student reflections, tracing metaphorical embodiment, or analyzing classroom habitus, at a scale previously reserved for standardized testing. By detailing the theoretical basis, opportunities, and limitations of these methods, we argue for a "computational turn" in educational science and research that remains faithful to embodied, multimodal meaning-making while overcoming traditional scalability constraints.

KEYWORDS
large language models, educational research, computational hermeneutics, semantic similarity rating, AI-based qualitative content analysis, computational ethnography

1 Introduction

A central aim of educational science is to understand the transformative constitution of the self in relation to the world (Laros et al., 2017). This goal aligns with global frameworks emphasizing transformative learning, identity formation, and the cultivation of 21st-century citizenship (Fullan and Scott, 2014). The objective is not merely the acquisition of static, decontextualized knowledge, but the development of reflective subjects capable of navigating complex social, material, and technological environments (Toffler, 2022; Turner, 2001). Measuring such transformative, situated processes presents a fundamental challenge. Traditional computational methods in education such as learning analytics tend to reduce complex trajectories of development to static performance metrics. By prioritizing efficiency and standardized outcomes, these approaches frequently fail to map the lived reality of educational experience, flattening the multi-layered complexity of human development into linear data points. In contrast, qualitative methods such as ethnography, videography, and content analysis can capture the nuance of embodied and situated transformation but are labor-intensive and difficult to scale (e.g., Junker et al., 2024; Rosa et al., 2020).

In parallel, theorists in cognitive science and linguistics have argued that cognition and communication are fundamentally embodied, situated, and action-oriented (e.g., Banaruee and Khatin-Zadeh, 2026; Barsalou, 2008; Gibbs, 2006a; Gibson, 1977, 1986; Lakoff and Johnson, 1980, 1999; Wilson, 2002; Ziemke, 2003; Zlatev, 2009). Meaning is not merely encoded in abstract symbols but is grounded in bodily interaction, affect, gesture, and socio-material practices (Banaruee and Khatin-Zadeh, 2026). Educational processes; teaching, learning, assessment, are likewise enacted through bodies in space, through multimodal communication, and through the affordances of artifacts and environments (Borghi and Binkofski, 2014; Tomasello, 2003). Recent advances in Natural Language Processing (NLP; Bittermann and Fischer, 2024) and Generative AI, including multimodal models that process text, image, audio, and video, enable a new synthesis of these strands. We propose that educational researchers can now employ computational tools to perform "Computational Hermeneutics": the interpretation of embodied, situated meaning at scale. Rather than merely optimizing existing metrics, this approach aims to mathematically operationalize hermeneutic processes, such as understanding students' metaphorical expressions of emotion, or tracing gesture-speech coordination in the classroom, within large datasets.

We exemplify this perspective through three methodological pillars:

1. Semantic Similarity Rating (SSR), which moves beyond discrete Likert scales by assessing the semantic and often metaphorically grounded content of texts relative to theoretically defined anchors.

2. AI-based Qualitative Content Analysis (AI-QCA), which integrates LLMs into established interpretive workflows to structure, categorize, and explicate large textual corpora.

3. Computational Ethnography via multimodal video analysis, which operationalizes embodied interaction by analyzing gesture, gaze, posture, and material engagement at scale.

Before detailing these methodologies, we first position them within the broader landscape of embodied cognition, language, and communication.

2 Semantic Similarity Rating (SSR) in educational contexts

2.1 Theoretical basis

Semantic Similarity Rating (SSR) represents a shift from discrete Likert scales (Likert, 1932) to high-dimensional vector mapping. Traditionally, measuring an attitude, such as a student's Reflective Stance or Motivation, required respondents to select a number (1–5). This approach often fails to capture the reasoning behind the choice (Maier et al., 2025). SSR, however, elicits free-text responses and projects them onto a scale by computing the cosine similarity of their embeddings against predefined reference statements.

Technically, embeddings are mathematical representations where text is converted into a vector of numbers representing its semantic meaning in a multi-dimensional space. Cosine similarity measures the angle between two vectors; if two texts share a similar meaning, their vectors point in the same direction, yielding a score close to 1. By defining anchors (reference texts representing the extremes of a scale, e.g., Anchor 1: "Surface Recall" vs. Anchor 5: "Deep Transformation"), SSR allows researchers to measure the semantic position of any text with mathematical precision. This methodology can potentially support two distinct applications in educational science: Simulation (e.g., stress-testing curriculum via synthetic personas) and Measurement (e.g., analyzing real student artifacts).

