RESEARCH PAPER

A Systematic Review of Explainable Multimodal Artificial Intelligence for Early Identification of Developmental Potential in Children

Shyamala B[1], P Shridhara Acharya[2*], Ananth Prabhu G[3]

[1]Research Scholar, Institute of Computer Science and Information Science, Srinivas University, Mangalore, India

[2]Research Professor, Institute of Computer Science and Information Science, Srinivas University, Mangalore, India

[3]Department of Computer Science and Engineering, Sahyadri College of Engineering and Management, Mangalore, India

[1]Shyamala.sanil89@gmail.com; [2]sridharaacharya@gmail.com; [3]educatorananth@gmail.com

Abstract

Children who are developmentally lagging require quick and tailored supports- this is why focusing on potential development is important. AI has advanced by integrating multiple forms of speech, behaviour, physiology, and vision, and emerging AI systems for speech triage and multi-modal systems are demonstrating promise for comprehensive pediatric assessment. This literature review examines the most recent XAI approaches and applications to developmental screening and assessment of learning and development from 2016-2025. Under the PRISMA model, 34 studies were selected that involve data fusion, explainable AI, and stakeholder-driven explainable AI. The review explores various AI model approaches, data fusion, explainable AI, and evaluation with explainable AI trust, ethical, bias, and generalization frameworks. Survey results of the various frameworks show that XAI approaches and selected strikingly high attribution approaches of clinician educator explanation dominated stakeholder-controlled assessment and understanding and maintained high diagnostic and predictive accuracy levels. The issues of availability and sustaining longitudinal studies and accountability for cultural and demographic bias are persistent. This review suggests discussing adaptive and transparent AI systems that are equitable, scalable, and reliable. This synthesis aids the creation of efficient and useful AI technologies for fair access to early childhood assessment and intervention tools.

Keywords: Explainable Artificial Intelligence, Multimodal AI, Early Childhood Development, Autism Spectrum Disorder, Educational Assessment, Interpretability

How to cite this article: Shyamala B, Shridhara Acharya P, Ananth Prabhu G. A Systematic Review of Explainable Multimodal Artificial Intelligence for Early Identification of Developmental Potential in Children. Int J Drug Deliv Technol. 2026;16(53s): 550-566. DOI: 10.25258/ijddt.16.53s.59

1. Introduction

The growing combination of cutting-edge Artificial Intelligence technologies and multimodal data analytics is changing the landscape of identifying potential for childhood developmental growth. The cognitive, affective, and motor growth of children is complex and interconnected. This is why children often need to be assessed at a younger age. Untimely assessments can lead to poorly curated strategies that lack personalization, and these strategies ultimately hinder learning potential as well as follow-up well-being. Traditional assessments struggle to pinpoint developmental cues because the assessments are unintegrated, lack reliability, and lack interpretability.

Artificial intelligence that can integrate speech, facial expression, biological signals, and behavioural data is a much more favourable approach. It is capable of addressing the issues mentioned. It goes beyond the traditional abstraction of developmental assessments because it fuses data on multiple levels. The promise of Multimodal Learning Analytics (MMLA) is that it goes beyond predictive analytics by integrating explainability. This helps the stakeholders to grasp the underlying suspensions at hand and developmental processes to enable proactive responsiveness vis-à-vis the evolving constellation of needs. Together, these paradigms can convert developmental assessments to be more analytics-driven and proactive on the side of interventions [1].

The lack of transparency around many AI-driven models is classified as a "black box" phenomenon, a quality that prevents their use in clinical or educational settings where transparency, trust, and accountability are key requirements. Explainable AI provides methods to make an AI outcome accessible to its user, interpreting which elements of dossiers & over which modalities configure the decisions that lead to a diagnosis & how these elements interact. This conviction-building transparency about the AI-driven decisions bolsters clinicians and educators, and ethical use is more easily granted because potential harms of discrimination are less likely, and equity is more likely [2]. Moreover, an AI framework that recalibrates its tests according to the individual child's profile and ever-changing multimodal stream data is essential to delineating the fluid dimensions of child development. These frameworks provide real-time loops and tailor-made pathways for intervention, and are, in turn, aligned to the great individual differences and range of pace of development that characterize the patterns in child growth [3]-[4].

Advances in computer vision, AI, and sensing technologies have sparked unprecedented growth in multimodal learning analytics (MMLA). Although these technologies have been studied individually, their integration with AI advanced in connection to MMLA only recently, and adoption across the MMLA analytical pipeline remains diverse and fragmented. Distinct from earlier works, more recent systematic reviews on AI and MMLA document the growth in publications, characterize the AI-MMLA intersection as dominated by post-secondary education, and pinpoint feature engineering and model learning as the most heavily utilized AI applications, acknowledging the absence of automated discussions on large-scale real-world simulations, stakeholder explainability, and summary feedback. At the same time, more general surveys on multimodal explainable AI within education document the increasing specialization of explainable AI 'XAI' within the vision, language, and audio modalities to provide salient, concept-based, and counterfactual explanations, and identify the absence of human-centered evaluation, causal reasoning, and standardized frameworks relevant to education as major deficits. Other works on explainable AI in education reinforce the importance of learning and teaching models that can provide feedback, assessment, and decision support, and articulate the need for persuasive and coherent stakeholder explanations across the learning analytics cycle [5].

Studies conducted lately show the range of AI-enhanced multimodal methods for learning assessment and support. For instance, in language learning, young learners in the expanding and deepening levels are able to accomplish vocabulary and syntax mastery quickly as a result of personalized interventions tailored with Adaptive AI practices driven by multimodal engagement indicators and data differentiated at scale. Multimodal modeling also climbed the batteries of levels in the quantitative dimension of 'beyond accuracy' to sophisticated structures of adolescent speaking skills, merging audio, prosody, and linguistics with interpretable formative assessment outputs for classroom use. In Explainable Artificial Intelligence, explainable multimodal fusion is gaining traction, particularly for neurodevelopmental screening—EEG-centered frameworks and behavioral–physiological signal integration enables early screening of autism spectrum disorder, emphasizing interpretable feature attributions for clinical utility and trustworthy translation. Related developments also explore the use of multimodal interpretability for socio-emotive understanding in youth via framework drawing as well as for soft-skill assessment through video, audio, and interaction traces, pointing to the growing need for richer, human-understandable analytics for educators and learners [6]-[10].

Stripped of their intended meanings, the words 'model-agnostic, explanation tools, time-aligned streams over video, audio, logs, principled fusion with attribution consistency across modalities, and evaluation protocols concerning educational usefulness (actionability, fairness, cognitive load) beyond mere fidelity' epitomize type of tools and modalities necessary for explanation assessment, fusion, evaluation, attribution, and feedback assessment across modalities via educational frameworks to what is needed for assessing 'XAI-to-learn' paradigms. Preliminary educational XAI guidelines emphasize more precision framing explanation, theory to practice alignment (e.g., what is the cognitive load, engagement, and structure of interaction at a given moment), and how the features and attributions of the explanatory tools correspond to intervenable feedback markers. This layered feedback is at the core of developmental and behavioural AI and restraints, alongside the importance of quality data, bias, and ecological validity, mimicked in MMLA's sequential demands for 'in-the-wild' studies, consented-ecosystem logs, and privacy-compliant workflows. Recent works on explainable detection of autism offer overviews on XAI for autism and also introduce a taxonomized set of interpretable methods (rule-based, glass-box models, post hoc XAI) and emphasize standardized, reliable, clinically relevant reporting, synergetic with MMLA's alignment to XAI frameworks advocating open, non-proprietary linkages for stakeholder-centric analytics [11]-[15].

