Multimodal Methods for Analyzing Learning and Training Environments: A Systematic Literature Review

CLAYTON COHN, Vanderbilt University, USA
EDUARDO DAVALOS, Vanderbilt University, USA
CALEB VATRAL, Tennessee State University, USA
JOYCE HORN FONTELES, Vanderbilt University, USA
HANCHEN DAVID WANG, Vanderbilt University, USA
AUSTIN COURSEY, Vanderbilt University, USA
SURYA RAYALA, Vanderbilt University, USA
ASHWIN T S, Vanderbilt University, USA
MEIYI MA, Vanderbilt University, USA
GAUTAM BISWAS, Vanderbilt University, USA

This document serves as the appendix to our literature review on multimodal methods applied to learning and training environments. It provides supplementary material not included in the main manuscript, including a comprehensive table of all publications in the review corpus, a detailed description of the literature search and screening procedures, and a discussion of the limitations of the review.

CCS Concepts: • **Applied computing → Education**; **Computer-assisted instruction**; **Interactive learning environments**; **Collaborative learning**; **E-learning**; **Computer-managed instruction**;

Additional Key Words and Phrases: multimodal data, data analytics, learning analytics, multimodal learning analytics, mmla, learning environments, training environments

ACM Reference Format:
Clayton Cohn, Eduardo Davalos, Caleb Vatral, Joyce Horn Fonteles, Hanchen David Wang, Austin Coursey, Surya Rayala, Ashwin T S, Meiyi Ma, and Gautam Biswas. 2025. Multimodal Methods for Analyzing Learning and Training Environments: A Systematic Literature Review. 1, 1 (December 2025), 22 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

This work is supported under National Science Foundation grants IIS-2327708, DRL-2112635, and IIS-2017000; and US Army CCDC Soldier Center Award #W912CG2220001. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or United States Government, and no official endorsement by either party should be inferred.
Authors' Contact Information: Clayton Cohn, clayton.a.cohn@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Eduardo Davalos, eduardo.davalos.anaya@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Caleb Vatral, cvatral@tnstate.edu, Tennessee State University, Nashville, TN, USA; Joyce Horn Fonteles, joyce.h.fonteles@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Hanchen David Wang, hanchen.wang.1@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Austin Coursey, austin.c.coursey@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Surya Rayala, surya.chand.rayala@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Ashwin T S, ashwindixit9@gmail.com, Vanderbilt University, Nashville, TN, USA; Meiyi Ma, meiyi.ma@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA; Gautam Biswas, gautam.biswas@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA.

A Corpus Table

Table 1 enumerates the 122 papers in this literature review's corpus.

UUID	First Author	Title	Year	Publication	Corpus
2456887548 [1]	Alyuz	An Unobtrusive And Multimodal Approach For Behavioral Engagement Detection Of Students	2017	MIE	A
818492192 [2]	Andrade	Understanding Student Learning Trajectories Using Multimodal Learning Analytics Within An Embodied-Interaction Learning Environment	2017	LAK	A
425012016 [3]	Anton	The Human Condition: Modal and Interactive Advantages of Teacher over AI Feedback on Children's Mathematical Performance	2025	IDC	B
3637456466 [4]	Ashwin	Impact Of Inquiry Interventions On Students In E-Learning And Classroom Environments Using Affective Computing Framework	2020	UMUAI	A
3448122334 [6]	Aslan	Investigating The Impact Of A Real-Time, Multimodal Student Engagement Analytics Technology In Authentic Classrooms	2019	CHI	A
2668965770 [5]	Aslan	Exploring kid space in the wild: a preliminary study of multimodal and immersive collaborative play-based learning experiences	2022	ETRD	B
1886134458 [7]	Azcona	Personalizing Computer Science Education By Leveraging Multimodal Learning Analytics	2018	FIE	A
3146393211 [9]	Birt	Mobile Mixed Reality For Experiential Learning And Simulation In Medical And Health Sciences Education	2018	Information	A
1326191931 [12]	Chan	Multimodal Learning Analytics In A Laboratory Classroom	2019	MLPALA	A
4089325423 [13]	Chan	Predicting behavior change in students with special education needs using multimodal learning analytics	2023	Access	B
2936220551 [14]	Chango	Multi-Source And Multimodal Data Fusion For Predicting Academic Performance In Blended Learning University Courses	2020	CEE	A
4277812050 [15]	Chango	Improving Prediction Of Students' Performance In Intelligent Tutoring Systems Using Attribute Selection And Ensembles Of Different Multimodal Data Sources	2021	JCHE	A

1196965665 [16]	Chejara	How to build more generalizable models for collaboration quality? lessons learned from exploring multi-context audio-log datasets using multimodal learning analytics	2023	LAK	B
1731146538 [17]	Chejara	Impact of window size on the generalizability of collaboration quality estimation models developed using Multimodal Learning Analytics	2023	LAK	B
1426267857 [19]	Chen	Affect, Support, And Personal Factors: Multimodal Causal Models Of One-On-One Coaching	2021	JEDM	A
2764645776 [20]	Chen	MindScratch: A Visual Programming Support Tool for Classroom Learning Based on Multimodal Generative AI	2025	IJHCI	B
328477558 [18]	Chen	Unpacking help-seeking process through multimodal learning analytics: A comparative study of ChatGPT vs Human expert	2025	CompEdu	B
1225141845 [21]	Cheung	Exploring students’ multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools	2025	JSET	B
3304069824 [24]	Civit	Class integration of ChatGPT and learning analytics for higher education	2024	Expert Sys	B
3809293172 [25]	Closser	Blending Learning Analytics And Embodied Design To Model Students’ Comprehension Of Measurement Using Their Actions, Speech, And Gestures	2021	IJCCI	A
570697424 [26]	Cohn	A multimodal approach to support teacher, researcher and AI collaboration in STEM+ C learning environments	2025	BJET	B
3537775194 [29]	Contero	Personalized and Timely Feedback in Online Education: Enhancing Learning with Deep Learning and Large Language Models	2025	MTI	B
4019205162 [27]	Cornide-Reyes	Introducing Low-Cost Sensors Into The Classroom Settings: Improving The Assessment In Agile Practices With Multimodal Learning Analytics	2019	Sensors	A
2846172025 [28]	Cosentino	Generative AI and multimodal data for educational feedback: Insights from embodied math learning	2025	BJET	B
1576545447 [30]	Cukurova	Artificial Intelligence And Multimodal Data In The Service Of Human Decision-Making: A Case Study In Debate Tutoring	2019	BJET	A
1609706685 [31]	Di Mitri	Learning Pulse: A Machine Learning Approach For Predicting Performance In Self-Regulated Learning Using Multimodal Data	2017	LAK	A

2070224207 [77]	Di Mitri	Detecting Medical Simulation Errors With Machine Learning And Multimodal Data	2019	CAIM	A
3009548670 [33]	Di Mitri	Real-Time Multimodal Feedback With The Cpr Tutor	2020	AIED	A
1763513559 [32]	Di Mitri	Keep Me In The Loop: Real-Time Feedback With Multimodal Data	2021	IJAIED	A
1296637108 [35]	Echeverria	Towards Collaboration Translucence: Giving Meaning To Multimodal Group Data	2019	CHI	A
1040787959 [36]	Echeverria	TeamSlides: A multimodal teamwork analytics dashboard for teacher-guided reflection in a physical learning space	2024	LAK	B
1581261659 [38]	Emerson	Early Prediction Of Visitor Engagement In Science Museums With Multimodal Learning Analytics	2020	ICMI	A
1598166515 [37]	Emerson	Multimodal Learning Analytics For Game-Based Learning	2020	BJET	A
4035649049 [39]	Fernández-Nieto	Storytelling With Learner Data: Guiding Student Reflection On Multimodal Team Data	2021	TLT	A
151988148 [41]	Fernández-Nieto	Data storytelling editor: A teacher-centred tool for customising learning analytics dashboard narratives	2024	LAK	B
483140962 [42]	Fwa	Investigating Multimodal Affect Sensing In An Affective Tutoring System Using Unobtrusive Sensors	2018	PPIG	A
4278392816 [43]	Giannakos	Multimodal Data As A Means To Understand The Learning Experience	2019	IJIM	A
2243240858 [44]	Goslen	Llm-based student plan generation for adaptive scaffolding in game-based learning environments	2025	IJAIED	B
853680639 [46]	Henderson	Sensor-Based Data Fusion For Multimodal Affect Detection In Game-Based Learning Environments	2019	EDM	A
3398902089 [50]	Järvelä	What Multimodal Data Can Tell Us About The Students' Regulation Of Their Learning Process?	2019	LAI	A
86191824 [48]	Jiang	Examining How Different Modes Mediate Adolescents' Interactions During Their Collaborative Multimodal Composing Processes	2019	ILE	
141378338 [47]	Jiang	How Did the Generative Artificial Intelligence-Assisted Digital Multimodal Composing Process Facilitate the Production of Quality Digital Multimodal Compositions: Toward a Process-Genre Integrated Model	2025	TESQ	B