2.2 Simulation and curriculum stress-testing (Persona-based)

Before a curriculum is deployed, researchers often lack data on how diverse student groups will react. SSR allows for "Synthetic Field Testing" by creating AI-simulated student personas to evaluate educational materials. The process could be applied like this:

- Define Personas: Instead of generic prompts (e.g., "You are a student"), researchers define high-fidelity personas based on sociological frameworks. For international applicability, models like the Sinus-Milieus®, which segment populations based on social values and lifestyle (e.g., "Traditionalists" vs. "Expeditive/Modern"), provide a robust foundation.
- Stimulus Presentation: An LLM primed with persona-data is presented with a curriculum artifact (e.g., a history text or a controversial ethical case study).
- Response Generation: The personas generate synthetic journal entries or reactions to the material.
- SSR Analysis: These synthetic texts are measured against educational goals (Anchors) to predict engagement or alienation.

The use of synthetic personas is grounded in recent empirical validation. Maier et al. (2025) demonstrated that when LLMs are conditioned with demographic profiles, they reproduce human response patterns with approximately 90% correlation attainment relative to human test-retest reliability. This suggests that sophisticated personas can serve as valid proxies for predicting group-level reactions to educational content. An example (see Figure 1) could be:

A university develops a new module on "Global Ethics." To ensure inclusivity, the module is "taught" to synthetic student personas representing the *Sinus-Meta-Milieus* (e.g., a "Cosmopolitan Avant-Garde" student vs. a "Traditional Working Class" student). SSR analysis reveals that the "Traditional" persona consistently generates text semantically close to "Alienation," signaling a need to revise the curriculum before real students enter the classroom.

2.3 Longitudinal trajectory mapping (Artifact-based)

The second application applies SSR to existing artifacts. Here, no personas are used; AI acts strictly as a "Semantic Ruler" to measure the development of actual students over time. The process could potentially be applied as follows:

- Data Collection: Researchers collect a corpus of student writing (e.g., weekly reflective journals) over an academic year.
- Anchor Construction: The researcher defines the "Formative Scale" by writing Gold Standard reference statements. Anchor A might represent "Instrumental Knowledge Acquisition" (learning for the test), while Anchor B represents "Transformative Self-Relation" (learning that changes one's worldview).
- Trajectory Mapping: The system calculates the semantic distance of every journal entry to these anchors.
- Visualization: This yields a longitudinal curve for every student, visualizing their movement through "meaning space" from instrumentalism toward transformation.

This method solves the problem of "reductionism" in learning analytics. Instead of reducing a student's growth to a grade (Efficiency), it maps their changing relationship to the world (Depth). It detects non-linear patterns, such as a "crisis of meaning" (a drop in certainty), that traditional metrics might misinterpret as failure, but which e.g., transformative educational theory (Koller, 2017) recognizes as a necessary step in formation.

While powerful, both applications face constraints. The Simulation approach is limited by the training data of the LLM; if the model has not seen sufficient text from a specific socio-economic milieu, the persona may default to stereotypes (Maier et al., 2025). The Measurement approach is entirely dependent on the validity of the "Anchors"; if the reference statements for "Transformation" are poorly defined, the resulting metric will be meaningless. Thus, the role of the educational researcher shifts from "grader" to "architect of meaning," responsible for designing rigorous semantic frameworks.

3 AI-based qualitative content analysis (AI-QCA)

3.1 Theoretical basis

Qualitative Content Analysis (QCA) is a systematic method for interpreting text through categorization, traditionally relying on human expertise to ensure validity (Mayring, 2015; Kuckartz and Rädiker, 2023). AI-based QCA (Fischer, 2025) extends this tradition by integrating LLMs into the analytical workflow. This integration does not simply replace the researcher but changes their role, operating in what Dell'Acqua et al. (2025) describe as a "Centaur" mode (where human and AI split tasks) or a "Cyborg" mode (where they continually interact).

To ensure scientific rigor, AI-QCA must follow a strict "Process Model" rather than unstructured prompting. Fischer (2025) delineates a six-step framework for this methodology: Context (defining the corpus and research question), Technique (selecting specific analytical modes like summarization or explication), Units (defining coding units vs. context units), Process (designing the prompt workflow), Implementation (executing and monitoring), and Evaluation (validating results). This structure prevents the analysis from becoming a "black box" and ensures that AI's output, whether it is summarizing vast corpora or deducing categories, remains methodologically controlled and replicable.

3.2 Application scenario: iterative analysis of student reflections

To visualize how this methodology functions in practice, consider a study analyzing 500 student journals for evidence of "Critical Reflection." In a traditional manual approach, a researcher might code 50 journals and generalize the findings. Using AI-QCA, the researcher employs a multi-stage, iterative workflow to analyze the entire corpus.