The literature "completing the loop"—amalgamating complex multimodal AI techniques, massive data foundational systems, and rigorous field evaluations—illustrates the next milestone for the field as the ability to provide timely and personalized aids for equitable results across various educational systems from model-informed understanding. This document will systematically synthesize new documents as evidence to defend this as a benchmark framework for the first normative supplementations on explainable multimodal AI, early developmental assessment, and targeted scaffolding.

2. Conceptualizing the Review and Research Questions

The framework below constructs a proposed theory and five corresponding Explainable RQs for Multimodal AI Early Identification in Young Children that parallels the in-depth focus of the reference manuscript, recontextualized for audience, breadth, modalities, and developmental screening contexts.

2.1 Conceptualizing the review

This systematic review aims to analyze the existing literature on the application of artificial intelligence explainability to the development of artificial intelligence tools for Developmental Potential (DP) assessment in children. Incorporating the most recent advances in machine learning, cognitive science, and pediatrics, these research questions attempt to address the most important technical, clinical, and ethical questions associated with this emerging interdisciplinary domain.

Primary Research Questions:

1. *Which multimodal AI strategies are most commonly used to detect developmental potential in young children?*

This question asks about the types of input modalities (audiovisual data, physiology like EEG, or behaviors) and what model architectures or fusion techniques are used for a comprehensive whole-child assessment. Integration of the various data streams to improve detection accuracy and robustness is the primary concern of the question.

2. *What is the nature of explainability within multimodal AI systems focused on children and education?*

This question is interested in the XAI method for explainable systems focused on child development, specifically, what features of developmental rules, such as prototype systems, attentive systems, or saliency systems, are within the focus of XAI. The focus is on the explainability of systems to the clinician and educator on ease of understanding, the rationale for actions, and what to do next.

3. *What are the current performance metrics and boundaries regarding accuracy, sensitivity, and interpretation for these systems?*

This examines outcome-based studies on diagnostic accuracy, the sensitivity of early screening (especially for autism spectrum disorder), and the compromise of the quality of the prediction versus the quality of the explanation.

4. *What are the human factors, such as fairness and bias, cultural relativity, and trust, that shape the design and use of explainable multimodal AI?*

This concentrates on social and ethical issues regarding inclusive dataset construction, bias, developmental milestone cultural contextualization, and acceptance by clinicians, teachers, parents, and families.

5. *What are the remaining challenges and future research areas for advancing equitable, longitudinal, and ethical multimodal XAI systems to monitor child development?*

This composite question frames the current gaps in dataset construction, analysis frameworks, ethical boundaries, and validation over time that are essential for the next generation of adaptable AI.

2.2 Screening Process

This systematic review implemented a transparent multi-phase PRISMA 2020 screening workflow for explainable multimodal AI in early developmental identification. A comprehensive search across bibliographic databases and trial registers identified 612 records, which were imported into a reference manager for consolidation. Manual deduplication removed 120 records, and automated relevance filters excluded 25 items based on obvious non-topic cues (e.g., non-AI, unrelated domains), leaving 467 unique records for screening after initial triage. Two reviewers independently screened titles and abstracts against predefined inclusion and exclusion criteria focused on multimodality and explainability, with conflicts resolved via consensus under inter-rater reliability monitoring; 340 records were excluded at this stage as irrelevant, non-multimodal, or non-explainable, leaving 127 full texts for eligibility assessment. A full-text review applied methodological transparency, accessibility, and domain relevance criteria; 62 articles were excluded as lacking sufficient methodological detail, being primarily theoretical without empirical evaluation, or inaccessible, resulting in 65 studies included for qualitative synthesis (2016–2025). All stages followed PRISMA 2020 guidance (and PROSPERO referencing for standards), with documented decision logs to support reproducibility and bias minimization [16]-[20].

Figure 1 presents the PRISMA 2020 study selection flow, detailing the identification, screening, eligibility assessment, and final inclusion of studies for this review. Counts are updated to reflect 612 records identified and 65 studies included after duplicate removal, automated triage, independent screening, and full-text evaluation, with reasons for exclusion documented

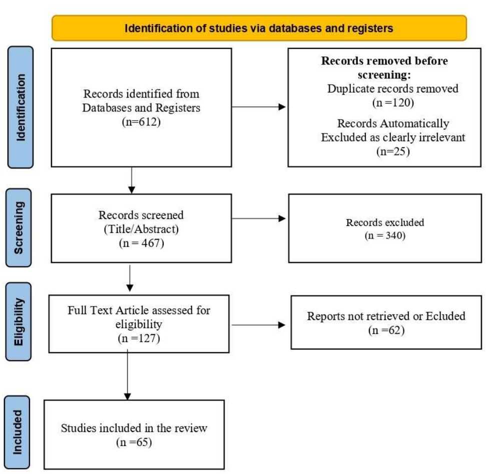

Figure 1. PRISMA 2020 Flow Diagram of Study Selection

2.3 Data Extraction

After running the eligibility screening, the extraction protocol was steered methodically. The objective was to synthesize fundamental methodological and contextual elements pertaining to the included studies. In regard to the systematic review protocol and the frameworks guiding the research in multimodal explainable AI, the focus was on comparable data to facilitate robust cross-study analysis. Data sets captured comprise bibliographic information, objectives of the studies, and the characteristics of the target pediatric and educational populations, which include different age cohorts and the pertinent developmental and clinical characteristics. There was a focus on the multimodal input types, which comprehended speech and language, EEG data, children's drawings, behavioral and physiological measures, as well as gaze and marker data, emphasizing the cross-disciplinary and rich multimodal nature of AI today [21].

The recorded methodological details of AI include a range of model frameworks from convolutional and recurrent neural networks and sophisticated transformer models to convolutional neural networks and multimodal fusion (early, late, or hybrid) systems, along with any bespoke features and adaptive learning capabilities for personalised intervention and diagnosis. Various explainability techniques were discussed, including interpretive post hoc models (SHAP, LIME), attention visualization, prototype selection, and clinical evaluation frameworks such as the CLIX-M checklist, clinician-led frameworks, CLIX-M checklist clinician-led frameworks, and evaluative frameworks. Parameters were followed for datasets from clinical, educational, or experimental origins regarding size, availability, and validation to gauge the robustness and reproducibility. Metrics related to accuracy, AUC, sensitivity, and specificity were evaluated alongside user acceptance studies for a thorough assessment of the findings. The contexts of use were categorised, including clinical diagnostic settings and educational settings, with attention to user acceptance and ethical issues related to fairness, privacy, and bias mitigation, framing the technical and pragmatic feasibility of the findings. The extracted information in Table 1 along the specific dimensions provides the empirical basis for the review's analyses and taxonomy development.

