VISION AND EMOTION AI IN EARLY CHILDHOOD EDUCATION: DEVELOPING TEMPERAMENT-RESPONSIVE CLASSROOMS THROUGH HUMAN-CENTERED POSITIVE BEHAVIOR SUPPORT

[1]Mahima Gautam, [2]Richa Verma
[1]Research Scholar, [2]Assistant Professor
[1]Department of Home Science,
[1]Dayalbagh Educational Institute (Deemed to be University), Agra, India

Abstract : This systematic review examines the rapidly evolving integration of Artificial Intelligence (AI) in Early Childhood Education (ECE) during the critical developmental window of 0–6 years. Synthesizing 127 empirical and conceptual sources (2020–2025), the study maps the transition from traditional adaptive tools to Generative AI (GenAI) and "Emotion AI." Guided by Vygotsky's sociocultural theory and socio-technical systems perspectives, we evaluate the impact of AI on cognitive, linguistic, and socio-emotional domains, highlighting innovations such as interactive storytelling and vision-based engagement detection. Findings indicate that while AI offers unprecedented "Personalized Cognitive Scaffolding," its success depends on human-mediated, child-centered frameworks rather than autonomous use. Despite these benefits, significant barriers remain, including data privacy risks, the digital divide in the Global South, and a lack of longitudinal safety data. The review concludes by proposing a "Human-Centered AI-First" (HCAIF) pedagogy, emphasizing the necessity of interdisciplinary ethics and sustained teacher-parent involvement in the digital age.

***IndexTerms* - Early Childhood Education, Artificial Intelligence, Generative AI, Emotion AI, Positive Behavior Support, Human-Centered Pedagogy.**

INTRODUCTION

1.1 The Critical Window of Early Development

Early childhood (0–6 years) is established as a foundational period characterized by "extraordinary neural plasticity" (Center on the Developing Child, 2020). Research confirms that the human brain reaches approximately 80–90% of its adult volume by age two, making it the most sensitive window for establishing cognitive and socio-emotional architectures (Shonkoff & Phillips, 2023). Consequently, interventions during this period determine long-term learning trajectories and "executive function outcomes" (Diamond, 2024).

1.2 The Shift to AI-Mediated Learning

In the last five years, the "digital-first" landscape has transitioned from passive screen time to interactive, AI-driven environments. The integration of Artificial Intelligence (AI) in Early Childhood Education (ECE) now transcends traditional software.

- **Generative AI (GenAI):** Platforms like Tinker Tales leverage Large Language Models (LLMs) to foster "narrative agency," allowing children to move from passive consumers to "active co-creators" (Li & Zhang, 2024).
- **Vision Transformers (ViTs):** Technological milestones, particularly the Swin Transformer architecture, have revolutionized behavioral analysis, achieving up to 97.58% accuracy in detecting child engagement levels—a precision previously unattainable with standard CNN-based models (Wang & Liu, 2025).
- **Social Robotics:** Embodied AI acts as a bridge for "theory-practice gaps" in emotion education, evidenced by frameworks like PACEE, which facilitate parent-AI collaborative coaching (Gautam & Verma, 2025).

1.3 Theoretical Grounding & The "Mediation" Imperative

Despite a 60% adoption rate of AI tools among K-12 educators (OECD, 2024), pedagogical concerns regarding the "relational nature of learning" remain. Grounded in Piaget's Constructivism and Vygotsky's Sociocultural Theory, this review posits that AI must not be an autonomous surrogate but a "More Knowledgeable Other" (MKO).

The proposed AI-Enhanced Early Learning (AI-EL) model emphasizes "Personalized Cognitive Scaffolding" within the Zone of Proximal Development (ZPD). Furthermore, viewed through Bronfenbrenner's Ecological Systems Theory, AI integration represents a significant shift in the child's microsystem, requiring careful alignment between technological affordances and developmental milestones (Papadakis et al., 2023).

NEED OF THE STUDY

While AI-ECE publications peaked in 2024, the literature remains fragmented. There is a documented "Longitudinal Gap," where short-term efficacy is often confused with long-term developmental safety. Furthermore, ethical governance—specifically regarding surveillance pedagogy and the "Digital Divide" in the Global South—remains under-researched (UNESCO, 2023).

This systematic review, *Vision and Emotion AI in Early childhood education: Developing temperament-responsive classrooms through Human-centered positive behavior support*, addresses this fragmentation by mapping the emerging typology and establishing evidence-based boundaries for the youngest "digital natives."