The process begins with AI-based Summarization to condense the data. The researcher defines the individual journal entry as the processing unit and designs a specific prompt template, such as: "Given the following journal entry {Text}, paraphrase all statements related to the student's changing self-view." This transforms 500 pages of raw text into a structured set of core thematic statements, effectively creating a manageable "meta-corpus" for further analysis. Next, the researcher employs AI-based Induction to generate the category system. Instead of imposing a pre-existing theory, AI is provided with the summarized statements and prompted to identify clusters of recurring themes, grouping them into a coherent category system (e.g., "Questioning Authority," "Personal Crisis," "New Perspective"). This mimics the "Open Coding" phase of Grounded Theory (Glaser and Strauss, 2008) but operates at a scale and speed unattainable by human coding alone. Finally, the researcher moves to AI-based Deduction and Explication to apply these categories back to the raw data. AI is instructed to classify each original journal entry according to the newly induced system and to explicate its decision. A prompt might read: "Assign this text to one of the following categories {Category List}. Provide a one-sentence justification citing specific

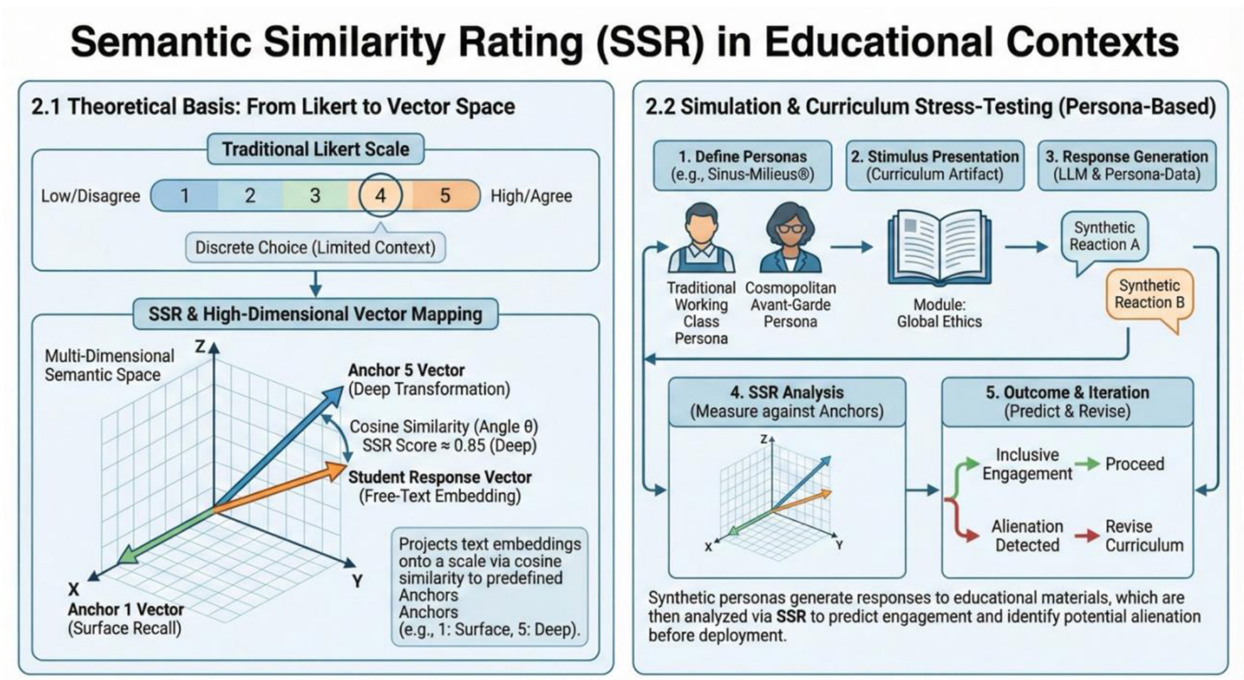

FIGURE 1
Semantic similarity rating (SSR) in educational contexts.

text evidence." This step converts AI from a simple classifier into a hermeneutic partner, providing a "thick description" of why a specific student reflection constitutes a moment of transformative learning. The researcher then reviews a random sample of these explications to validate AI's interpretive logic, ensuring the analysis remains grounded in human pedagogical understanding.

3.3 Critical reflection on opportunities and challenges

This methodological shift offers profound opportunities for educational science and educational research. By automating the mechanical aspects of coding, AI-QCA allows researchers to move from analyzing small, non-representative samples to conducting "Mass Processing" of entire educational datasets, revealing patterns that would otherwise remain invisible. Furthermore, the capability of AI-based explication adds a layer of interpretive depth; an LLM can be prompted to contextualize specific segments, e.g., explaining the socio-economic implications of a student's statement, thereby enriching the analysis with a broader context that a tired human coder might overlook.