2166765216 [49]	Jin	Chatting with a learning analytics dashboard: The role of generative AI literacy on learner interaction with conventional and scaffolding chatbots	2025	LAK	B
2280467946 [66]	Kim	Multimodal Writing Evaluation in Digital Storytelling using Video-Based Output: Comparing performance of AI and Human Raters.	2024	ICMET	B
32184286 [51]	Kubsch	Once More With Feeling: Emotions In Multimodal Learning Analytics	2022	MMLA Handbook	A
205660768 [52]	Larmuseau	Multimodal Learning Analytics To Investigate Cognitive Load During Online Problem Solving	2020	BJET	A
1877483551 [56]	Lee-Cultura	Motion-Based Educational Games: Using Multi-Modal Data To Predict Player'S Performance	2020	COG	A
3660066725 [53]	Lee-Cultura	Children'S Play And Problem Solving In Motion-Based Educational Games: Synergies Between Human Annotations And Multi-Modal Data	2021	IDC	A
3856280479 [54]	Lee-Cultura	Children'S Play And Problem-Solving In Motion-Based Learning Technologies Using A Multi-Modal Mixed Methods Approach	2021	IJCCI	A
962997360 [57]	Lehtonen	Multimodal Communication and Peer Interaction during Equation-Solving Sessions with and without Tangible Technologies	2023	MTI	B
2429627610 [59]	Lin	Advancing self-directed learning in STEM education: integrating GPT-based learning aid with multimodal learning analytics	2025	JRTE	B
227355655 [58]	Lin	Recognitions of image and speech to improve learning diagnosis on STEM collaborative activity for precision education	2024	EIT	B
804659204 [65]	Liu	Towards Smart Educational Recommendations With Reinforcement Learning In Classroom	2018	TALE	A
3783339081 [64]	Liu	A Novel Method For The In-Depth Multimodal Analysis Of Student Learning Trajectories In Intelligent Tutoring Systems	2018	JLA	A
3796180663 [63]	Liu	Learning Linkages: Integrating Data Streams Of Multiple Modalities And Timescales	2018	JCAL	A
1161441004 [62]	Liu	Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing	2024	CompEdu	B

518268671 [67]	López	Using Multimodal Learning Analytics To Explore Collaboration In A Sustainability Co-Located Tabletop Game	2021	ECGBL	A
566043228 [11]	Ma	Automatic Student Engagement In Online Learning Environment Based On Neural Turing Machine	2021	IJIET	A
3754172825 [68]	Ma	Detecting Impasse During Collaborative Problem Solving With Multimodal Learning Analytics	2022	LAK	A
147203129 [70]	Mangaroska	Multimodal Learning Analytics To Inform Learning Design: Lessons Learned From Computing Education	2020	JLA	A
603534886 [69]	Mangaroska	Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity	2022	JCAL	B
1847468084 [71]	Martin	Computationally Augmented Ethnography: Emotion Tracking And Learning In Museum Games	2019	ICQE	A
2879332689 [72]	Martinez-Maldonado	From Data To Insights: A Layered Storytelling Approach For Multimodal Learning Analytics	2020	CHI	A
549526582 [73]	Martinez-Maldonado	Lessons learnt from a multimodal learning analytics deployment in-the-wild	2023	TOCHI	B
2737776963 [75]	Milesi	"It's Really Enjoyable to See Me Solve the Problem like a Hero": GenAI-enhanced Data Comics as a Learning Analytics Tool	2024	CHI EA	B
1552158788 [76]	Mills	Smart glasses for 3D multimodal composition	2025	LMT	B
1278817005 [78]	Moon	Using multimodal learning analytics as a formative assessment tool: Exploring collaborative dynamics in mathematics teacher education	2024	JCAL	B
2155422499 [79]	Morell	A Multimodal Analysis Of Pair Work Engagement Episodes: Implications For Emi Lecturer Training	2022	JEAP	A
190066185 [80]	Mzwri	Bridging LMS and Generative AI: Dynamic Course Content Integration (DCCI) for Connecting LLMs to Course Content–The Ask ME Assistant	2025	JCE	B
2273914836 [81]	Nasir	Many Are The Ways To Learn Identifying Multi-Modal Behavioral Profiles Of Collaborative Learning In Constructivist Activities	2022	IJCSCL	A
1469065963 [82]	Nguyen	Examining Socially Shared Regulation And Shared Physiological Arousal Events With Multimodal Learning Analytics	2022	BJET	A

3224774131 [83]	Nguyen	Providing Automated Feedback on Formative Science Assessments: Uses of Multimodal Large Language Models	2025	LAK	B
3888330750 [40]	Nieto	Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling	2022	LAK	B
2345021698 [84]	Noël	Exploring Collaborative Writing Of User Stories With Multimodal Learning Analytics: A Case Study On A Software Engineering Course	2018	Access	A
2609260641 [85]	Noël	Visualizing Collaboration In Teamwork: A Multimodal Learning Analytics Platform For Non-Verbal Communication	2022	DAMLE	A
2497456347 [87]	Ochoa	The Rap System: Automatic Feedback Of Oral Presentation Skills Using Multimodal Analysis And Low-Cost Sensors	2018	LAK	A
2634033325 [86]	Ochoa	Controlled Evaluation Of A Multimodal System To Improve Oral Presentation Skills In A Real Learning Setting	2020	BJET	A
3051560548 [88]	Olsen	Temporal Analysis Of Multimodal Data To Predict Collaborative Learning Outcomes	2020	BJET	A
116733479 [89]	Ouyang	Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course	2023	ETHE	B
2005607968 [123]	Ouyang	Multimodal learning analytics of collaborative patterns during pair programming in higher education	2023	ETHE	B
2995141815 [90]	Ouyang	An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective	2023	IJCSCL	B
123412197 [91]	Papamitsiou	Utilizing Multimodal Data Through Fsqca To Explain Engagement In Adaptive Learning	2020	TLT	A
85990093 [93]	Petukhova	Multimodal Markers Of Persuasive Speech : Designing A Virtual Debate Coach	2017	INTERSPEECH	A
957160695 [92]	Petukhova	Virtual Debate Coach Design: Assessing Multimodal Argumentation Performance	2017	ICMI	A
1374035721 [94]	Pham	Attentivelearner2: A Multimodal Approach For Improving Mooc Learning On Mobile Devices	2017	AIED	A

2836996318 [95]	Pham	Predicting Learners’ Emotions In Mobile Mooc Learning Via A Multimodal Intelligent Tutor	2018	ITS	A
3135645357 [96]	Prieto	Multimodal Teaching Analytics: Automated Extraction Of Orchestration Graphs From Wearable Sensor Data	2018	JCAL	A
3408664396 [97]	Psaltis	Multimodal Student Engagement Recognition In Prosocial Games	2017	T-CIAIG	A
3308658121 [98]	Reilly	Exploring Collaboration Using Motion Sensors And Multi-Modal Learning Analytics	2018	EDM	A
1500258376 [99]	Sabuncuoglu	Developing a multimodal classroom engagement analysis dashboard for higher-education	2023	PACM HCI	B
1844320601 [100]	Santhosh	Gaze-Driven Adaptive Learning System with ChatGPT-Generated Summaries	2024	Access	B
3625722965 [74]	Sanusi	Table Tennis Tutor: Forehand Strokes Classification Based On Multimodal Data And Neural Networks	2021	Sensors	A
2000036002 [102]	Sharma	Predicting Learners’ Effortful Behaviour In Adaptive Assessment Using Multimodal Data	2020	LAK	A
261302708 [55]	Sharma	Multimodal teacher dashboards: Challenges and opportunities of enhancing teacher insights through a case study	2023	TLT	B
780281159 [103]	Smith	Multimodal composing with generative AI: Examining preservice teachers’ processes and perspectives	2025	CompComp	B
1118315889 [104]	Spikol	Using Multimodal Learning Analytics To Identify Aspects Of Collaboration In Project-Based Learning	2017	CSCL	A
3339002981 [106]	Spikol	Estimation Of Success In Collaborative Learning Based On Multimodal Learning Analytics Features	2017	ICALT	A
1637690235 [105]	Spikol	Supervised Machine Learning In Multimodal Learning Analytics For Estimating Success In Project-Based Learning	2018	JCAL	A
3796643912 [107]	Standen	An Evaluation Of An Adaptive Learning System Based On Multimodal Affect Recognition For Learners With Intellectual Disabilities	2020	BJET	A
2181637610 [108]	Starr	Toward Using Multi-Modal Learning Analytics To Support And Measure Collaboration In Co-Located Dyads	2018	ICLS	A
1315379489 [109]	Sümer	Multimodal Engagement Analysis From Facial Videos In The Classroom	2021	TAC	A

3093310941 [110]	Tanaka	Embodied Conversational Agents For Multimodal Automated Social Skills Training In People With Autism Spectrum Disorders	2017	PLOS	A
1345598079 [111]	Tancredi	Intermodality In Multimodal Learning Analytics For Cognitive Theory Development: A Case From Embodied Design For Mathematics Learning	2022	MMLA Handbook	A
1687167932 [112]	Tang	Using multimodal analytics to systemically investigate online collaborative problem-solving	2022	DistEdu	B
1285699194 [34]	Tang	A multimodal analysis of college students' collaborative problem solving in virtual experimentation activities: A perspective of cognitive load	2023	JCHE	B
433919853 [115]	Tisza	Understanding Fun In Learning To Code: A Multi-Modal Data Approach	2022	IDC	A
1770989706 [117]	Vrzakova	Focused Or Stuck Together: Multimodal Patterns Reveal Triads' Performance In Collaborative Problem Solving	2020	LAK	A
2055153191 [118]	Vujovic	Round Or Rectangular Tables For Collaborative Problem Solving? A Multimodal Learning Analytics Study	2020	BJET	A
3095923626 [120]	Worsley	A Multimodal Analysis Of Making	2017	IJAIED	A
3309250332 [119]	Worsley	(Dis)Engagement Matters: Identifying Efficacious Learning Practices With Multimodal Learning Analytics	2018	LAK	A
666050348 [121]	Worsley	Multicraft: A Multimodal Interface For Supporting And Studying Learning In Minecraft	2021	HCII	A
1441411748 [61]	Wu	Enhancing self-directed learning and Python mastery through integration of a large language model and learning analytics dashboard	2025	BJET	B
3313249608 [122]	Xu	Classroom Simulacra: Building Contextual Student Generative Agents in Online Education for Learning Behavioral Simulation	2025	CHI	B
3522635517 [124]	Yan	Evidence-based multimodal learning analytics for feedback and reflection in collaborative learning	2024	BJET	B
1019093033 [125]	Yang	Prime: Block-Wise Missingness Handling For Multi-Modalities In Intelligent Tutoring Systems	2019	MMM	A
1436887306 [60]	Yeh	Enhancing EFL vocabulary learning with multimodal cues supported by an educational robot and an IoT-Based 3D book	2022	System	B