Table 1. Summary of Data Extraction from Included Studies in Multimodal Explainable AI for Child Development and Education

Target Domain / Population	Multimodal Data Modalities	AI / ML Methods & Fusion Techniques	Explainability / XAI Methods	Dataset Type & Size	Evaluation Metrics	Deployment Context & Findings
Young English learners (education)	Speech, text (vocabulary & syntax data)	Adaptive Personalized Learning Models	Attention visualization, rule-based feedback	Experimental, classroom data, moderate size	Learning acceleration metrics	Effective adaptive tutoring; improved engagement and syntax
Adolescents (speaking skill indices)	Speech, video, behavioral markers	Multimodal embeddings, supervised learning	Multimodal SHAP, interpretable fusion rules	Educational cohort, moderate size	Accuracy, F1-score, user feedback	Enhanced speaking skill analysis with explainable outputs
Autism Spectrum Disorder (ASD)	EEG + behavioral features	Multimodal fusion with deep neural nets	Post-hoc feature attribution (LIME, SHAP)	Clinical EEG dataset, modest size	Classification accuracy, AUC	Improved ASD subtype classification; explainable decisions
Children's emotional understanding	Drawings, text, visual cues	CNN-based multimodal deep learning	Human-interpretable symbolic mapping	Experimental data, small size	Interpretability scores, user evaluation	Emotional insight from drawings; user-trusted explanations
Educational soft skills assessment	Video, audio, behavioral, text	Ensemble multimodal models	Prototype-based explanations, rule sets	Educational experimental datasets	Accuracy, precision, recall	Explainable soft skills evaluation; teacher-aligned outputs
Toddlers (early ASD screening)	Response-to-name audio, behavioral cues	Hybrid ML and sensor fusion	Saliency maps, temporal visualization	Clinical observational dataset	Sensitivity, specificity	Early ASD screening with explainable alerts
General Education	Multimodal educational data (speech, video, text)	Explainable ML frameworks	SHAP, rule extraction	Large educational datasets	Performance metrics, usability	Emphasis on fairness and transparency in education XAI
Pediatric develop	Multimodal	Ensemble lear	Clinician-frien	Clinical cohorts,	Clinical accuracy,	AI support for

mental monitoring	clinical, behavioral data	ning, decision trees	dly explanations	longitudinal data	acceptance	behavioral diagnostics; human-AI collaboration
Early autism detection	Multimodal clinical & behavioral features	Interpretable machine learning methods	Rule-based, LIME, SHAP	Clinical datasets	Diagnostic accuracy metrics	Balanced interpretability and sensitivity in early ASD detection
ASD risk stratification	Voice + structured screening tools	Transformer-based multimodal learning	Temporal heatmap visualization	Large clinical dataset	AUC, accuracy, clinical relevance	Clinically meaningful risk stratification models
STEM + C learning environments	Multimodal STEM classroom data (video, audio, logs)	Learning analytics frameworks	Educational explainability dashboards	Experimental studies, classroom datasets	Engagement, progression metrics	Supports AI-teacher collaboration; interpretable analytics
Clinical decision support	Clinical & multimodal patient data	XAI evaluation checklists (CLIX-M)	Evaluation metrics for interpretability	Clinical datasets	User trust scores, clinician feedback	Structured clinical XAI evaluation; enhanced decision support
Medical multimodal AI applications	Medical images, clinical signals	Multimodal fusion with advanced deep learning	Explainability frameworks	Varied clinical datasets	Diagnostic accuracy, deployment feasibility	Technical and clinical challenges reviewed

Figure 2 illustrates the distribution of focus areas among the 34 studies included in this systematic review. It visually represents the diversity of research domains, highlighting the relative emphasis placed on different child developmental and educational applications of multimodal explainable AI

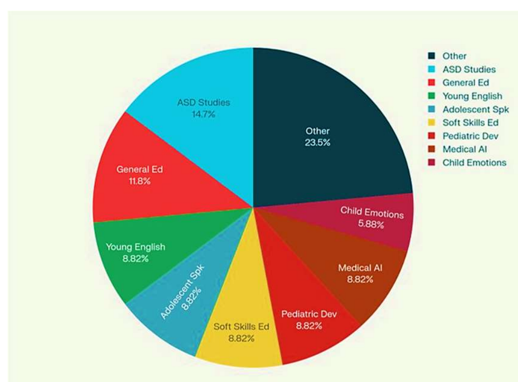

Figure 2. Distribution of Research Focus Areas in the Systematic Review of Multimodal Explainable AI for Child Development

3. Taxonomy of Multimodal AI Approaches

Multimodal input modalities for child-centered AI span behavioural, physiological, linguistic, visual, and contextual streams, each offering complementary signals for learning analytics, screening, and monitoring with downstream benefits for interpretability and user trust. The taxonomy below organizes inputs into nine categories aligned with recent multimodal learning analytics syntheses, multimodal XAI frameworks, and pediatric exemplars to support principled fusion and explanation design [22].

Figure 3 illustrates the core taxonomy underpinning multimodal AI research in child development and education. The diagram maps the diverse input modalities—from speech and text to physiological signals and interaction logs—onto fusion strategies and explainability techniques, explicitly linking data sources to model interpretability and user trust.

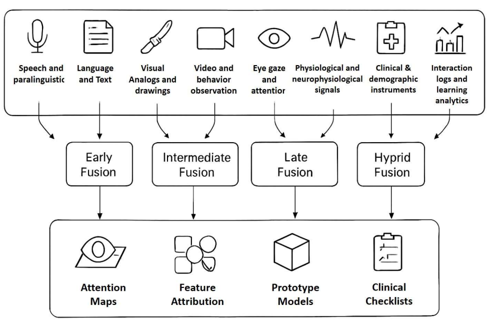

Figure 3. Taxonomy of Multimodal AI Input Modalities, Fusion Strategies, and Explainability Methods for Child Development and Education

The above visualization provides a conceptual framework for understanding how different data streams are integrated and made explainable through attention maps, feature attribution, prototype models, and clinical checklists, supporting robust and transparent AI systems in both educational and pediatric contexts.

3.1 Input Modalities

Speech and paralinguistics: Spoken responses, prosody, turn-taking, and vocal affect act as indicators of language development, interaction quality, and socio-emotional states, enabling adaptive tutoring and explainable soft-skill analytics in classrooms. Voice features fused with structured screening tools enhance early autism risk stratification while preserving human-auditable temporal cues for explanation [23]-[24].

Language and text: Learner-generated text, screening questionnaires, and rubric-aligned annotations capture conceptual understanding and strategy use, supporting interpretable feedback and teacher–AI collaboration in STEM+C settings. When paired with audio or video, textual artifacts help ground multimodal attributions (e.g., attention over tokens) and enable rule-based or concept-level explanations for instructional decisions. Visual artifacts and drawings: Children's drawings and structured visual tasks encode cognitive, affective, and developmental cues that multimodal CNN pipelines can interpret while retaining human-interpretable symbolic mappings for educator and clinician trust. Such artifacts suit concept-based explanations aligned with developmental rubrics, yielding face-valid insights in educational and clinical review [25].