However, the reliance on probabilistic models currently introduces distinct validity challenges. LLMs are prone to "hallucinations," generating plausible but factually incorrect summaries. There is also the risk of reproducing biases present in the training data; models may subconsciously associate certain dialects with lower academic competence (Hofmann et al., 2024; Bui et al., 2025). Moreover, the stochastic nature of these models means that re-running the same prompt may yield different categorizations, challenging the traditional scientific criterion of replicability. Despite these current limitations, emerging computer science research suggests the "black box" may soon become far more transparent. Nikolaou et al. (2025) have recently proven that Transformer language models are mathematically injective, meaning they map every distinct input text to a unique internal representation without losing information. This challenges the assumption that neural networks inherently "blur" or "compress" meaning in unrecoverable ways. Their work demonstrates that it is possible to algorithmically invert these models, reconstructing the exact input text solely from the model's hidden activations. For educational researchers, this implies that the "hermeneutic path" of AI is not structurally unknowable. Future tools built on this property could allow researchers to audit exactly how a student's text is represented within the model, potentially resolving current reproducibility issues by providing a mathematical guarantee that the analysis is based on a precise, lossless reading of the qualitative data.

4 Computational ethnography and multimodal video analysis

4.1 Theoretical basis

Education and learning often occurs in the "conjunctive" realm of implicit knowledge and research frequently needs to consider how people interact (using gestures, proxemics and non-verbal language), rather than just what they say (Khatin-Zadeh et al., 2022b, 2023c; Rosa and Farsani, 2021). Videography, as

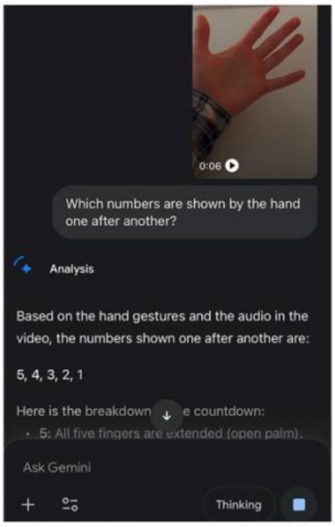

FIGURE 2
Multimodal analysis of temporal gestures. The screenshot demonstrates the model's ability to identify, sequence, and explicate a physical countdown (5–1) from a video stream. The QR code provides access to the original video data showing the hand and LLM response. Link to video: https://t1p.de/videoAnalysisExample.

a key method within ethnographic research (Knoblauch et al., 2014), seeks to capture these fleeting, embodied dynamics e.g., of collaborative environments, analyzing the shared glances, the rhythm of cooperation, and the subtle shifts in physical orientation that define the social order of creative work. Historically, analyzing this "choreography of interaction" was a micro-sociological endeavor, requiring researchers to manually annotate seconds of footage. Modern Multimodal AI, capable of processing video, audio, and text simultaneously, allows for the operationalization of this method at scale (Google, 2025). By treating video as a high-dimensional data stream, computational tools can identify patterns of tacit knowledge transfer that remain invisible in transcripts alone (see example in Figure 2).

Figure 2 provides a proof-of-concept for this capability. Here, a multimodal model analyzes a video stream of a hand counting down from 5 to 1. Crucially, the AI does not merely classify a static image but correctly interprets the *temporal sequence* of the gestures, explicating exactly which anatomical changes (e.g., "The thumb tucks in") correspond to the numerical descent. This demonstrates the potential to automate the coding of complex, time-dependent and detailed physical interactions in educational settings.

4.2 Application scenario: the "Attention Heatmap" in teacher training

To illustrate the transformative potential of this approach, consider a study examining Peer-to-Peer Learning in a Creator Space or Innovation Lab. The research question is: How are knowledge and skill acquisition structured between novices and experts without explicit instruction?

The research team collects 50 h of video footage from open sessions. In a traditional videographic study, the researcher would rely on "purposeful sampling," selecting snippets based on intuition or limited observation.

In a computational workflow, the researcher uses AI as a "Pre-Screening Ethnographer." The model is prompted to sift through the footage to identify specific "Interactional Modes." For example, a prompt might request: "Identify all sequences of 'Scaffolded Mimicry'—defined as one participant silently demonstrating a technique while another observes with focused gaze, followed by the observer attempting the same motion." Alternatively, AI could map the "Attention Economy" of the room, outlining how the group's joint attention shifts from a whiteboard (abstract planning) to a physical prototype (embodied problem solving). AI identifies and curates a set of "decisive moments," enabling the researcher to conduct a process-oriented analysis of skill acquisition. This analysis focuses not on the summative assessment of the final product, but on the observable, embodied sequence through which proficiency is developed. This allows the human researcher to skip hours of irrelevant footage and focus their "thick description" (Geertz, 1973) on these critical moments.