177743022 [126]	You	AI-Driven Intelligent Learning Companions: A Multimodal Fusion Framework for Personalized Education	2025	WOCC	B
1935812764 [127]	Yusuf	Using multimodal learning analytics to model students' learning behavior in animated programming classroom	2024	EIT	B
1675503665 [128]	Zapata	AI and peer reviews in higher education: students' multimodal views on benefits, differences and limitations	2025	TPE	B
2737977054 [8]	Zhang	Can AI-generated pedagogical agents (AIPA) replace human teacher in picture book videos? The effects of appearance and voice of AIPA on children's learning	2025	EIT	B
209328204 [130]	Zhao	METS: Multimodal learning analytics of embodied teamwork learning	2023	LAK	B
3602263061 [129]	Zhao	Towards automated transcribing and coding of embodied teamwork communication through multimodal learning analytics	2024	BJET	B

Table 1. Each of the 122 works in our corpus.

B Corpus Distillation Procedure

This appendix contains a detailed account of the steps we took to gather relevant works for our literature review and how we distilled the initial search results to the 73 and 49 papers for Corpora A and B, respectively.

B.1 Literature Search

The literature search for both corpora was based on search strings collaboratively defined and agreed upon by the authors as representative of the target research space. Rather than conducting queries manually, we used SerpAPI [101], a third-party Google Scholar scraping API selected for its ability to return organic search results—unlike alternatives such as `scholarly` [23] and `gscholar` [116], whose outputs differ from browser-based queries.

For Corpus A, we queried Google Scholar via API for papers published between January 2017 and October 2022. The 2017 cutoff was chosen to capture developments from the past five years while excluding earlier foundational work, which is discussed in Sections 1 and 2 of the main manuscript but not included in the corpus. The Corpus B search was conducted in August 2025 and backdated to begin in November 2022, covering the period following the release of ChatGPT. We timed the search to follow major conference publication cycles (LAK, AIED, EDM, and L@S) to ensure comprehensive coverage.

The Corpus A search included 14 distinct phrases, each queried three times using variations of the word *multimodal* (*multimodal*, *multi-modal*, and *multi modal*) as prefixes.[1] For Corpus B, we used 12 updated queries reflecting recent developments in GenAI and LLMs, employing only the standard spelling of *multimodal* after confirming that alternative spellings had no impact on results. We also omitted broad terms such as "multimodal survey" and "multimodal literature review," which surfaced naturally in other targeted searches. The complete list of search phrases is shown in Table 2.

Table 2. Full Corpus Search Terms

education technology
explainable artificial intelligence
learning analytics
learning environments
learning environments literature review
learning environments survey
literature review
simulation environments
survey
training environments
training environments literature review
training environments survey
tutoring systems
xai

(a) Corpus A Search Terms

education technology
learning analytics
learning environments
training environments
simulation environments
llm learning environments
llm training environments
llm learning analytics
pedagogical agents
llm pedagogical agents
ChatGPT in education
generative AI in education

(b) Corpus B Search Terms

For each search string, we collected the top five pages (100 publications) returned by Google Scholar. This top-5 cutoff was imposed for practical and financial reasons related to the subsequent construction of a citation graph (see

[1] The term "xai" was included to identify works on explainable AI in learning and training contexts; however, no relevant results were returned during the initial search.

Appendix B.2.1). SerpAPI limits citation queries to 20 citations per API call, requiring multiple calls for highly cited papers (e.g., five calls for a paper with 100 citations). Without a cutoff, the number of API calls would become intractable.

The initial search yielded 4,200 papers for Corpus A (14 search terms ×3 multimodal spelling variants ×100 results) and 1,200 papers for Corpus B (12 search terms ×1 multimodal spelling variant). The full corpus reduction procedure is detailed in Table 3 and discussed in the following subappendices. Each step is referenced using the corresponding Step ID in Table 3.

Step	Procedure	Removed A	Remain A	Removed B	Remain B
0	Literature search	0	4200	0	1200
1	Remove duplicates	2079	2121	355	845
2	Remove non-English	1	2120	0	845
3	Remove degree-0 nodes/disconnected components	589	1531	33	812
4	Iteratively remove degree-1 nodes	468	1063	253	559
5	Title reads	675	388	305	254
6	Abstract reads	261	127	110	144
7	Full paper reads	54	**73**	95	**49**

Table 3. Corpus reduction procedure.

We removed 2,079 duplicates from Corpus A and 355 from Corpus B by hashing paper titles (Table 3, Step 1), retaining the official published version when multiple copies existed. We then excluded one non-English paper from Corpus A (Step 2), identified using spaCy FastLang [114] and verified through manual inspection. After these steps, the combined search yielded 2,120 unique English-language papers for Corpus A and 855 for Corpus B.

B.2 Study Selection

To reduce the corpora to a reviewable set, we applied both quantitative and qualitative methods. First, we performed citation graph pruning (CGP) to distill the corpus algorithmically (Appendix B.2.1). This was followed by qualitative screening, detailed in Appendix B.2.2.

B.2.1 Citation Graph Pruning (Quantitative Corpus Reduction). For visualization, analysis, and corpus distillation, we used NetworkX [45] to construct a directed citation graph for all remaining papers. Each node corresponds to a paper identified by its Google Scholar UUID, and each directed edge denotes a citation from one corpus paper to another. Following SerpAPI's "cited by" results, only inbound citation queries were required; citations from papers outside the list of remaining papers were ignored.

We first removed all 0-degree nodes and disconnected components (Step 3)—papers that neither cited nor were cited by any other paper in the corpus and components with no edges to or from the primary (i.e., largest by number of nodes) component. Because incoming and outgoing citations jointly determine degree, this approach balances early papers (with few outgoing edges) and recent papers (with few incoming edges). Step 3 removed 589 papers from Corpus A and 33 from Corpus B, resulting in connected citation graphs of 1,531 and 812 papers, respectively.

We then applied iterative degree-1 pruning (Step 4), removing nodes with only one citation edge and repeating the process until none remained. Corpus A required four iterations, removing 468 papers and yielding 1,063; Corpus B required two, removing 253 and yielding 559. This approach allowed us to eliminate loosely connected papers unlikely to be central to the field. Given that multimodal learning and training research spans multiple disciplines (e.g., computer science, education, psychology), the authors agreed that papers with minimal citation connectivity were unlikely to meet the scope of this review. The CGP algorithm is detailed in Algorithm 1.

Algorithm 1 Citation Graph Pruning Algorithm

Require: Acyclic directed graph $G = (V, E)$
procedure Degree Trimming(G, n)
  $S, D \leftarrow \{\}, \{\}$
  for all $v \in V$ **do**
    if $\deg(v) <= n$ **then** $S = S \cup \{v\}$
  for all $v \in S$ **do**
    for all $e \in E$ **do**
      if $v \in e \wedge e \notin D$ **then** $D = D \cup \{e\}$
  return $(V \setminus S, E \setminus D)$
procedure Subconnected Graph Trimming(G)
  $[S_1, S_2, S_3, ..., S_n]$ = ConnectedComponent(G), where each $S_i = (V_i, E_i)$
  $j = \arg\max\{|V_1|, |V_2|, |V_3|, ..., |V_n|\}$
  return (V_j, E_j)
procedure Iterative Trimming(G)
  while True **do**
    G' = DegreeTrimming$(G, 1)$, where $G' = (V', E')$
    if $|V| == |V'|$ **then**
      break
  return (V', E')
G' = DegreeTrimming$(G, 0)$ ▷ Remove 0-deg vertices
G' = SubconnectedGraphTrimming(G') ▷ Keep largest connected subgraph
G' = IterativeTrimming(G') ▷ Iteratively remove 1-deg vertices until equilibrium
return G'

At this point, we concluded the quantitative pruning procedure. The resulting citation graphs served as the basis for subsequent qualitative screening.

B.2.2 Quality Control (Qualitative Corpus Reduction). Following quantitative pruning, qualitative screening further reduced each corpus according to the procedures summarized in Table 3. For Corpus A, the remaining 1,063 papers proceeded through title, abstract, and full-paper review. For Corpus B, due to time constraints, we used an LLM-as-a-Judge workflow [131] for title, abstract, and full-paper decisions, with human verification on the final distilled set.

Title Screening. For Corpus A, four reviewers independently evaluated all 1,063 titles for relevance to multimodal learning or training. Inclusion and exclusion were determined by majority vote, with ties resolved by a fifth reviewer. This resulted in 388 retained titles and 675 exclusions (Table 3, Step 5 for A). For Corpus B, title decisions were made jointly by GPT-4o and Gemini 2.5; agreement between both models determined inclusion or exclusion, and disagreements were adjudicated by a human reviewer. Title screening retained 254 papers, excluding 305 (Step 5 for B).