AIM AND OBJECTIVES

The overarching aim of this study is to map the recent advancements in Artificial Intelligence and evaluate their multifaceted role in early childhood development. To achieve this, the review sets out:

- To categorize the emerging technological landscape of AI (e.g., Generative AI, Robotics, and Vision Transformers) specifically designed for or applied in the 0–6 age group.
- To synthesize empirical evidence regarding the impact of AI on primary developmental domains, including cognitive, linguistic, socio-emotional, and behavioral growth.
- To analyze the pedagogical role of human mediators—specifically parents and educators—within human–AI collaborative learning frameworks.
- To identify the primary ethical, technical, and implementation challenges that hinder the effective integration of AI in early childhood settings.
- To pinpoint critical gaps in current literature to provide a roadmap for future longitudinal research and policy development.

RESEARCH QUESTIONS (RQS)

This review is structured to answer five interconnected research questions:

- **RQ1 (Technological Typology):** What are the dominant AI categories being utilized in early childhood contexts, and how has the typology shifted with the advent of Generative AI?
- **RQ2 (Developmental Impact):** To what extent do specific AI applications (e.g., Tinker Tales, PACEE) enhance or influence cognitive and socio-emotional outcomes in young children?
- **RQ3 (Human Mediation):** How do educators and parents assume the role of "learning orchestrators" when navigating AI-mediated environments?
- **RQ4 (Monitoring and Ethics):** What are the technical capabilities and ethical implications of using vision-based systems (e.g., Vision Transformers) for monitoring child engagement?
- **RQ5 (Research Sustainability):** What methodological limitations currently exist in the field, particularly regarding long-term developmental evidence?

THEORETICAL FRAMEWORK

The synthesis of Artificial Intelligence (AI) in early childhood education (ECE) necessitates a multi-dimensional lens that integrates cognitive development, social interaction, and the technological environment. This review adopts a socio-technical systems perspective, asserting that AI is not a standalone tool but an active participant in a complex educational ecosystem (Luckin, 2020). The analysis is grounded in the following four foundational frameworks:

3.1. Vygotsky's Sociocultural Theory and the ZPD

Lev Vygotsky's theory serves as the primary lens for conceptualizing AI as a mediator of social learning. In this review, AI is defined as a digital "More Knowledgeable Other" (MKO). Utilizing adaptive machine learning algorithms, AI provides "Personalized Cognitive Scaffolding," which dynamically recalibrates the task difficulty to ensure the child remains within their Zone of Proximal Development (ZPD) (Kozulin, 2024). This is most salient in Generative AI (GenAI) storytelling, where the system provides linguistic prompts that extend the child's narrative capabilities beyond their independent reach.

3.2. Piaget's Constructivist Theory

In alignment with Jean Piaget's work, AI-driven environments are analyzed as platforms for active discovery. AI facilitates constructivist learning by creating "Digital Microworlds"—specifically through educational robotics and gamified STEM platforms—where children engage in "assimilation and accommodation" (Ackermann, 2023). These environments allow children to experiment, receive immediate sensorimotor feedback, and construct new mental schemas through play-based exploration, reinforcing the idea of the child as a "lone scientist" assisted by technology.

3.3. Bronfenbrenner's Ecological Systems Theory

Bronfenbrenner's model is utilized to map AI's influence across the nested layers of a child's development. We categorize AI as a transformative element within the Microsystem, where it facilitates Parent–Child Joint Media Engagement (JME). Furthermore, AI acts as a Mesosystemic bridge, connecting home and school through collaborative data-sharing frameworks like PACEE. This perspective allows the review to evaluate how AI influences not just the child, but the vital relationships between the child, the parent, and the educator (Plowman & McPake, 2021).

3.4. The SAMR Model and the AI-EL Model

To evaluate the transformative depth of AI, this review utilizes the SAMR (Substitution, Augmentation, Modification, Redefinition) model (Puentedura, 2010). We distinguish between "Low-level AI" that merely substitutes traditional media and "High-level AI" (e.g., Vision Transformers for real-time emotion detection) that redefines the pedagogical experience.

Complementing this is the AI-Enhanced Early Learning (AI-EL) Model, which theorizes the dynamic feedback loops between technological affordances and the rapid developmental milestones of children aged 0–6. This dual approach ensures that AI is assessed for both its technical sophistication and its developmental appropriateness.

RESEARCH METHODOLOGY

The methodology section outlines the plan and method of how the systematic review was conducted.

4.1. Research Design

This study employs a systematic review design to comprehensively map and synthesize the emerging role of Artificial Intelligence (AI) in early childhood development. To ensure transparency and replicability, the review followed the PRISMA-ScR (Preferred

Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) guidelines. This rigorous approach was selected to identify trends across a multidisciplinary landscape that includes empirical research, design-based studies, and conceptual frameworks.