4.3 Critical reflection on opportunities and challenges

While powerful, this "Interpretive Sifting" introduces significant validity challenges. The primary hurdle is "Interpretive Opacity": while an AI can flag a behavior (e.g., a student looking down), the meaning of that gesture is to some extent culturally dependent and context-specific (Banaruee and Khatin-Zadeh, 2026; Banaruee et al., 2024; Farsani et al., 2022; Yu, 2008), particularly bodily experiences which are the realization of abstract concepts. According to multi-layered perspective of embodied language and abstract concepts (Banaruee and Khatin-Zadeh, 2026), language is in one layer universally embodied and in another layer culturally-specific. This is the convergence of both layers which leads to the accurate processing and comprehension of an abstract or metaphorically embodied information. Take for example, a student looking down might be disengaged, being sad, showing respect, or concentrating deeply; such realizations are both context- and culture-specific. AI "explication" of video therefore requires heavy human validation to avoid misinterpretation; the machine identifies the signal, but the human must ascribe the meaning. Furthermore, analyzing e.g., classroom video data via cloud-based AI models raises significant data privacy and ethical concerns. The processing of biometric data (face, voice, gaze), especially of minors, requires strict anonymization protocols and, ideally, local processing solutions to ensure that the "surveillance" capacity of the tool does not undermine the pedagogical safety of the learning environment.

5 Embodied cognition, language, and communication

5.1 From amodal symbols to grounded cognition

Embodied approaches in cognitive science reject the idea that cognition is primarily the manipulation of amodal, abstract symbols. Instead, they posit that conceptual structures are grounded in patterns of sensorimotor, affective, and social experience (Banaruee and Khatin-Zadeh, 2026; Barsalou, 2008, 2010; Khatin-Zadeh et al., 2023a,b,c,d; Lakoff and Johnson, 1980, 1999; Wilson, 2002). Perception and action are not external to "higher" cognition but constitute its very fabric. Gibson's (1977, 1986) ecological psychology introduced the concept of *affordances*, action possibilities offered by the environment relative to an organism's bodily capacities. Affordances link perception directly to action and have been influential in discussions of learning environments, tools, and educational artifacts. Within grounded cognition frameworks, concepts are seen as partially constituted by simulations of such perception–action cycles (Barsalou, 2008, 2010). Conceptual processing involves re-enactments or simulations of prior bodily engagements with the world. Neuroscientific findings lend strong support to this view. Studies of action words and action-related language show somatotopically organized activations in motor and premotor cortex (Aziz-Zadeh and Damasio, 2008; Hauk et al., 2004; Pulvermüller, 2005). Fischer and Zwaan (2008) review evidence that language comprehension recruits modality-specific systems, including motor simulation, suggesting that understanding linguistic meaning involves reenacting bodily states and actions rather than decoding abstract symbols.

5.2 Embodied metaphor, abstract concepts, and social interaction

Embodiment is central not only for concrete actions but also for abstract thought and metaphor (Khatin-Zadeh et al., 2023d). Conceptual Metaphor Theory proposes that abstract domains (e.g., time, emotion, morality) are structured through mappings from bodily grounded source domains (e.g., motion, space, temperature; Lakoff and Johnson, 1980; Lakoff and Turner, 1989). For instance, emotional states are often conceptualized as vertical movements ("feeling down," "uplifted"), or as physical burdens. Empirical work on metaphor interpretation demonstrates that understanding metaphors involves embodied simulations and sensorimotor constraints (Gibbs, 2006a,b; Yu, 2008; Banaruee et al., 2017).

Recent research has extended embodiment to the representation of abstract concepts more generally. Borghi and colleagues argue that even highly abstract concepts are grounded in complex constellations of sensorimotor experience, linguistic co-occurrence, and social interaction, with words functioning as *social tools* for coordinating actions and shared understanding (Borghi and Binkofski, 2014; Borghi, 2018; Borghi et al., 2018). In educational settings, this implies that learning abstract domains—such as mathematics or ethics—depends not just on symbolic representation but on the embodied, dialogical practices through which learners engage with these domains.

5.3 Gesture, multimodality, and extended cognition

Language and thought are inherently multimodal. Gesture research has shown that hand movements, head movements, and facial expressions are integral components of cognitive processes and communication (McNeill, 1992, 2005; Kendon, 2004; Goldin-Meadow, 2003; Hostetter and Alibali, 2008). Gestures can reveal aspects of thinking that are not yet stabilized in speech, scaffold reasoning, and shape how others interpret utterances. Gibbs and Colston (1995) and Johnson and Rohrer (2007) highlight how image schemas, recurring sensorimotor structures, are expressed both linguistically and gesturally. In addition, in education, gestures and bodily actions have been shown to contribute to concept formation and meaning-making (e.g., Goldin-Meadow, 2003; Khatin-Zadeh et al., 2022b, 2023c; Krause and Farsani, 2022; Rosa and Farsani, 2021). Studies of metaphorical and iconic gestures demonstrate how emotions, intensifiers, and mathematical concepts are systematically embodied through specific hand, head, and eyebrow movements (Khatin-Zadeh et al., 2022a, 2023a,b). These findings reinforce the view that communication is not merely verbal but deeply embodied and distributed. Extended and distributed cognition frameworks further argue that cognitive processes extend into tools, artifacts, and environments (Clark and Chalmers, 1998; Clark, 2008). Writing, diagrams, digital devices, and other material resources are not just outputs of cognition but part of the cognitive system itself. In educational contexts, this suggests that learning is distributed across bodies, inscriptions, technologies, and social interactions, and that studying learning requires attention to this broader system.