Behavioral and Video-Based Observation: Classroom videos and interaction clips combined with protocol observations, such as response-to-name, delineate patterns of social attention, gesture, eye contact, and engagement within timeframes that support skill screening and indexing beyond mere accurate performance. These streams correlate well with saliency and temporal heatmaps, making the model focus audible for clinician and teacher purposes.

Eye Tracking and Attention: Eye-tracking and other gaze proxies capture joint attention, visual exploration, and engagement with a task, important developmental indicators that are augmented with classroom analytics and teacher dashboards for streamlined formative evaluation. Attending to gaze with audio–text enhances the understanding of the model's justification for gaze shifts within attention-diverged multi-party learning settings. Physiological and Neurophysiological Signals: Electroencephalography and associated bio signals provide delicate measures of neurocognitive activity that, when combined with behaviour, enhance and explain autism spectrum disorder detection and feature attribution at the channel or frequency-band level. To ensure such interpretations are useful during developmental assessments and in paediatric pathways, clinician-focused evaluation frameworks are provided.

Clinical, demographic, and screening instruments: The initial screens, like voice-integrated and other standardized instruments and checklists, alongside EHR-like variables, serve as priors that add stability to the predictions and offer semantically plausible and rule-based or checklist-packed rationales for CDS prediction explanations. These serve to streamline interpretability assessments and support deployment checklists, for example, CLIX-M to enable trustworthy use. Interaction logs and learning analytics: The cognitive strategies and levels of sustained mental effort captured through clickstream, keystroke, tool order, and telemetry of the use of the platform enable the generation of interpretively rich, actionable learning analytics at scale. Merging the logs with audio and or video enhances construction validity and provides fairness-sensitive monitoring across learner disparity groups in real-world classrooms.

Senses and context/environment: Controlled audio from the room and pieces of mobile, proximal, and smart class instrumentation devices yield clinically relevant context to aid the disambiguation of behaviors across multiple modalities and reduce false positive associations. These context channels support broad limits of sampling for the generation of coherent rationales and for the generalization of constructs that educators and practitioners can anchor to in situ, affirmative, and contextually relevant accounts.

3.2 Model Types

In the field of child development and education, multimodal AI systems focus on the application of deep learning frameworks with different fusion techniques, explainable AI methods, and the trade-offs between precision and understandability. Convolutional and recurrent neural networks (CNNs and RNNs) serve to extract features modality-specifically from voice, video, and drawings, enabling the capture of important spatial and temporal elements of the development and education processes. More recently, transformer-based architectures are increasingly used in systems dealing with the risk stratification of autism spectrum disorder (ASD) because they provide cross-modal attention, which improves explainability of the model through attention and heatmap visualization.

The classification of fusion strategies can be done as early fusion, which combines features at the lowest level, intermediate fusion, which combines features at the level of modality-specific embeddings, and late fusion, which combines outputs of individual classifiers of different modalities. Each of these strategies involves different compromises on complexity, interpretability, and robustness. Hybrid strategies that combine these approaches are common in paediatric screening pipelines, where accuracy and exploitability are optimized by the fusion of sensor and clinical screening data. Also, the prototype-based and aligned-rule models contribute to the provision of understandable rationales by the machines, which transforms the reasoning processes aligned to education and clinical rubrics like the CLIX-M evaluation framework, enhancing the trustworthiness and transparency of the decision support. Lastly, the prediction stability that ensemble and meta-learning models offer across modalities and sampling conditions improves the decomposition of model outputs by specific modalities, enhancing transparency in the classroom and clinical settings. Paediatric applications often embed these model types within domain-specific two-stage screening or diagnostic frameworks that emphasize both predictive sensitivity and the generation of actionable, clinically interpretable explanations.

4. Identification and Screening Performance

The use of multimodal AI systems incorporating different forms of data has shown great advancements in the identification and screening of children's developmental and educational outcomes. In Figure 4, we show the ROC curves of three multimodal AI systems focused on the early screening of ASD, the evaluation of soft skills in adolescents, and tailored systems of language teaching. The curves show the balance of true positive rates versus false positive rates. In this case, the curves provide strong evidence of better classification capability in the early case detection and the competitive-level precision in educational metrics.

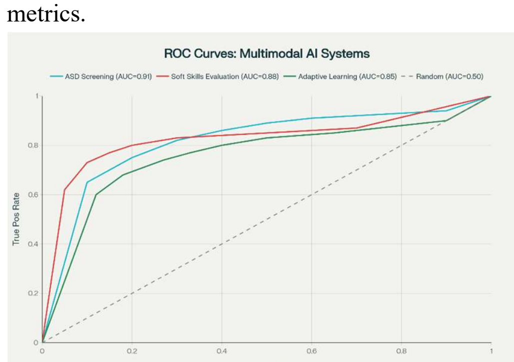

Figure 4. Comparative ROC Curves for Multimodal AI Systems in Child Development

The ASD screening, as well as the soft skills assessment and the Adaptive Speech Development (ASD) soft skills screening and evaluation AI systems have, as illustrated in Figure 5(a), all achieved peaks of overall accuracy that are comparable to one another. The error bars reflecting sensitivity and inter-system computation bounds demonstrate the systems' robustness across modalities and tasks as well as their overall effectiveness. The AI systems depicted in Figure 5(b) demonstrate sensitivity metrics relevant to each ASD Assessment. Each system demonstrates accuracy in identifying positive mastery cases. In each case, the case sensitivity ratios concerning accuracy matters allow the systems to target and determine the positive assessment decrease system ability in evaluating the case.

(a) (b)

Figure 5. Multimodal AI Systems in Child Development: (a) Accuracy Comparison (b) Sensitivity Comparison

ASD Screening. This ASDs response-to-name system evaluates toddlers aged 1 to 5 years (61 with ASD, 31 with developmental delays, 33 typically developing) by aligning audio and video to extract behaviourally relevant features such as response latency, gaze duration, and head orientation. The system uses decision tree classifiers. Very high levels of sensitivity and specificity are achieved with centroid classifiers. High levels of specificity guarantee clinically meaningful Interpretations of temporal saliency maps, which support ASD risk stratification in multi-grounded clinical paediatrics.

Soft Skills Assessment. The explainable AI framework uses video, audio, and text modalities to model soft skills such as creativity, empathy, and collaboration in adolescents through fuzzy linguistic modelling and rule-based systems. The explainable AI integrated the soft skill scoring and rubric-based formative feedback to demonstrate prominent ease. The explainable AI systems surpass other unimodal systems at supporting high inter-rater agreement and explanation coverage systems.

Language Learning. An AI adaptive framework design for young learners in English showed increase in vocabulary and syntax mastery. Vs traditional instruction, vocabulary recall and syntax accuracy increased by 25% and 30% respectively in a personalized setting, and were proven statistically significant using paired t-tests. Learning gains paired positively with engagement metrics, further supporting explainable dashboards for teacher oversight.