Abstract Screening. For abstract screening (Step 6), each Corpus A abstract was reviewed by two reviewers using the exclusion criteria in Table 4. Papers without unanimous reviewer agreement underwent a second round of review using majority voting. This yielded 127 retained abstracts and 261 exclusions. For Corpus B, both LLM judges independently evaluated all abstracts under the same criteria. Agreement resulted in automatic inclusion or exclusion; disagreements were resolved by a human reviewer. A total of 144 abstracts were retained.

Full-Paper Screening. Full-paper review followed the same exclusion framework with two additional criteria introduced during reading (Table 5). For Corpus A, 127 papers were divided among five reviewers. Papers were labeled "immediate accept," "immediate exclude," or "borderline." Exclusion required unanimous agreement across all reviewers. After this stage, 73 papers remained (Table 3, Step 7). For Corpus B, both LLM judges evaluated all 144 papers end-to-end using the cumulative exclusion criteria, selecting 79 papers (including a human tie-breaker) for inclusion. Two human reviewers then manually reviewed and discussed each of these papers to assess their alignment with the scope of this review. Based on consensus coding [22], 30 papers were excluded, resulting in a final set of 49 papers. This human-in-the-loop validation ensured that all retained papers met the inclusion criteria and were within the scope of this literature review.

1. Paper does not involve a learning or training environment
2. Environment is VR-only
3. No multimodal data are analyzed
4. No multimodal analysis methods are applied
5. Paper is not original applied research

Table 4. Exclusion criteria for abstract screening.

Across both corpora, qualitative screening ensured that only papers presenting original multimodal methods applied to learning or training environments advanced to the final analysis set: 73 papers for Corpus A and 49 for Corpus B.

1. Results are not informative about learning or training
2. Analysis methods cannot be determined from the manuscript

Table 5. Additional exclusion criteria for full-paper screening.

B.3 Feature Extraction

Feature extraction was performed after the full paper review stages (Table 3, Steps 7) and was conducted manually by two human reviewers for all 73 papers in Corpus A and all 49 papers in Corpus B. Extracted features included identifying information (e.g., title, author, year) and methodological descriptors (e.g., data collection media, modalities, fusion strategies, and analysis methods). Table 6 lists the initial feature set.

To ensure consistency, feature categories were initially discretized through inductive coding [113]. Four reviewers each coded a portion of the papers in Corpus A to define discrete feature sets. For example, "video camera," "webcam," and "Kinect" were consolidated under the medium "video." Reviewers then re-extracted features into these discrete sets. The resulting circumscribing features are shown in Table 7 (Cohen's $\kappa = 0.87$).

A second Corpus A feature extraction round gathered additional features supporting later analysis. These circumscribing features—environment setting, domain, participant interaction structure, didactic nature, level of instruction or training, analysis approach, and analysis results—are listed in Table 8. All were discretized except analysis results, which were recorded in free form for thematic analysis [10]. As with the first round, feature extraction for the second

Feature	Description
UUID	Universally unique identifier on Google Scholar
Title	Publication title
First Author	Publication's first author
Year	Year first publicly available
Environment Type	Type of environment analyzed
Data Collection Media	Types of data collected
Modalities	Modalities used during analysis
Analysis Methods	Methods applied in the analysis
Fusion Type	Data fusion strategies used
Publication Source	Journal, conference, workshop, etc.

Table 6. Initial features extracted from each paper.

Feature	Feature Set
Environment Type	learning, training
Data Collection Media	video, audio, screen recording, eye tracking, logs, physiological sensor, interview, survey, participant produced artifacts, researcher produced artifacts, motion, text
Modalities	affect, pose, gesture, activity, prosodic speech, transcribed speech, qualitative observation, logs, gaze, interview notes, survey, pulse, EDA, body temperature, blood pressure, EEG, fatigue, EMG, participant artifacts, researcher artifacts, audio spectrogram, text, pixel
Analysis Methods	classification, regression, clustering, qualitative, statistical methods, network analysis, pattern extraction
Fusion Type	early, mid, late, hybrid, other

Table 7. First set of circumscribing features and their feature sets.

feature set of Corpus A involved independent coding by two reviewers followed by consensus. For this round, Cohen's κ prior to consensus was 0.71.

Once the feature sets were finalized, this process was applied to Corpus B using two human reviewers for consensus coding (Cohen's $\kappa = 0.68$ prior to consensus). For each corpus, final feature sets represent agreement between the reviewers who coded each paper.

C Literature Review Limitations

This review has three primary limitations: the use of Google Scholar for the literature search, the application of citation graph pruning for corpus reduction, and inconsistencies in versioning across published papers. Each is discussed below.

C.1 Google Scholar

Google Scholar, while widely used in academia and industry, presents challenges for reproducibility. Its proprietary ranking algorithm is opaque and presumably variable, with results influenced by factors such as user location, search

Feature	Feature Set
Environment Setting	physical, virtual, blended, unspecified
Domain of Study	STEM, humanities, psychomotor skills, other, unspecified
Participant Interaction Structure	individual, multi-person
Didactic Nature	instructional, training, informal, unspecified
Level of Instruction or Training	K-12, university, professional development, unspecified
Analysis Approach	model-free, model-based
Feedback	direct, indirect

Table 8. Second set of circumscribing features and their feature sets.

history, the time of query, and possible A/B testing. Because the algorithm is constantly evolving and individualized, exact reproduction of our search results cannot be guaranteed.

To mitigate this, we used SerpAPI—a web scraping service that queries Google Scholar via randomized headers and proxies without user account data (Figure 1). SerpAPI's documentation confirms that no personal information is attached to API requests, and the API supports reproducibility through standardized, non-personalized queries. Additionally, SerpAPI recommends validating results using the included search URLs in browser incognito mode.

We also contacted SerpAPI directly, who confirmed: "No, we don't add any personal information," and noted that others can reproduce results by using the same search parameters. While we agree this may be optimistic given Google's opacity, we are reasonably confident that our initial search was free from personal bias due to the nature of the SerpAPI interface. For reference, all searches were conducted in Nashville, TN, USA.

C.2 Citation Graph Pruning

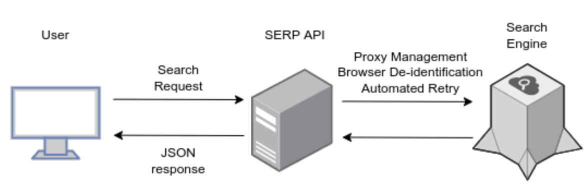

Fig. 1. Searching Google Scholar via SerpAPI.

As detailed in Appendix Section B.2.1, we used citation graph pruning to programmatically reduce our corpus. This approach may have excluded some relevant works that had few citation connections within the corpus. However, our goal was to identify core contributions in the field—papers that either built upon, or were built upon by, other relevant works. We reasoned that isolated papers with few citation ties were less likely to represent foundational or widely used methods.

Importantly, even after CGP, over half the remaining papers were ultimately excluded during qualitative screening. This suggests that CGP still retained many irrelevant works, reinforcing our confidence that the method did not omit significant in-scope works.

C.3 Versioning

Many papers in our corpus appeared in multiple forms across preprint servers and publication venues, often with inconsistent metadata. We used the earliest available public release date when possible. However, discrepancies may remain, particularly for papers from 2022-2023, where preprint and publication dates may differ by months.

As a result, it is possible that some Corpus B papers were written before the public release of ChatGPT (November 2022). Nonetheless, the sharp rise in publications in 2025 (see Figure 1 in the main manuscript) strongly suggests that generative AI was the driving force behind these works, which our own analysis reinforces. Any misclassification in dating is expected to be minor and unlikely to affect our overall findings or conclusions.

References

[1] Nese Alyuz, Eda Okur, Utku Genc, Sinem Aslan, Cagri Tanriover, and Asli Arslan Esme. 2017. An unobtrusive and multimodal approach for behavioral engagement detection of students. In *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education*. ACM, Glasgow UK, 26–32. https://doi.org/10.1145/3139513.3139521

[2] Alejandro Andrade. 2017. Understanding student learning trajectories using multimodal learning analytics within an embodied-interaction learning environment. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, Vancouver British Columbia Canada, 70–79. https://doi.org/10.1145/3027385.3027429

[3] Jacqueline Anton, Giulia Cosentino, Kshitij Sharma, Mirko Gelsomini, Micah Mok, Michail Giannakos, and Dor Abrahamson. 2025. *The Human Condition: Modal and Interactive Advantages of Teacher over AI Feedback on Children's Mathematical Performance*. Association for Computing Machinery, New York, NY, USA, 183–203. https://doi.org/10.1145/3713043.3728863

[4] T. S. Ashwin and Ram Mohana Reddy Guddeti. 2020. Impact of inquiry interventions on students in e-learning and classroom environments using affective computing framework. *User Modeling and User-Adapted Interaction* 30, 5 (Nov. 2020), 759–801. https://doi.org/10.1007/s11257-019-09254-3

[5] Sinem Aslan, Ankur Agrawal, Nese Alyuz, Rebecca Chierichetti, Lenitra M Durham, Ramesh Manuvinakurike, Eda Okur, Saurav Sahay, Sangita Sharma, John Sherry, et al. 2022. Exploring kid space in the wild: a preliminary study of multimodal and immersive collaborative play-based learning experiences. *Educational technology research and development* 70, 1 (2022), 205–230.