5.4 Implications for educational research and computational methods

From an embodied and extended perspective, educational phenomena, such as identity formation, epistemic stances, or classroom interaction, are realized through bodily, affective, and socio-material processes. Research methods must therefore attend to language, gesture, gaze, spatial arrangements, and the affordances of materials. Traditional qualitative methods (e.g., ethnography, conversation analysis, videography) are well suited to this task but are resource-intensive and difficult to scale. Multimodal AI and LLMs, however, can operate on text, audio, and video, providing new ways to model and analyze embodied meaning-making at scale. If concepts are grounded in distributed sensorimotor and social patterns, then computational models that capture semantic similarity, metaphorical structure, gesture; speech coordination, and interactional dynamics can serve as proxies for aspects of embodied cognition, provided their use remains theoretically informed and methodologically reflexive.

In the following sections, we present three methodological approaches, Semantic Similarity Rating, AI-based Qualitative Content Analysis, and Computational Ethnography via multimodal video analysis, as concrete instantiations of this embodied, computational hermeneutics.

6 Discussion and conclusion

The computational methods outlined in this article do not merely offer a way to process more data faster; they drastically alter the resolution at which we can observe educational phenomena. While researchers have long employed mixed methods to bridge the gap between statistical generalizability and interpretive depth, these approaches have typically operated in parallel, using interviews to contextualize survey data, or coding schemes to quantify open responses. The "computational turn" proposed here moves beyond this parallel structure. By integrating methods like the ones we introduced, we can now operationalize hermeneutics mathematically, preserving the semantic complexity of human expression while subjecting it to rigorous, scalable analysis. Among these numerous possibilities, we highlight the following two pivotal areas as illustrative examples of the paradigm shift:

First, it enables the shift from static snapshots to dynamic trajectory mapping. Conventional assessment models, whether qualitative or quantitative, often capture competence at discrete intervals. Computational hermeneutics allows us to visualize learning as a continuous motion through semantic space. We can now mathematically model the "arc" of a learner's changing worldview, identifying non-linear patterns, such as periods of regression, confusion, or sudden conceptual leaps, that outcome-based measurement often discards as noise. This allows researchers to study the process of change in diverse settings more broadly and at scale, from professional development workshops to informal creator spaces, validating the erratic nature of learning and education itself. Second, these tools open the door to real-time hermeneutic feedback. Currently, deep qualitative insights often arrive too late to benefit the cohort under study. The speed of computational analysis creates the possibility of a "pedagogical loop" where an AI system acts as a real-time mirror for the learning environment. Ultimately, these computational methods empower educational science to defend the complexity of human development in a digital age. They allow us to reject the reductionist urge to collapse a learner's identity into a single score. Instead, we can now build a science that honors the nuance, diversity, and unpredictability of the learning process.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

DA: Writing – original draft, Software, Investigation, Visualization, Validation, Conceptualization, Methodology. NA: Conceptualization, Validation, Investigation, Software, Writing – original draft, Methodology, Visualization. DF: Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. The authors thank NTNU for paying for the APC.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor HB declared a past co-authorship with the author DF.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. AI was at the heart of forming Figure 2 in the manuscript as this manuscript has a potential methodological orientation (which incorporates, instrumentalises, and addresses AI).

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Aziz-Zadeh, L., and Damasio, A. (2008). Embodied semantics for actions: findings from functional brain imaging. *J. Physiol.* 102, 35–39. doi: 10.1016/j.jphysparis.2008.03.012

Banaruee, H., Farsani, D., Khatin-Zadeh, O. and Eskandari, Z. (2024). Metaphors across cultures. *Poznan Stud. Contemp. Linguist.* 60, 133–158. doi: 10.1515/psicl-2023-0035

Banaruee, H., and Khatin-Zadeh, O. (2026). *Embodied Language Across Cultures: Life and Death in German,* Persian, and English. London: Palgrave Macmillan.