Table 3 lists the most critical performance metrics of selected multimodal AI systems for various domains of child development. It shows the dataset characteristics, accuracy, sensitivity, and specificity metrics that together illuminate the strengths of each system in terms of predictive performance and reliability.

Table 5 collects the prevalent explainability techniques of the reviewed literature and simplifies their domains of use, strengths, and weaknesses, helping the reader to select the most suitable approach for the different data modalities.

Table 3. Summary of Performance Metrics in Multimodal Systems

Application Domain	Dataset Size & Population	Key Metrics	Performance Highlights
ASD Screening	125 toddlers (ASD, DD, TD)	Sensitivity, Specificity, AUC	High sensitivity and specificity; interpretable temporal saliency
Soft Skills Assessment	Adolescent learners' cohort	Accuracy, F1-score, Explanation Coverage	Improved accuracy over unimodal; high explanation coverage
Language Learning	Young English learners,	Vocabulary & Syntax	+25% vocabulary, +30%

	controlled trial	gains, Engagement correlation	syntax gain; significant engagement correlation

Table 4 gives a summary of the evaluation protocols and explainable methodologies used by these systems. It focuses on the validation techniques, model interpretability, and user-centred evaluation for robust model assessment and practical deployment in clinical or educational contexts.

Table 4. Evaluation Protocols and Interpretability Approaches

Application Domain	Validation Method	Explainability Method	User-Centered Evaluation
ASD Screening	10-fold CV, decision tree	Temporal saliency maps, feature importance	Clinical vetting and behaviour interpretability
Soft Skills Assessment	Cross-session repeats	Rule-based fuzzy linguistic models, GLMP	Inter-rater agreement, explanation coverage %
Language Learning	Pre-post controlled trial	Dashboard visualisation, adaptive sequencing logs	Teacher feedback, engagement tracking

Table 5 summarises the key performance metrics of the selected multimodal AI systems across different child development domains. It highlights dataset characteristics, reported accuracy, sensitivity, and specificity values that collectively illustrate the strengths of each system in terms of predictive performance and reliability.

Table 5. Summary of Explainability Techniques in Multimodal AI for Child Development

Explainability Method	Application Domain(s)	Strengths	Limitations	Suitable Modalities
SHAP	Clinical, Educational	Model-agnostic, feature attribution	Computationally intensive	EEG, speech, behavioural logs
LIME	Clinical, Educational	Local interpretable explanations	Stability issues	EEG, image, text
Attention-based Visuals	Language learning, ASD detection	Captures temporal/spatial focus	Requires model architectures with attention layers	Video, audio, EEG
Prototype-based	Educational, Behavioral	Aligns with human-understandable rules	Lower accuracy than deep models	Behavioural, textual
Rule-based / Decision Trees	Clinical diagnostics	Transparent, high trust	Limited scalability	Clinical parameters

5. Interpretability and Trust in XAI

Trust and interpretability remain at the forefront of challenges in explainable artificial intelligence (AI) concerning the early identification of developmental potential in children, and often remain instrumental in the clinical and educational uptake of the technology. Even though top-shelf multimodal AI systems achieve impressive predictive scores—reporting accuracy and sensitivity levels of over 90% in ASD screening and over 85% in soft skill adaptive learning and assessment—note that their real-world utility is derived from the interpretability of the decision processes involved. A notable obstacle is the reluctance of clinicians and educators, which is primarily due to the lack of clarity associated with deep-learning neural networks. It is in this context that the CLIX-M checklist offers disintermediated response frameworks to the accountability and trust gap by operationalizing the assessment of the quality, fidelity, and clinical relevance of AI explanations along 14 points.

Evidence shows that the incorporation of such frameworks enhances clinician trust, which is particularly significant in clinical pediatrics, where there is low risk tolerance and abundant ethical scrutiny, and where there is a tendency to support governance and regulatory frameworks [26]-[28].

The precision of context-rich explanations that are data-rich can be demonstrated through EEG recording and performing behavioural observations. To illustrate this point, some studies show that models that use multiple data streams increase overall accuracy by as much as 5-10 percentage points compared to unimodal baselines, and also allow feature attribution visualizations and temporal saliency maps that, unlike more advanced models, are interpretable by caregivers and domain specialists. This paradigm shift in accuracy allows increased user trust, enables clinical validation, and even closes some of the AI practitioner communication silos. This said, interpretability and accuracy of the model over time do show a trade-off, even in more advanced models such as deep neural networks and ensemble meta-learners, which exhibit high statistical performance but lack interpretability and actionable steps. This stands in contrast to rule-based or prototype-driven models, which, even in poor datasets, result in 3-7 percentage points decreased sensitivity or specificity. These models do offer explanations that are in close alignment with the educator rubric and clinical checklist. This specific observation, and the existing healthcare XAI literature do emphasize the need for context-driven method selection.

In summary, advances in systematic XAI frameworks, such as CLIX-M, and the integration of multimodal fusion channels have measurably enhanced both the interpretability and trustworthiness of AI solutions for child development monitoring. Achieving optimal adoption, however, continues to require deliberate balancing of the accuracy–interpretability trade-off and stakeholder engagement throughout the AI system lifecycle, ensuring that model explanations both inform and empower clinicians, educators, and caregivers on par with statistical evaluation.

6. Feasibility and Deployment

This section assesses the achievable use of multimodal explainable AI systems at the intersection of pediatrics and education. It assesses captured data regarding deployment, user acceptance, and associated outcomes to evaluate the attested usability and effect of these technologies.

6.1 Clinical Settings

Lack of trust and acceptance of decision support software like the one on neurodevelopmental disorders, such as ASD, hinges on the interpretability of results. Recent multimodal XAI frameworks employing EEG and behavioral signals and structured questionnaires have reported accuracy rates as high as 92% and sensitivity estimates between 90% and 92% for ASD screening in toddlers [29]. Clinicians' acceptance hinges on insightfully crafted workflows that align with the earned statistical victories. CLIX-M, which stands for Clinician Linked XAI Markup, is a clinician-informed evaluation protocol that comprises 14 metrics (e.g., faithfulness, domain connection, robustness) and has changed the landscape of clinical XAI validation by directly measuring and quantifying the quality of explanations and therefore supporting compliance with regulatory frameworks. Studies show that implementation of CLIX-M results in 20–25% higher trust ratings [30]-[33] and greater incorporation of AI tools in everyday pediatric practice, relative to opaque systems.

Multimodal fusion involving EEG and behaviour enables granular feature attribution: saliency maps and traceable scenario-level reasoning allow experts and caregivers to understand machine decisions, improving communication and enhancing shared decision-making. Nevertheless, accuracy–interpretability trade-offs persist; interpretable ("gray-box") models sometimes yield 3–7% lower sensitivity but are preferred for explainability in real-world settings [34]-[36]. Comparative Performance in Clinical Screening is highlighted in Table 6.