[6] Sinem Aslan, Nese Alyuz, Cagri Tanriover, Sinem E. Mete, Eda Okur, Sidney K. D'Mello, and Asli Arslan Esme. 2019. Investigating the Impact of a Real-time, Multimodal Student Engagement Analytics Technology in Authentic Classrooms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300534

[7] David Azcona, I-Han Hsiao, and Alan F. Smeaton. 2018. Personalizing Computer Science Education by Leveraging Multimodal Learning Analytics. In *2018 IEEE Frontiers in Education Conference (FIE)*. IEEE, San Jose, CA, USA, 1–9. https://doi.org/10.1109/FIE.2018.8658596

[8] Jie Bai, Xiulan Cheng, Hui Zhang, Yihang Qin, Tao Xu, and Yun Zhou. 2025. Can AI-generated pedagogical agents (AIPA) replace human teacher in picture book videos? The effects of appearance and voice of AIPA on children's learning. *Education and Information Technologies* (2025), 1–21.

[9] James Birt, Zane Stromberga, Michael Cowling, and Christian Moro. 2018. Mobile Mixed Reality for Experiential Learning and Simulation in Medical and Health Sciences Education. *Information* 9, 2 (Jan. 2018), 31. https://doi.org/10.3390/info9020031

[10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77.

[11] Capital Normal University, Beijing, China, Xiaoyang Ma, Min Xu, Yao Dong, and Zhong Sun. 2021. Automatic Student Engagement in Online Learning Environment Based on Neural Turing Machine. *International Journal of Information and Education Technology* 11, 3 (2021), 107–111. https://doi.org/10.18178/ijiet.2021.11.3.1497

[12] Man Ching Esther Chan, Xavier Ochoa, and David Clarke. 2020. Multimodal Learning Analytics in a Laboratory Classroom. In *Machine Learning Paradigms*, Maria Virvou, Efthimios Alepis, George A. Tsihrintzis, and Lakhmi C. Jain (Eds.). Vol. 158. Springer International Publishing, Cham, 131–156. http://link.springer.com/10.1007/978-3-030-13743-4_8

[13] Rosanna Yuen-Yan Chan, Chun Man Victor Wong, and Yen Na Yum. 2023. Predicting behavior change in students with special education needs using multimodal learning analytics. *IEEE Access* 11 (2023), 63238–63251.

[14] Wilson Chango, Rebeca Cerezo, and Cristóbal Romero. 2021. Multi-source and multimodal data fusion for predicting academic performance in blended learning university courses. *Computers & Electrical Engineering* 89 (Jan. 2021), 106908. https://doi.org/10.1016/j.compeleceng.2020.106908

[15] Wilson Chango, Rebeca Cerezo, Miguel Sanchez-Santillan, Roger Azevedo, and Cristóbal Romero. 2021. Improving prediction of students' performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. *Journal of Computing in Higher Education* 33, 3 (Dec. 2021), 614–634. https://doi.org/10.1007/s12528-021-09298-8

[16] Pankaj Chejara, Luis P Prieto, Maria Jesus Rodriguez-Triana, Reet Kasepalu, Adolfo Ruiz-Calleja, and Shashi Kant Shankar. 2023. How to build more generalizable models for collaboration quality? lessons learned from exploring multi-context audio-log datasets using multimodal learning analytics. In *LAK23: 13th International Learning Analytics and Knowledge Conference*. 111–121.

[17] Pankaj Chejara, Luis P Prieto, Maria Jesus Rodriguez-Triana, Adolfo Ruiz-Calleja, and Mohammad Khalil. 2023. Impact of window size on the generalizability of collaboration quality estimation models developed using Multimodal Learning Analytics. In *LAK23: 13th International Learning*

Analytics and Knowledge Conference. 559–565.

[18] Angxuan Chen, Mengtong Xiang, Junyi Zhou, Jiyou Jia, Junjie Shang, Xinyu Li, Dragan Gašević, and Yizhou Fan. 2025. Unpacking help-seeking process through multimodal learning analytics: A comparative study of ChatGPT vs Human expert. *Computers & Education* 226 (2025), 105198.

[19] Lujie Karen Chen. 2021. Affect, Support, and Personal Factors: Multimodal Causal Models of One-on-one Coaching. *Journal of Educational Data Mining* 13, 3 (2021), 36–68.

[20] Yunnong Chen, Shuhong Xiao, Yaxuan Song, Zejian Li, Lingyun Sun, and Liuqing Chen. 2025. MindScratch: A Visual Programming Support Tool for Classroom Learning Based on Multimodal Generative AI. *International Journal of Human–Computer Interaction* (2025), 1–19.

[21] Kason Ka Ching Cheung, Jack Pun, Wangyin Kenneth-Li, and Jiayi Mai. 2025. Exploring students' multimodal representations of ideas about epistemic reading of scientific texts in generative AI tools. *Journal of Science Education and Technology* 34, 2 (2025), 284–297.

[22] Bonnie Chinh, Himanshu Zade, Abbas Ganji, and Cecilia Aragon. 2019. Ways of Qualitative Coding: A Case Study of Four Strategies for Resolving Disagreements. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312879

[23] Steven A. Cholewiak, Panos Ipeirotis, Victor Silva, and Arun Kannawadi. 2021. *SCHOLARLY: Simple access to Google Scholar authors and citation using Python*. N/A. https://doi.org/10.5281/zenodo.5764801

[24] Miguel Civit, María José Escalona, Francisco Cuadrado, and Salvador Reyes-de Cozar. 2024. Class integration of ChatGPT and learning analytics for higher education. *Expert Systems* 41, 12 (2024), e13703.

[25] Avery H. Closser, John A. Erickson, Hannah Smith, Ashvini Varatharaj, and Anthony F. Botelho. 2022. Blending learning analytics and embodied design to model students' comprehension of measurement using their actions, speech, and gestures. *International Journal of Child-Computer Interaction* 32 (June 2022), 100391. https://doi.org/10.1016/j.ijcci.2021.100391

[26] Clayton Cohn, Caitlin Snyder, Joyce Horn Fonteles, Ashwin TS, Justin Montenegro, and Gautam Biswas. 2025. A multimodal approach to support teacher, researcher and AI collaboration in STEM+ C learning environments. *British Journal of Educational Technology* 56, 2 (2025), 595–620.

[27] Hector Cornide-Reyes, René Noël, Fabián Riquelme, Matías Gajardo, Cristian Cechinel, Roberto Mac Lean, Carlos Becerra, Rodolfo Villarroel, and Roberto Munoz. 2019. Introducing Low-Cost Sensors into the Classroom Settings: Improving the Assessment in Agile Practices with Multimodal Learning Analytics. *Sensors* 19, 15 (July 2019), 3291. https://doi.org/10.3390/s19153291

[28] Giulia Cosentino, Jacqueline Anton, Kshitij Sharma, Mirko Gelsomini, Michail Giannakos, and Dor Abrahamson. 2025. Generative AI and multimodal data for educational feedback: Insights from embodied math learning. *British Journal of Educational Technology* (2025).

[29] Óscar Cuéllar, Manuel Contero, and Mauricio Hincapié. 2025. Personalized and Timely Feedback in Online Education: Enhancing Learning with Deep Learning and Large Language Models. *Multimodal Technologies and Interaction* 9, 5 (2025), 45.

[30] Mutlu Cukurova, Carmel Kent, and Rosemary Luckin. 2019. Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring. *British Journal of Educational Technology* 50, 6 (Nov. 2019), 3032–3046. https://doi.org/10.1111/bjet.12829

[31] Daniele Di Mitri, Maren Scheffel, Hendrik Drachsler, Dirk Börner, Stefaan Ternier, and Marcus Specht. 2017. Learning pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, Vancouver British Columbia Canada, 188–197. https://doi.org/10.1145/3027385.3027447

[32] Daniele Di Mitri, Jan Schneider, and Hendrik Drachsler. 2022. Keep Me in the Loop: Real-Time Feedback with Multimodal Data. *International Journal of Artificial Intelligence in Education* 32, 4 (Dec. 2022), 1093–1118. https://doi.org/10.1007/s40593-021-00281-z

[33] Daniele Di Mitri, Jan Schneider, Kevin Trebing, Sasa Sopka, Marcus Specht, and Hendrik Drachsler. 2020. Real-time multimodal feedback with the CPR tutor. In *International conference on artificial intelligence in education*. Springer, 141–152.

[34] Xu Du, Miao Dai, Hengtao Tang, Jui-Long Hung, Hao Li, and Jinqiu Zheng. 2023. A multimodal analysis of college students' collaborative problem solving in virtual experimentation activities: A perspective of cognitive load. *Journal of Computing in Higher Education* 35, 2 (2023), 272–295.

[35] Vanessa Echeverria, Roberto Martinez-Maldonado, and Simon Buckingham Shum. 2019. Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–16. https://doi.org/10.1145/3290605.3300269

[36] Vanessa Echeverria, Lixiang Yan, Linxuan Zhao, Sophie Abel, Riordan Alfredo, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, Abra Osborne, Simon Buckingham Shum, et al. 2024. TeamSlides: A multimodal teamwork analytics dashboard for teacher-guided reflection in a physical learning space. In *Proceedings of the 14th learning analytics and knowledge conference*. 112–122.