Banaruee, H., Khoshsima, H., Khatin-Zadeh, O., and Askari, A. (2017). Suppression of semantic features in metaphor comprehension. *Cogent Psychol.* 4, 1–6. doi: 10.1080/23311908.2017.1409323

Barsalou, L. W. (2008). Grounded cognition. *Ann. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2010). Grounded cognition: past, present, and future. *Topics Cogn. Sci.* 2, 716–724. doi: 10.1111/j.1756-8765.2010.01115.x

Bittermann, A., and Fischer, A. (2024). Natural language processing in psychology. *Z. Psychol.* 232, 143–146. doi: 10.1027/2151-2604/a000568

Borghi, A. M. (2018). Affordances, context and sociality. *Synthese* 199, 12485–12515. doi: 10.1007/s11229-018-02044-1

Borghi, A. M., Barca, L., Binkofski, F., and Tummolini, L. (2018). Varieties of abstract concepts: development, use, and representation in the brain. *Philos. Trans. Royal Soc. B Biol. Sci.* 373:20170121. doi: 10.1098/rstb.2017.0121

Borghi, A. M., and Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts.* New York: Springer.

Bui, M. D., Holtermann, C., Hofmann, V., Lauscher, A., and von der Wense, K. (2025). "Large language models discriminate against speakers of German dialects," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, 8223–8251. doi: 10.48550/ARXIV.2509.13835

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7

Dell'Acqua, F., Ayoubi, C., Lifshitz, H., Sadun, R., Mollick, E., Mollick, L., et al. (2025). *The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise*. Boston, MA: Harvard Business School.

Farsani, D., Lange, T., and Meaney, T. (2022). Gestures, systemic functional grammar and mathematics education. *Mind Cult. Act.* 29, 75–95. doi: 10.1080/10749039.2022.2060260

Fischer, M. H., and Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 825–850. doi: 10.1080/17470210701623605

Fischer, A. (2025). "Die Zukunft der Qualitativen Inhaltsanalyse im Zeitalter Künstlicher Intelligenz," in *Ein Ablaufmodell KI-basierter Inhaltsanalyse*. f-bb-online.

Fullan, M., and Scott, G. (2014). *Education PLUS*. Available online at: https://www.michaelfullan.ca/wp-content/uploads/2014/09/Education-Plus-A-Whitepaper-July-2014-1.pdf (Accessed December 9, 2022).

Geertz, C. (1973). "Thick description: toward an interpretive theory of culture," in *The Interpretation of Cultures: Selected Essays* (New York: Basic Books), 3–30.

Gibbs, R. W. (2006a). *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press.

Gibbs, R. W. (2006b). Metaphor interpretation as embodied simulation. *Mind Lang.* 21, 434–458. doi: 10.1111/j.1468-0017.2006.00285.x

Gibbs, R. W., and Colston, H. (1995). The cognitive psychological reality of image schemas and their transformations. *Cogn. Linguist.* 6, 347–378. doi: 10.1515/cogl.1995.6.4.347

Gibson, J. J. (1977). "The theory of affordances," in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology,* eds. R. E. Shaw and J. Bransford (Mahwah, NJ: Lawrence Erlbaum), 67–82.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. London: Psychology Press.

Glaser, B. G., and Strauss, A. L. (2008). *Grounded theory: Strategien qualitativer Forschung., 2. korrigierte Aufl.* Bern: Huber.

Goldin-Meadow, S. (2003). *Hearing Gesture: How our Hands Help us Think.* Cambridge, MA: Belknap Press of Harvard University Press.

Google (2025). *Gemini 3 Pro Model Card.* New York: Google.

Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9

Hofmann, V., Kalluri, P. R., Jurafsky, D., and King, S. (2024). AI generates covertly racist decisions about people based on their dialect. *Nature* 633, 147–154. doi: 10.1038/s41586-024-07856-5

Hostetter, A. B., and Alibali, M. W. (2008). Visible embodiment: gestures as simulated action. *Psychon. Bull. Rev.* 15, 495–514. doi: 10.3758/PBR.15.3.495

Johnson, M., and Rohrer, T. (2007). "We are live creatures: embodiment, American pragmatism, and the cognitive organism," in *Body, language and mind: Volume 1. Embodiment*, eds. T. Ziemke, J. Zlatev, and R. Frank (Berlin: Mouton de Gruyter), 17–54.

Junker, A., Nortvedt, G. A., and Farsani, D. (2024). Patterning strategies in grade 1 students with low and high number sense proficiency. *Educ. Stud. Math.* 118, 29–51. doi: 10.1007/s10649-024-10341-5

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Pres.