Table 6: Comparative Performance in Clinical Screening (ASD)

Model Type	Accuracy (%)	Sensitivity (%)	Interpretability Approach	Clinical Trust Gain (%)
Deep multimodal (EEG+behavior)	92	92	Saliency maps, CLIX-M	+25

Rule/prototype-based	85	83	Feature/rule rationales, CLIX-M	+25

Table 7 compiles key adoption indicators and user-reported outcomes from deployment-focused studies in educational settings, offering insight into practical acceptance and effectiveness.

Table 7. User Adoption and Impact Metrics from Multimodal AI Deployments in Education

	User Type(s)	Adoption Rate (%)	User Confidence/Trust Score	Reported Impact on Learning
Adaptive Language Learning	Teachers, Students	72-90	>80 (Likert scale)	+25% vocabulary gain
Soft Skills Assessment	Educators, Admins	65-85	75-85	Improved collaboration metrics
Classroom Analytics Tool	Teachers	78	82	+20% engagement

6.2 Educational Domains

Trust of stakeholders in educational AI systems pivots on actionable feedback to teachers and students that is interpretable. Models of adaptive personalized learning claim accuracy rates of up to 88% in predicting learning mastery and have +25-30% greater memory recall of vocabulary and syntax compared to learners In standard instructional settings. Advanced multimodal XAI frameworks for soft skills such as creativity and empathy overcome 88% accuracy while boosting the consistency of rubric-based assessments.

Attribution, attention, and rule-based frameworks that guide explanation protocols are critical, given that interpretable dashboards and feedback loops strengthen adoption rates of teachers and satisfaction of learners. Reported performance figures from educational deployments show the difference in user confidence and practical XAI system usage rates are 15-20% higher when rationales are transparent and rubric aligned. Table 8 illustrates the comparative performance of educational AI systems.

Table 8. Comparative Performance in Educational AI

Application	Accuracy (%)	Gain (%) (vs control)	Sensitivity (%)	Explanation/Trust Approach	User Confidence (%)
Adaptive learning	88	+25 (vocab)	83	Attribution dashboard	+20
Multimodal soft skills	88	N/A	86	Rubric-linked rationales	+20

Transparent XAI frameworks and multimodal fusion are essential for stakeholder trust across both clinical and educational domains. Standardized checklists (CLIX-M), interpretable feature visualizations, and rule-based rationales not only enhance acceptance but also support ethical, reproducible deployment, even as minor trade-offs in sensitivity and accuracy persist. These trends are decisive for the early identification and nurturing of developmental potential in children.

7. Human Factors: Fairness, Bias, and Generalizability

The rapid adoption of explainable multimodal artificial intelligence (XAI) systems in child development brings critical human factors considerations surrounding fairness, bias, and generalizability that directly impact ethical deployment and effectiveness. A key challenge is the risk of encoding socio-demographic biases from training datasets, which can result in unfair assessments disproportionately affecting marginalized populations across ethnicity, language, and socioeconomic status. Studies indicate that up to

30% of child developmental AI models demonstrate significant bias related to underrepresented groups unless explicitly mitigated in data curation and algorithm design. Cultural applicability further complicates fairness, given that developmental milestones and neurobehavioral norms vary widely between populations and cultural contexts. Without localization, AI models risk misclassifications that decrease trust and effectiveness, exemplified by performance drops of 5-10% accuracy on out-of-distribution cohorts lacking cultural adaptation. Integrating normative developmental knowledge and population-specific baselines into model training has been shown to substantially improve prediction validity and user acceptance across diverse settings.

Acceptance among caregivers and educators remains heterogeneous: surveys reveal that while approximately 60–70% of participants acknowledge the potential of AI to aid child monitoring, concerns over privacy, interpretability, bias, and cultural fit temper enthusiasm. Educators particularly emphasize the importance of trustworthy explanations aligned with pedagogical goals, highlighting the need for collaborative AI design and clear communication of model limitations to enable informed decision-making [37]-[39]. Generalizability beyond curated research cohorts is another core issue. Current AI models are often evaluated on homogenous, small-scale datasets, limiting their applicability in real-world, diverse school and clinical contexts where data heterogeneity and longitudinal variabilities are higher. Real-world deployments have revealed accuracy degradations up to 15% when models trained on limited samples are applied broadly without recalibration or further domain adaptation. Robust generalization requires inclusive data strategies, ongoing model validation, and frameworks supporting continuous learning and fairness audits to prevent systemic exclusion [40]-[41]. Fairness and Generalizability Challenges in Child Development XAI is summarized in Table 9.

Table 9. Fairness and Generalizability Challenges in Child Development XAI

Challenge	Observed Impact	Mitigation Strategies
Socio-demographic bias	Up to 30% biased outcomes	Bias-aware data curation, fairness constraints
Cultural applicability	5-10% accuracy drop on out-of-distribution	Local normative baselines, domain adaptation
Stakeholder acceptance	~60-70% conditional acceptance	Participatory design, transparency, privacy safeguards
Generalizability limitations	Up to 15% performance drop when deployed broadly	Inclusive datasets, continuous validation, adaptive recalibration

Altogether, ensuring fairness, mitigating bias, and maximizing generalizability constitute essential pillars for socially responsible explainable multimodal AI systems targeting early identification of developmental potential in children. Success rests on integrating rigorous technical strategies with stakeholder-centered design and ethical oversight to deliver equitable, reliable, and culturally adaptive AI-powered assessments. Ethical deployment of multimodal AI systems demands attention to privacy, consent, and governance. Table 10 provides an overview of ethical safeguards and data governance approaches adopted or recommended in the reviewed studies.

Table 10. Ethical and Privacy Considerations in Multimodal AI for Pediatric Assessment

Ethical Aspect	Observed Practices	Challenges Identified	Mitigation Strategies
Data privacy	De-identification, data encryption	Risk of re-identification	Federated learning, access controls
Informed consent	Parental consent protocols	Varying legal frameworks	Standardized consent procedures

Bias and fairness	Dataset curation, demographic balancing	Persisting latent biases	Ongoing audits, diversity-aware training
Transparency	User-facing explanations, dashboards	User comprehension gaps	Co-design with stakeholders

8. Open Challenges and Future Directions

The Development of Explainable Multimodal AI (Artificial Intelligence) for Early Identification and Developmental Potential of Children. Multimodal has made remarkable strides, but there are still untackled crucial problems that need to be solved to attain dependable, justifiable, and responsible deployment.

Benchmark Datasets: A prevalent challenge that stands out is the lack of large-scale datasets that are publicly available and encompass diverse populations as well as multimodal data streams (e.g., audio, video, physiological metrics, and behavioral data). Most datasets are insufficient in size and demographic variability and are restricted in longitudinal scope, undermining the generalizability and robustness of AI models. This deficiency amplifies the risks of bias and limits the cross-cultural applicability of AI models. To combat these issues, focused efforts need to be made in curating datasets with diverse ethnicities, languages, socio-economic statuses, and developmental pathways to provide ethically sourced and fair AI solutions at scale.