[37] Andrew Emerson, Elizabeth B. Cloude, Roger Azevedo, and James Lester. 2020. Multimodal learning analytics for game-based learning. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1505–1526. https://doi.org/10.1111/bjet.12992

[38] Andrew Emerson, Nathan Henderson, Jonathan Rowe, Wookhee Min, Seung Lee, James Minogue, and James Lester. 2020. Early Prediction of Visitor Engagement in Science Museums with Multimodal Learning Analytics. In *Proceedings of the 2020 International Conference on Multimodal Interaction*. ACM, Virtual Event Netherlands, 107–116. https://doi.org/10.1145/3382507.3418890

[39] Gloria Milena Fernandez-Nieto, Vanessa Echeverria, Simon Buckingham Shum, Katerina Mangaroska, Kirsty Kitto, Evelyn Palominos, Carmen Axisa, and Roberto Martinez-Maldonado. 2021. Storytelling With Learner Data: Guiding Student Reflection on Multimodal Team Data. *IEEE Transactions on Learning Technologies* 14, 5 (Oct. 2021), 695–708. https://doi.org/10.1109/TLT.2021.3131842

[40] Gloria Milena Fernandez Nieto, Kirsty Kitto, Simon Buckingham Shum, and Roberto Martinez-Maldonado. 2022. Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling. In *LAK22: 12th international learning analytics and knowledge conference*. 219–229.

[41] Gloria Milena Fernandez-Nieto, Roberto Martinez-Maldonado, Vanessa Echeverria, Kirsty Kitto, Dragan Gašević, and Simon Buckingham Shum. 2024. Data storytelling editor: A teacher-centred tool for customising learning analytics dashboard narratives. In *Proceedings of the 14th Learning Analytics and Knowledge Conference*. 678–689.

[42] Fwa, Hua Leong and Lindsay Marshall. 2018. Investigating multimodal affect sensing in an Affective Tutoring System using unobtrusive sensors. *Psychology of Programming Interest Group* 29 (Oct. 2018), 78–85.

[43] Michail N. Giannakos, Kshitij Sharma, Ilias O. Pappas, Vassilis Kostakos, and Eduardo Velloso. 2019. Multimodal data as a means to understand the learning experience. *International Journal of Information Management* 48 (Oct. 2019), 108–119. https://doi.org/10.1016/j.ijinfomgt.2019.02.003

[44] Alex Goslen, Yeo Jin Kim, Jonathan Rowe, and James Lester. 2025. Llm-based student plan generation for adaptive scaffolding in game-based learning environments. *International journal of artificial intelligence in education* 35, 2 (2025), 533–558.

[45] Aric Hagberg, Pieter J. Swart, and Daniel A. Schult. 2008. Exploring network structure, dynamics, and function using NetworkX. *N/A* N/A, N/A (1 2008), N/A. https://www.osti.gov/biblio/960616

[46] Nathan L Henderson, Jonathan P Rowe, Bradford W Mott, and James C Lester. 2019. Sensor-based Data Fusion for Multimodal Affect Detection in Game-based Learning Environments. In *Proceedings of the EDM and Games Workshop at the 12th International Conference on Educational Data Mining*, Vol. 2592. International Educational Data Mining Society, Montreal, CA, 1–7.

[47] Lianjiang Jiang and Chun Lai. 2025. How Did the Generative Artificial Intelligence-Assisted Digital Multimodal Composing Process Facilitate the Production of Quality Digital Multimodal Compositions: Toward a Process-Genre Integrated Model. *TESOL Quarterly* (2025).

[48] Shiyan Jiang, Blaine E. Smith, and Ji Shen. 2021. Examining how different modes mediate adolescents' interactions during their collaborative multimodal composing processes. *Interactive Learning Environments* 29, 5 (July 2021), 807–820. https://doi.org/10.1080/10494820.2019.1612450

[49] Yueqiao Jin, Kaixun Yang, Lixiang Yan, Vanessa Echeverria, Linxuan Zhao, Riordan Alfredo, Mikaela Milesi, Jie Xiang Fan, Xinyu Li, Dragan Gasevic, et al. 2025. Chatting with a learning analytics dashboard: The role of generative AI literacy on learner interaction with conventional and scaffolding chatbots. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference*. 579–590.

[50] Sanna Järvelä, Jonna Malmberg, Eetu Haataja, Marta Sobocinski, and Paul A. Kirschner. 2021. What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction* 72 (April 2021), 101203. https://doi.org/10.1016/j.learninstruc.2019.04.004

[51] Marcus Kubsch, Daniela Caballero, and Pablo Uribe. 2022. Once More with Feeling: Emotions in Multimodal Learning Analytics. In *The Multimodal Learning Analytics Handbook*, Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad (Eds.). Springer International Publishing, Cham, 261–285. https://link.springer.com/10.1007/978-3-031-08076-0_11

[52] Charlotte Larmuseau, Jan Cornelis, Luigi Lancieri, Piet Desmet, and Fien Depaepe. 2020. Multimodal learning analytics to investigate cognitive load during online problem solving. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1548–1562. https://doi.org/10.1111/bjet.12958

[53] Serena Lee-Cultura, Kshitij Sharma, Giulia Cosentino, Sofia Papavlasopoulou, and Michail Giannakos. 2021. Children's Play and Problem Solving in Motion-Based Educational Games: Synergies between Human Annotations and Multi-Modal Data. In *Interaction Design and Children*. ACM, Athens Greece, 408–420. https://doi.org/10.1145/3459990.3460702

[54] Serena Lee-Cultura, Kshitij Sharma, and Michail Giannakos. 2022. Children's play and problem-solving in motion-based learning technologies using a multi-modal mixed methods approach. *International Journal of Child-Computer Interaction* 31 (March 2022), 100355. https://doi.org/10.1016/j.ijcci.2021.100355

[55] Serena Lee-Cultura, Kshitij Sharma, and Michail N Giannakos. 2023. Multimodal teacher dashboards: Challenges and opportunities of enhancing teacher insights through a case study. *IEEE Transactions on Learning Technologies* 17 (2023), 181–201.

[56] Serena Lee-Cultura, Kshitij Sharma, Sofia Papavlasopoulou, and Michail Giannakos. 2020. Motion-Based Educational Games: Using Multi-Modal Data to Predict Player's Performance. In *2020 IEEE Conference on Games (CoG)*. IEEE, Osaka, Japan, 17–24. https://doi.org/10.1109/CoG47356.2020.9231892

[57] Daranee Lehtonen, Jorma Joutsenlahti, and Päivi Perkkilä. 2023. Multimodal Communication and Peer Interaction during Equation-Solving Sessions with and without Tangible Technologies. *Multimodal Technologies and Interaction* 7, 1 (2023), 6.

[58] Chia-Ju Lin, Wei-Sheng Wang, Hsin-Yu Lee, Yueh-Min Huang, and Ting-Ting Wu. 2024. Recognitions of image and speech to improve learning diagnosis on STEM collaborative activity for precision education. *Education and Information Technologies* 29, 11 (2024), 13859–13884.

[59] Chia-Ju Lin, Wei-Sheng Wang, Hsin-Yu Lee, Pin-Hui Li, Yueh-Min Huang, and Ting-Ting Wu. 2025. Advancing self-directed learning in STEM education: integrating GPT-based learning aid with multimodal learning analytics. *Journal of Research on Technology in Education* (2025), 1–19.

[60] Vivien Lin, Hui-Chin Yeh, Huai-Hsuan Huang, and Nian-Shing Chen. 2022. Enhancing EFL vocabulary learning with multimodal cues supported by an educational robot and an IoT-Based 3D book. *System* 104 (2022), 102691.

[61] Ming Liu, Zhongming Wu, Haimin Dai, Yifei Su, Laiba Malik, Jian Liao, Wei Zhang, Shuo Guo, Li Liu, and Junqiang Zhao. 2025. Enhancing self-directed learning and Python mastery through integration of a large language model and learning analytics dashboard. *British Journal of Educational Technology* (2025).

[62] Meilu Liu, Lawrence Jun Zhang, and Christine Biebricher. 2024. Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education* 211 (2024), 104977.

[63] Ran Liu, John Stamper, Jodi Davenport, Scott Crossley, Danielle McNamara, Kalonji Nzinga, and Bruce Sherin. 2019. Learning linkages: Integrating data streams of multiple modalities and timescales. *Journal of Computer Assisted Learning* 35, 1 (Feb. 2019), 99–109. https://doi.org/10.1111/jcal.12315

[64] Ran Liu, John C Stamper, and Jodi Davenport. 2018. A Novel Method for the In-Depth Multimodal Analysis of Student Learning Trajectories in Intelligent Tutoring Systems. *Journal of Learning Analytics* 5, 1 (April 2018), 41–54. https://doi.org/10.18608/jla.2018.51.4

[65] Su Liu, Ye Chen, Hui Huang, Liang Xiao, and Xiaojun Hei. 2018. Towards Smart Educational Recommendations with Reinforcement Learning in Classroom. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE, Wollongong, NSW, 1079–1084. https://doi.org/10.1109/TALE.2018.8615217

[66] Sichen Liu and Eunyoung Kim. 2024. Multimodal Writing Evaluation in Digital Storytelling using Video-Based Output: Comparing performance of AI and Human Raters.. In *Proceedings of the 6th International Conference on Modern Educational Technology*. 117–123.

[67] María Ximena López, Francesco Strada, Andrea Bottino, and Carlo Fabricatore. 2021. Using Multimodal Learning Analytics to Explore Collaboration in a Sustainability Co-located Tabletop Game. In *15th European Conference on Game-Based Learning*. Academic Conferences LTD, Brighton, UK, 482–489.