Khatin-Zadeh, O., Farsani, D., and Banaruee, H. (2022a). A study of the use of iconic and metaphoric gestures with motion-based, static space-based, static object-based, and static event-based statements. *Behav. Sci.* 12, 1–10. doi: 10.3390/bs12070239

Khatin-Zadeh, O., Farsani, D., Hu, J., Eskandari, Z., and Banaruee, H. (2023c). Gestural embodiment of intensifiers in iconic, metaphoric, and beat Gestures. *Behav. Sci.* 13:174. doi: 10.3390/bs13020174

Khatin-Zadeh, O., Farsani, D., Hu, J., Eskandari, Z., Zhu, Y. and Banaruee, H. (2023d). A review of studies supporting metaphorical embodiment. *Behav. Sci.* 13:585. doi: 10.3390/bs13070585

Khatin-Zadeh, O., Farsani, D., Hu, J., Farina, M., Banaruee, H., and Marmolejo-Ramos, F. (2023a). Distributed embodiment of metaphorical hope in hand, head, and eyebrow gestures. *Front. Psychol.* 14:1139881. doi: 10.3389/fpsyg.2023.1139881

Khatin-Zadeh, O., Farsani, D., and Reali, F. (2022b). A study of using metaphoric and beat gestures with motion-based and non-motion-based metaphors during retelling stories. *Behav. Sci.* 12:129. doi: 10.3390/bs12050129

Khatin-Zadeh, O., Hu, J., Banaruee, H., and Marmolejo-Ramos, F. (2023b). How emotions are metaphorically embodied: measuring hand and head action strengths of typical emotional states. *Cogn. Emot.* 37, 486–498. doi: 10.1080/02699931.2023.2181314

Knoblauch, H., Tuma, R., and Schnettler, B. (2014). "Video analysis and videography," in *The SAGE Handbook of Qualitative Data Analysis*, ed. U. Flick (London: SAGE Publications), 435–450.

Koller, H. C. (2017). "Bildung as a transformative process," in *Transformative Learning Meets Bildung: an International Exchange*, eds. A. Laros, T. Fuhr, and E. W. Taylor (Rotterdam Boston Taipei: Sense Publishers), 33–43.

Krause, K., and Farsani, D. (2022). More than words: an integrated framework for exploring gestures' role in bilinguals' linguistic practices for making mathematical meaning. *Canad. J. Sci. Math. Technol. Educ.* 22, 773–705. doi: 10.1007/s42330-022-00253-y

Kuckartz, U., and Rädiker, S. (2023). *Qualitative Content Analysis: Methods, Practice and Software,* 2nd Edn. Los Angeles, London, New Delhi, Singapore, Washington, DC, Melbourne: SAGE.

Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Lakoff, G., and Turner, M. (1989). *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago, IL: University of Chicago Press.

Laros, A., Fuhr, T., and Taylor, E. W. (Eds.). (2017). *Transformative Learning Meets Bildung: An International Exchange*. Rotterdam Boston Taipei: Sense Publishers.

Likert, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.* 22, 1–55.

Maier, B. F., Aslak, U., Fiaschi, L., Rismal, N., Fletcher, K., Luhmann, C. C., et al. (2025). LLMs reproduce human purchase intent via semantic similarity elicitation of Likert ratings. *arXiv* [Preprint]. *arXiv:2510.08338.* doi: 10.48550/arXiv.2510.08338

Mayring, P. (2015). "Qualitative content analysis: theoretical background and procedures," in *Approaches to Qualitative Research in Mathematics Education*, eds. A. Bikner-Ahsbahs, C. Knipping, and N. Presmeg (Dordrecht: Springer Netherlands), 365–380.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought.* Chicago, IL: The University of Chicago Press.

Nikolaou, G., Mencattini, T., Crisostomi, D., Santilli, A., Panagakis, Y., and Rodola', E. (2025). Language models are injective and hence invertible. *arXiv* [Preprint]. *arXiv:2510.15511.* doi: 10.48850/arXiv.2510.15511

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582. doi: 10.1038/nrn1706

Rosa, M., and Farsani, D. (2021). Two fish moving in their seas: how does the body language of teachers show itself who teach mathematical equations? *Acta Sci.* 23, 141–168. doi: 10.17648/acta.scientiae.6391

Rosa, M., Farsani, D., and Silva, C. (2020). Mathematics education, body and digital games: the perception of body-proper opening up horizons of mathematical knowledge constitution. *Math. Teach. Res. J.* 12, 310–324.

Toffler, A. (2022). *Future Shock*. New York: Penguin Random House LLC.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Turner, M. (2001). *Cognitive Dimensions of Social Science*. Oxford: Oxford University Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322

Yu, N. (2008). "Metaphor from body and culture," in *The Cambridge handbook of metaphor and thought,* ed. R. W. Gibbs, Jr. (Cambridge: Cambridge University Press), 247–261.

Ziemke, T. (2003). "What's that thing called embodiment?," in *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, eds. R. Alterman and D. Kirsch (Mahwah, NJ: Lawrence Erlbaum), 1305–1310.

Zlatev, J. (2009). Levels of meaning, embodiment, and communication. *Cybern. Hum. Knowing* 16, 149–174.