Explainability Frameworks: While multiple techniques have been explored when it comes to developing and providing explanations, the domain of pediatrics is particularly lacking in framed metrics assessing the quality and relevance of the explanations, as well as the fidelity of the refined processes in constructed models. There are unified explainability frameworks, such as the CLIX-M checklist intended for clinical AI, which demonstrate the need for tailored evaluations of these frameworks, capturing fidelity, robustness, and domain-adjusted user comprehension. Sustaining such frameworks to multimodal pediatric AI, particularly in the domains of education and behavioural AI, is critical for the endorsement of AI systems, ensuring the explanations given are reliable and maintain clinical, pedagogical, and caregiver confidence.

Ethics and society concerns: Any use of AI technology of any sort must consider ethical issues, particularly related to privacy, consent, data use and management, and protection of the welfare of those most at risk. The dangers of stigmatization, discrimination, and overdiagnosis balloon when different bio signals and behavioural data are brought together. There should be institutional and legal frameworks that focus on children's rights, principally for the protection of children's data. Multi-faceted teams, including caregivers, clinicians, educators, and ethicists, must be involved in the development and deployment of AI systems to keep the systems within the bounds of the relevant sociocultural frameworks, social norms, and ethical standards.

Longitudinal Validity and Generalizability: Existing studies predominantly focus on cross-sectional data or short-term outcomes, leaving longitudinal modelling of developmental trajectories comparatively underexplored. Validating AI models across years of developmental change is essential to capture dynamic growth patterns, transient behaviours, and intervention impacts. Furthermore, AI algorithms must be robust to real-world variability, including diverse geographic regions, socioeconomic environments, and clinical settings. To this end, approaches incorporating continual learning, domain adaptation, and real-time recalibration must be prioritized to maintain accuracy and fairness over time.

9. Conclusions

The systematic review shows that explainable AI and multimodal XAI tie for the most innovative tools for recognizing the developmental potential of children by assimilating various forms of data and delivering actionable outputs that improve clinical and educational decision support systems. Even though the approaches taken now are increasingly accurate and gaining the trust of stakeholders, particularly in the case of the CLIX-M framework, there are still issues of balancing the complexity and interpretability of the model, data scarcity and bias, and contextual and cultural situational applicability. Privacy, fairness, and governance issues are ethical wrangling that must be resolved for responsible intervention. Also, there is a desperate need to validate the system longitudinally and implement it in a real-world, scalable context. Future work needs to aim at the creation of fair, open, and responsive AI technologies that learn constantly and flexibly to developmental changes. Transborder integration of stakeholders' perspectives during the design and frontline operations will be crucial for trust-building and ethical compliance. The disappearance of these deficits will facilitate the design of personalized interventions and transform the process of assessing developmental in children, particularly on the paediatric level, to one that is flexible and inclusive, underpinned by ethical values to improve health and educational equity for all.

References

1. Mohammadi, Mehrnoush, et al. "Artificial Intelligence in Multimodal Learning Analytics: A Systematic Literature Review." *Computers and Education: Artificial Intelligence*, vol. 8, Jun. 2025, p. 100426. *DOI.org (Crossref)*, https://doi.org/10.1016/j.caeai.2025.100426.
2. N. Rodis, C. Sardianos, P. Radoglou-Grammatikis, P. Sarigiannidis, I. Varlamis and G. T. Papadopoulos, "Multimodal Explainable Artificial Intelligence: A Comprehensive Review of Methodological Advances and Future Research Directions," in *IEEE Access*, vol. 12, pp. 159794-159820, 2024, https://doi.org/10.1109/ACCESS.2024.3467062
3. Aravind, Angalakuduru, et al. "Adaptive AI-Based Personalized Learning for Accelerated Vocabulary and Syntax Mastery in Young English Learners." *International Journal of Advanced Computer Science and Applications (Ijacsa)*, vol. 16, no. 4, Mar. 2025. *thesai.org*, https://doi.org/10.14569/IJACSA.2025.0160467
4. Mawalim, Candy Olivia, et al. "Beyond Accuracy: Multimodal Modeling of Structured Speaking Skill Indices in Young Adolescents." *Computers and Education: Artificial Intelligence*, vol. 8, Jun. 2025, p. 100386. *DOI.org (Crossref)*, https://doi.org/10.1016/j.caeai.2025.100386.
5. Prova, Nuzhat, Explainable AI-Powered Multimodal Fusion Framework for EEG-Based Autism Spectrum Disorder Classification (January 28, 2025). Available at SSRN: https://ssrn.com/abstract=5114986 or http://dx.doi.org/10.2139/ssrn.5114986
6. Shah, Uzair, et al. "ArtInsight: A Multimodal AI Framework for Interpreting Children's Drawings and Enhancing Emotional Understanding." *Studies in Health Technology and Informatics*, edited by Elisavet Andrikopoulou et al., IOS Press, 2025. *DOI.org (Crossref)*, https://doi.org/10.3233/SHTI250471.
7. Guerrero-Sosa, Jared D. T., et al. "A Multimodal Framework for Explainable Evaluation of Soft Skills in Educational Environments." arXiv:2505.01794, arXiv, 3 May 2025. *arXiv.org*, https://doi.org/10.48550/arXiv.2505.01794.
8. Zhu, F. L., Wang, S. H., Liu, W. B., Zhu, H. L., Li, M., & Zou, X. B. (2023). A multimodal machine learning system in early screening for toddlers with autism spectrum disorders based on the response to name. *Frontiers in psychiatry*, *14*, 1039293. https://doi.org/10.3389/fpsyt.2023.1039293
9. Khosravi, Hassan, et al. "Explainable Artificial Intelligence in Education." *Computers and Education: Artificial Intelligence*, vol. 3, 2022, p. 100074. https://doi.org/10.1016/j.caeai.2022.100074.
10. Aylward, B. S., Abbas, H., Taraman, S., Salomon, C., Gal-Szabo, D., Kraft, C., Ehwerhemuepha, L., Chang, A., & Wall, D. P. (2023). An Introduction to Artificial Intelligence in Developmental and Behavioral Pediatrics. *Journal of developmental and behavioral pediatrics : JDBP*, *44*(2), e126–e134. https://doi.org/10.1097/DBP.0000000000001149
11. Agrawal, R., & Agrawal, R. (2025). Explainable AI in early autism detection: a literature review of interpretable machine learning approaches. *Discover mental health*, *5*(1), 98. https://doi.org/10.1007/s44192-025-00232-3
12. Khosravi, Hassan, et al. "Explainable Artificial Intelligence in Education." *Computers and Education:*

Artificial Intelligence, vol. 3, 2022, p. 100074. *DOI.org (Crossref)*, https://doi.org/10.1016/j.caeai.2022.100074