[68] Yingbo Ma, Mehmet Celepkolu, and Kristy Elizabeth Boyer. 2022. Detecting Impasse During Collaborative Problem Solving with Multimodal Learning Analytics. In *LAK22: 12th International Learning Analytics and Knowledge Conference*. ACM, Online USA, 45–55. https://doi.org/10.1145/3506860.3506865

[69] Katerina Mangaroska, Kshitij Sharma, Dragan Gašević, and Michail Giannakos. 2022. Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity. *Journal of Computer Assisted Learning* 38, 1 (2022), 40–59.

[70] Katerina Mangaroska, Kshitij Sharma, Dragan Gašević, and Michalis Giannakos. 2020. Multimodal Learning Analytics to Inform Learning Design: Lessons Learned from Computing Education. *Journal of Learning Analytics* 7, 3 (Dec. 2020), 79–97. https://doi.org/10.18608/jla.2020.73.7

[71] Kit Martin, Emily Q. Wang, Connor Bain, and Marcelo Worsley. 2019. Computationally Augmented Ethnography: Emotion Tracking and Learning in Museum Games. In *Advances in Quantitative Ethnography*, Brendan Eagan, Morten Misfeldt, and Amanda Siebert-Evenstone (Eds.). Vol. 1112. Springer International Publishing, Cham, 141–153. http://link.springer.com/10.1007/978-3-030-33232-7_12

[72] Roberto Martinez-Maldonado, Vanessa Echeverria, Gloria Fernandez Nieto, and Simon Buckingham Shum. 2020. From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. https://doi.org/10.1145/3313831.3376148

[73] Roberto Martinez-Maldonado, Vanessa Echeverria, Gloria Fernandez-Nieto, Lixiang Yan, Linxuan Zhao, Riordan Alfredo, Xinyu Li, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, et al. 2023. Lessons learnt from a multimodal learning analytics deployment in-the-wild. *ACM Transactions on Computer-Human Interaction* 31, 1 (2023), 1–41.

[74] Khaleel Asyraaf Mat Sanusi, Daniele Di Mitri, Bibeg Limbu, and Roland Klemke. 2021. Table Tennis Tutor: Forehand Strokes Classification Based on Multimodal Data and Neural Networks. *Sensors* 21, 9 (April 2021), 3121. https://doi.org/10.3390/s21093121

[75] Mikaela E Milesi, Riordan Alfredo, Vanessa Echeverria, Lixiang Yan, Linxuan Zhao, Yi-Shan Tsai, and Roberto Martinez-Maldonado. 2024. "It's Really Enjoyable to See Me Solve the Problem like a Hero"': GenAI-enhanced Data Comics as a Learning Analytics Tool. In *Extended abstracts of the CHI conference on human factors in computing systems*. 1–7.

[76] Kathy A Mills and Alinta Brown. 2025. Smart glasses for 3D multimodal composition. *Learning, Media and Technology* 50, 2 (2025), 156–177.

[77] Daniele Di Mitri. 2019. Detecting Medical Simulation Errors with Machine learning and Multimodal Data. In *17th Conference on Artificial Intelligence in Medicine*. Springer International Publishing, Poznan, Poland, 1–6.

[78] Jewoong Moon, Sheunghyun Yeo, Seyyed Kazem Banihashem, and Omid Noroozi. 2024. Using multimodal learning analytics as a formative assessment tool: Exploring collaborative dynamics in mathematics teacher education. *Journal of Computer Assisted Learning* 40, 6 (2024), 2753–2771.

[79] Teresa Morell, Vicent Beltrán-Palanques, and Natalia Norte. 2022. A multimodal analysis of pair work engagement episodes: Implications for EMI lecturer training. *Journal of English for Academic Purposes* 58 (July 2022), 101124. https://doi.org/10.1016/j.jeap.2022.101124

[80] Kovan Mzwri and Márta Turcsányi-Szabo. 2025. Bridging LMS and generative AI: dynamic course content integration (DCCI) for enhancing student satisfaction and engagement via the ask ME assistant. *Journal of Computers in Education* (2025), 1–38.

[81] Jauwairia Nasir, Aditi Kothiyal, Barbara Bruno, and Pierre Dillenbourg. 2021. Many are the ways to learn identifying multi-modal behavioral profiles of collaborative learning in constructivist activities. *International Journal of Computer-Supported Collaborative Learning* 16, 4 (Dec. 2021), 485–523. https://doi.org/10.1007/s11412-021-09358-2

[82] Andy Nguyen, Sanna Järvelä, Carolyn Rosé, Hanna Järvenoja, and Jonna Malmberg. 2023. Examining socially shared regulation and shared physiological arousal events with multimodal learning analytics. *British Journal of Educational Technology* 54, 1 (Jan. 2023), 293–312. https://doi.org/10.1111/bjet.13280

[83] Ha Nguyen and Saerok Park. 2025. Providing Automated Feedback on Formative Science Assessments: Uses of Multimodal Large Language Models. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference*. 803–809.

[84] Rene Noel, Fabian Riquelme, Roberto Mac Lean, Erick Merino, Cristian Cechinel, Thiago S. Barcelos, Rodolfo Villarroel, and Roberto Munoz. 2018. Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course. *IEEE Access* 6 (2018), 67783–67798. https://doi.org/10.1109/ACCESS.2018.2876801

[85] René Noël, Diego Miranda, Cristian Cechinel, Fabián Riquelme, Tiago Thompsen Primo, and Roberto Munoz. 2022. Visualizing Collaboration in Teamwork: A Multimodal Learning Analytics Platform for Non-Verbal Communication. *Applied Sciences* 12, 15 (July 2022), 7499. https://doi.org/10.3390/app12157499

[86] Xavier Ochoa and Federico Dominguez. 2020. Controlled evaluation of a multimodal system to improve oral presentation skills in a real learning setting. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1615–1630. https://doi.org/10.1111/bjet.12987

[87] Xavier Ochoa, Federico Domínguez, Bruno Guamán, Ricardo Maya, Gabriel Falcones, and Jaime Castells. 2018. The RAP system: automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors. In *Proceedings of the 8th International Conference on Learning*

Analytics and Knowledge. ACM, Sydney New South Wales Australia, 360–364. https://doi.org/10.1145/3170358.3170406

[88] Jennifer K. Olsen, Kshitij Sharma, Nikol Rummel, and Vincent Aleven. 2020. Temporal analysis of multimodal data to predict collaborative learning outcomes. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1527–1547. https://doi.org/10.1111/bjet.12982

[89] Fan Ouyang, Mian Wu, Luyi Zheng, Liyin Zhang, and Pengcheng Jiao. 2023. Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 4.

[90] Fan Ouyang, Weiqi Xu, and Mutlu Cukurova. 2023. An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning* 18, 1 (2023), 39–66.

[91] Zacharoula Papamitsiou, Ilias O. Pappas, Kshitij Sharma, and Michail N. Giannakos. 2020. Utilizing Multimodal Data Through fsQCA to Explain Engagement in Adaptive Learning. *IEEE Transactions on Learning Technologies* 13, 4 (Oct. 2020), 689–703. https://doi.org/10.1109/TLT.2020.3020499

[92] Volha Petukhova, Tobias Mayer, Andrei Malchanau, and Harry Bunt. 2017. Virtual debate coach design: assessing multimodal argumentation performance. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM, Glasgow UK, 41–50. https://doi.org/10.1145/3136755.3136775

[93] Volha Petukhova, Manoj Raju, and Harry Bunt. 2017. Multimodal Markers of Persuasive Speech: Designing a Virtual Debate Coach. In *Interspeech 2017*. ISCA, Stockholm, Sweden, 142–146. https://doi.org/10.21437/Interspeech.2017-98

[94] Phuong Pham and Jingtao Wang. 2017. AttentiveLearner2: A Multimodal Approach for Improving MOOC Learning on Mobile Devices. In *Artificial Intelligence in Education*, Elisabeth André, Ryan Baker, Xiangen Hu, Ma. Mercedes T. Rodrigo, and Benedict Du Boulay (Eds.). Vol. 10331. Springer International Publishing, Cham, 561–564. http://link.springer.com/10.1007/978-3-319-61425-0_64

[95] Phuong Pham and Jingtao Wang. 2018. Predicting Learners' Emotions in Mobile MOOC Learning via a Multimodal Intelligent Tutor. In *Intelligent Tutoring Systems*, Roger Nkambou, Roger Azevedo, and Julita Vassileva (Eds.). Vol. 10858. Springer International Publishing, Cham, 150–159. http://link.springer.com/10.1007/978-3-319-91464-0_15

[96] L.P. Prieto, K. Sharma, Ł. Kidzinski, M.J. Rodríguez-Triana, and P. Dillenbourg. 2018. Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning* 34, 2 (April 2018), 193–203. https://doi.org/10.1111/jcal.12232

[97] Athanasios Psaltis, Konstantinos C. Apostolakis, Kosmas Dimitropoulos, and Petros Daras. 2018. Multimodal Student Engagement Recognition in Prosocial Games. *IEEE Transactions on Games* 10, 3 (Sept. 2018), 292–303. https://doi.org/10.1109/TCIAIG.2017.2743341

[98] Joseph M Reilly, Milan Ravenell, and Bertrand Schneider. 2018. Exploring Collaboration Using Motion Sensors and Multi- Modal Learning Analytics. In *Proceedings of the 11th International Conference on Educational Data Mining*. International Educational Data Mining Society, Buffalo, NY, USA, 333–339.

[99] Alpay Sabuncuoglu and T Metin Sezgin. 2023. Developing a multimodal classroom engagement analysis dashboard for higher-education. *Proceedings of the ACM on Human-Computer Interaction* 7, EICS (2023), 1–23.