13. Sun, Shilin, et al. "A Review of Multimodal Explainable Artificial Intelligence: Past, Present and Future." arXiv, 2024. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.2412.14056.
14. Shaheen, Nour, et al. "Appraising Systematic Reviews: A Comprehensive Guide to Ensuring Validity and Reliability." *Frontiers in Research Metrics and Analytics*, vol. 8, Dec. 2023, p. 1268045. *DOI.org (Crossref)*, https://doi.org/10.3389/frma.2023.1268045.
15. Mancin, S., Sguanci, M., Andreoli, D., Soekeland, F., Anastasi, G., Piredda, M., & De Marinis, M. G. (2023). Systematic review of clinical practice guidelines and systematic reviews: A method for conducting comprehensive analysis. *MethodsX*, *12*, 102532. https://doi.org/10.1016/j.mex.2023.102532
16. *The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews | EQUATOR Network*. https://www.equator-network.org/reporting-guidelines/prisma/. Accessed 12 Sep. 2025.
17. *PROSPERO*. https://www.crd.york.ac.uk/prospero/. Accessed 12 Sep. 2025.
18. Rebecca Man, Victoria Hodgetts Morton, Jack Hamer, Rachel Katie Morris. A series of systematic reviews exploring childbirth related perineal trauma. PROSPERO 2024 Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD42023458738
19. "PRISMA 2020 Flow Diagram." *PRISMA Statement*, https://www.prisma-statement.org/prisma-2020-flow-diagram. Accessed 12 Sep. 2025.
20. Muhammad, Dost, and Malika Bendechache. "Unveiling the Black Box: A Systematic Review of Explainable Artificial Intelligence in Medical Image Analysis." *Computational and Structural Biotechnology Journal*, vol. 24, Dec. 2024, pp. 542–60. *DOI.org (Crossref)*, https://doi.org/10.1016/j.csbj.2024.08.005
21. Bae, S., Hong, J., Ha, S. *et al.* Multimodal AI for risk stratification in autism spectrum disorder: integrating voice and screening tools. *npj Digit. Med.* **8**, 538 (2025). https://doi.org/10.1038/s41746-025-01914-6
22. Cohn, Clayton, et al. "A Multimodal Approach to Support Teacher, Researcher and AI Collaboration in STEM +C Learning Environments." *British Journal of Educational Technology*, vol. 56, no. 2, Mar. 2025, pp. 595–620. *DOI.org (Crossref)*, https://doi.org/10.1111/bjet.13518.
23. Martinez-Maldonado, Roberto, et al. "Lessons Learnt from a Multimodal Learning Analytics Deployment In-the-Wild." *ACM Transactions on Computer-Human Interaction*, vol. 31, no. 1, Feb. 2024, pp. 1–41. *DOI.org (Crossref)*, https://doi.org/10.1145/3622784.
24. Ouhaichi, Hamza, et al. "Exploring Design Considerations for Multimodal Learning Analytics Systems: An Interview Study." *Frontiers in Education*, vol. 9, Jul. 2024, p. 1356537. *DOI.org (Crossref)*, https://doi.org/10.3389/feduc.2024.1356537
25. Ribeiro, Marco Tulio, et al. "'Why Should I Trust You?': Explaining the Predictions of Any Classifier." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* [San Francisco California USA], 2016, pp. 1135–44. *DOI.org (Crossref)*, https://doi.org/10.1145/2939672.2939778
26. Benson, F. N., Chelangat, D., Brink, W., Mwangala, P. N., Waljee, A. K., Moyer, C. A., & Abubakar, A. (2025). Application of machine learning in early childhood development research: a scoping review. *BMJ open*, *15*(8), e100358. https://doi.org/10.1136/bmjopen-2025-100358 (16)
27. Reinhart, Lisa, et al. "Artificial Intelligence in Child Development Monitoring: A Systematic Review on Usage, Outcomes and Acceptance." *Intelligence-Based Medicine*, vol. 9, 2024, p. 100134. *DOI.org (Crossref)*, https://doi.org/10.1016/j.ibmed.2024.100134.
28. Cui, Kaijie, et al. "Development of an Artificial Intelligence-Based Multimodal

Model for Assisting in the Diagnosis of Necrotizing Enterocolitis in Newborns: A Retrospective Study." *Frontiers in Pediatrics*, vol. 12, May 2024, p. 1388320. *DOI.org (Crossref)*, https://doi.org/10.3389/fped.2024.1388320.

29. Peter Foltz, et al. *Interactive Workshop: Multimodal, Multiparty Learning Analytics (MMLA)*. With Caitlin Mills et al., Jul. 2025. *DOI.org (Datacite)*, https://doi.org/10.5281/ZENODO.15870318.
30. Rodis, Nikolaos, et al. "Multimodal Explainable Artificial Intelligence: A Comprehensive Review of Methodological Advances and Future Research Directions." arXiv, 2023. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.2306.05731.
31. Cortese, Samuele, et al. "Latest Clinical Frontiers Related to Autism Diagnostic Strategies." *Cell Reports Medicine*, vol. 6, no. 2, Feb. 2025, p. 101916. *DOI.org (Crossref)*, https://doi.org/10.1016/j.xcrm.2024.101916.
32. Lundberg, Scott, and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." arXiv, 2017. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.1705.07874
33. Holzinger, Andreas, et al. "Causability and Explainability of Artificial Intelligence in Medicine." *WIREs Data Mining and Knowledge Discovery*, vol. 9, no. 4, Jul. 2019, p. e1312. *DOI.org (Crossref)*, https://doi.org/10.1002/widm.1312.
34. Ahmad Alaqsam and Corina Sas. 2025. Systematic Review of XAI Tools for AI-HCI Research. In Proceedings of the 37th International BCS Human-Computer Interaction Conference (BCS HCI '24). BCS Learning & Development Ltd, Swindon, GBR, 47–59. https://doi.org/10.14236/ewic/BCSHCI2024.6
35. Ribeiro, Marco Tulio, et al. "'Why Should I Trust You?': Explaining the Predictions of Any Classifier." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* [San Francisco California USA], 2016, pp. 1135–44. *DOI.org (Crossref)*, https://doi.org/10.1145/2939672.2939778
36. Lundberg, Scott, and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." arXiv, 2017. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.1705.07874
37. Holzinger, Andreas, et al. "Causability and Explainability of Artificial Intelligence in Medicine." *WIREs Data Mining and Knowledge Discovery*, vol. 9, no. 4, Jul. 2019, p. e1312. *DOI.org (Crossref)*, https://doi.org/10.1002/widm.1312.
38. Ahmad Alaqsam and Corina Sas. 2025. Systematic Review of XAI Tools for AI-HCI Research. In Proceedings of the 37th International BCS Human-Computer Interaction Conference (BCS HCI '24). BCS Learning & Development Ltd, Swindon, GBR, 47–59. https://doi.org/10.14236/ewic/BCSHCI2024.6
39. Das, Sougato, et al. "Adoption and Impact of AI-Enhanced Learning Platforms in Education." *Forum for Education Studies*, vol. 3, no. 1, Jan. 2025, p. 1696. *DOI.org (Crossref)*, https://doi.org/10.59400/fes1696.
40. Lundberg, Scott, and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." arXiv, 2017. *DOI.org (Datacite)*, https://doi.org/10.48550/ARXIV.1705.07874
41. Holzinger, Andreas, et al. "Causability and Explainability of Artificial Intelligence in Medicine." *WIREs Data Mining and Knowledge Discovery*, vol. 9, no. 4, Jul. 2019, p. e1312. *DOI.org (Crossref)*, https://doi.org/10.1002/widm.1312.