[100] Jayasankar Santhosh, Andreas Dengel, and Shoya Ishimaru. 2024. Gaze-Driven Adaptive Learning System with ChatGPT-Generated Summaries. *IEEE Access* 12 (2024), 173714–173733.

[101] SerpApi. N/A. Google Scholar API. https://serpapi.com/google-scholar-api. [Accessed 08-02-2024].

[102] Kshitij Sharma, Zacharoula Papamitsiou, Jennifer K. Olsen, and Michail Giannakos. 2020. Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 480–489. https://doi.org/10.1145/3375462.3375498

[103] Blaine E Smith, Amanda Yoshiko Shimizu, Sarah K Burriss, Melanie Hundley, and Emily Pendergrass. 2025. Multimodal composing with generative AI: Examining preservice teachers' processes and perspectives. *Computers and Composition* 75 (2025), 102896.

[104] Daniel Spikol, Emanuele Ruffaldi, and Mutlu Cukurova. 2017. Using Multimodal Learning Analytics to Identify Aspects of Collaboration in Project-Based Learning. In *Making a Difference: Prioritizing Equity and Access in CSCL*, Vol. 1. International Society of the Learning Sciences, Philadelphia, PA USA, 263–270.

[105] Daniel Spikol, Emanuele Ruffaldi, Giacomo Dabisias, and Mutlu Cukurova. 2018. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning* 34, 4 (Aug. 2018), 366–377. https://doi.org/10.1111/jcal.12263

[106] Daniel Spikol, Emanuele Ruffaldi, Lorenzo Landolfi, and Mutlu Cukurova. 2017. Estimation of Success in Collaborative Learning Based on Multimodal Learning Analytics Features. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, Timisoara, Romania, 269–273. https://doi.org/10.1109/ICALT.2017.122

[107] Penelope J. Standen, David J. Brown, Mohammad Taheri, Maria J. Galvez Trigo, Helen Boulton, Andrew Burton, Madeline J. Hallewell, James G. Lathe, Nicholas Shopland, Maria A. Blanco Gonzalez, Gosia M. Kwiatkowska, Elena Milli, Stefano Cobello, Annaleda Mazzucato, Marco Traversi, and Enrique Hortal. 2020. An evaluation of an adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1748–1765. https://doi.org/10.1111/bjet.13010

[108] Emma L Starr, Joseph M Reilly, and Bertrand Schneider. 2018. Toward Using Multi-Modal Learning Analytics to Support and Measure Collaboration in Co-Located Dyads. In *ICLS 2018*. International Society of the Learning Sciences, London, UK, 448–455.

[109] Ömer Sümer, Patricia Goldberg, Sidney D'Mello, Peter Gerjets, Ulrich Trautwein, and Enkelejda Kasneci. 2023. Multimodal Engagement Analysis From Facial Videos in the Classroom. *IEEE Transactions on Affective Computing* 14, 2 (April 2023), 1012–1027. https://doi.org/10.1109/TAFFC.2021.3127692

[110] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PLOS ONE* 12, 8 (Aug. 2017), e0182151. https://doi.org/10.1371/journal.pone.0182151

[111] Sofia Tancredi, Rotem Abdu, Ramesh Balasubramaniam, and Dor Abrahamson. 2022. Intermodality in Multimodal Learning Analytics for Cognitive Theory Development: A Case from Embodied Design for Mathematics Learning. In *The Multimodal Learning Analytics Handbook*, Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad (Eds.). Springer International Publishing, Cham, 133–158. https://link.springer.com/10.1007/978-3-031-08076-0_6

[112] Hengtao Tang, Miao Dai, Shuoqiu Yang, Xu Du, Jui-Long Hung, and Hao Li. 2022. Using multimodal analytics to systemically investigate online collaborative problem-solving. *Distance Education* 43, 2 (2022), 290–317.

[113] David R. Thomas. 2006. A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation* 27, 2 (2006), 237–246. https://doi.org/10.1177/1098214005283748 arXiv:https://doi.org/10.1177/1098214005283748

[114] Thomas Thiebaud". 2020. Spacy FastLang. https://spacy.io/universe/project/spacy_fastlang. [Accessed 08-02-2024].

[115] Gabriella Tisza, Kshitij Sharma, Sofia Papavlasopoulou, Panos Markopoulos, and Michail Giannakos. 2022. Understanding Fun in Learning to Code: A Multi-Modal Data approach. In *Interaction Design and Children*. ACM, Braga Portugal, 274–287. https://doi.org/10.1145/3501712.3529716

[116] Bastian Venthur. 2010. GitHub - venthur/gscholar: Query Google Scholar with Python. https://github.com/venthur/gscholar. [Accessed 08-02-2024].

[117] Hana Vrzakova, Mary Jean Amon, Angela Stewart, Nicholas D. Duran, and Sidney K. D'Mello. 2020. Focused or stuck together: multimodal patterns reveal triads' performance in collaborative problem solving. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 295–304. https://doi.org/10.1145/3375462.3375467

[118] Milica Vujovic, Davinia Hernández-Leo, Simone Tassani, and Daniel Spikol. 2020. Round or rectangular tables for collaborative problem solving? A multimodal learning analytics study. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1597–1614. https://doi.org/10.1111/bjet.12988

[119] Marcelo Worsley. 2018. (Dis)engagement matters: identifying efficacious learning practices with multimodal learning analytics. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. ACM, Sydney New South Wales Australia, 365–369. https://doi.org/10.1145/3170358.3170420

[120] Marcelo Worsley and Paulo Blikstein. 2018. A Multimodal Analysis of Making. *International Journal of Artificial Intelligence in Education* 28, 3 (Sept. 2018), 385–419. https://doi.org/10.1007/s40593-017-0160-1

[121] Marcelo Worsley, Kevin Mendoza Tudares, Timothy Mwiti, Mitchell Zhen, and Marc Jiang. 2021. Multicraft: A Multimodal Interface for Supporting and Studying Learning in Minecraft. In *HCI in Games: Serious and Immersive Games*, Xiaowen Fang (Ed.). Vol. 12790. Springer International Publishing, Cham, 113–131. https://link.springer.com/10.1007/978-3-030-77414-1_10

[122] Songlin Xu, Hao-Ning Wen, Hongyi Pan, Dallas Dominguez, Dongyin Hu, and Xinyu Zhang. 2025. Classroom Simulacra: Building Contextual Student Generative Agents in Online Education for Learning Behavioral Simulation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–26.

[123] Weiqi Xu, Yajuan Wu, and Fan Ouyang. 2023. Multimodal learning analytics of collaborative patterns during pair programming in higher education. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 8.

[124] Lixiang Yan, Vanessa Echeverria, Yueqiao Jin, Gloria Fernandez-Nieto, Linxuan Zhao, Xinyu Li, Riordan Alfredo, Zachari Swiecki, Dragan Gašević, and Roberto Martinez-Maldonado. 2024. Evidence-based multimodal learning analytics for feedback and reflection in collaborative learning. *British Journal of Educational Technology* 55, 5 (2024), 1900–1925.

[125] Xi Yang, Yeo-Jin Kim, Michelle Taub, Roger Azevedo, and Min Chi. 2020. PRIME: Block-Wise Missingness Handling for Multi-modalities in Intelligent Tutoring Systems. In *MultiMedia Modeling*, Yong Man Ro, Wen-Huang Cheng, Junmo Kim, Wei-Ta Chu, Peng Cui, Jung-Woo Choi, Min-Chun Hu, and Wesley De Neve (Eds.). Vol. 11962. Springer International Publishing, Cham, 63–75. http://link.springer.com/10.1007/978-3-030-37734-2_6

[126] Cunqian You, Huijuan Lu, Ping Li, Xiaoyu Zhao, and Yudong Yao. 2025. AI-Driven Intelligent Learning Companions: A Multimodal Fusion Framework for Personalized Education. In *2025 IEEE 34th Wireless and Optical Communications Conference (WOCC)*. IEEE, 424–428.

[127] Abdullahi Yusuf, Norah Md Noor, and Shamsudeen Bello. 2024. Using multimodal learning analytics to model students' learning behavior in animated programming classroom. *Education and Information Technologies* 29, 6 (2024), 6947–6990.

[128] Gabriela C Zapata, Bill Cope, Mary Kalantzis, Anastasia Olga Tzirides, Akash K Saini, Duane Searsmith, Jennifer Whiting, Nikoleta Polyxeni Kastania, Vania Castro, Theodora Kourkoulou, et al. 2025. AI and peer reviews in higher education: students' multimodal views on benefits, differences and limitations. *Technology, Pedagogy and Education* (2025), 1–19.

[129] Linxuan Zhao, Dragan Gašević, Zachari Swiecki, Yuheng Li, Jionghao Lin, Lele Sha, Lixiang Yan, Riordan Alfredo, Xinyu Li, and Roberto Martinez-Maldonado. 2024. Towards automated transcribing and coding of embodied teamwork communication through multimodal learning analytics. *British Journal of Educational Technology* 55, 4 (2024), 1673–1702.

[130] Linxuan Zhao, Zachari Swiecki, Dragan Gasevic, Lixiang Yan, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, Abra Osborne, Xinyu Li, Riordan Alfredo, et al. 2023. METS: Multimodal learning analytics of embodied teamwork learning. In *LAK23: 13th International learning analytics and knowledge conference*. 186–196.

[131] Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric Xing, et al. 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in neural information processing systems* 36 (2023), 46595–46623.