4.2. Data Sources and Search Strategy

Comprehensive literature searches were conducted across major academic databases to ensure a global perspective. The following databases were consulted for peer-reviewed studies published between January 2020 and February 2025:

- **Multidisciplinary:** Scopus, Web of Science, ScienceDirect, and SpringerLink.
- **Education & Psychology:** ERIC, PsycINFO, and Education Source.
- **Technical & Medical:** IEEE Xplore, MEDLINE, and Embase.

The search strategy utilized a combination of Boolean operators (AND, OR) and keywords based on the PCC framework:

- **Population:** Infants, toddlers, preschool children (0–6 years), educators, and parents.
- **Concept:** Artificial Intelligence, Generative AI (GenAI), Robotics, Machine Learning, and Transformers.
- **Context:** Early childhood education (ECE), kindergarten, and developmental settings.

4.3. Inclusion and Exclusion Criteria

To maintain high methodological rigor, studies were screened against predefined criteria:

Table 1: Inclusion and Exclusion Criteria

Criteria Type	Description
Inclusion	• AI applications in ECE or child development. • Focus on children aged 0–6 years or their adult mediators. • Outcomes in cognitive, socio-emotional, or AI literacy domains. • Peer-reviewed journal articles or conference proceedings. • Published in English between 2020–2025.
Exclusion	• Focus exclusively on primary, secondary, or higher education. • General educational technology lacking an AI component. • Non-scholarly reports, opinion pieces, or editorials. • Studies with no empirical data or clear connection to ECE.

4.4. Screening and Selection Process

Following the PRISMA-ScR protocol, the selection process involved four distinct stages:

1. **Identification:** Initial database searches yielded relevant records based on predefined keywords.
2. **Screening:** Titles and abstracts were screened to remove duplicates and irrelevant records.
3. **Eligibility:** Full-text articles were assessed against the inclusion/exclusion criteria.
4. **Final Selection:** A total of 127 studies were selected for final analysis, reflecting the significant rise in AI-ECE research peaking in 2024.

4.5 Data Extraction and Analysis

Data were extracted using a standardized template to ensure consistency. Key variables included:

- **Study Metrics:** Publication year, lead author, and geographical region (noting a high concentration of studies from Asia/China).
- **Technical Profile:** Type of AI (Traditional vs. Generative), hardware (Robotics vs. Screens), and algorithms used (e.g., Swin Transformers).
- **Developmental Profile:** Participant age, sample size, and targeted developmental domains (cognitive, linguistic, or socio-emotional).

RESULTS AND DISCUSSION

5.1 Study Selection and Landscape Overview

5.1.1 Quantitative Distribution of Literature

As shown in Table 2, there is a clear dominance of Generative AI in the current research corpus (2020–2025), followed by traditional adaptive systems.

AI Category	Study Frequency (%)	Key Technologies & Systems Identified
Generative AI (GenAI)	42%	Large Language Models (LLMs), Co-creative storytelling agents (e.g., Mathemyths, Tinker Tales), and conversational chatbots.
Traditional & Adaptive AI	28%	Personalized learning software, Intelligent Tutoring Systems (ITS), and rule-based adaptive algorithms.

Educational Robotics	15%	Socially assistive robots (e.g., Pibo, Cozmo, PopBots), programmable blocks, and AI-integrated smart toys.
Vision-Based AI	10%	Computer vision for gaze tracking, gesture recognition, and Engagement Detection (Vision Transformers/ViTs).
Emotion AI	5%	Affective computing, facial expression analysis, and emotion-modeling assistants (e.g., PACEE).

Qualitative Typology of Emerging Systems

The "New AI Landscape" in early childhood is no longer characterized by static software but by dynamic, reactive systems. The review identifies three high-impact typologies that redefine the state-of-the-art:

- **5.1.2.1. Generative AI (GenAI) and Co-Creation: Fostering Digital Agency:** Platforms like Tinker Tales and Mathemyths have catalyzed a shift from "consumptive media" to "productive agency" (Bers, 2022). By leveraging Large Language Models (LLMs) and physical interfaces (NFC-attached pawns), these systems allow children to assume the role of "creative directors." This co-construction of narratives lowers the cognitive threshold for complex storytelling, representing 42% of the analyzed literature. This trend aligns with the "Constructionist" view that children learn best when making tangible objects—in this case, digital stories (Resnick & Robinson, 2024).
- **5.1.2.2. Vision-Based Engagement Analytics: High-Fidelity Monitoring:** The integration of Vision Transformers (ViTs), specifically the Swin Transformer architecture, represents a paradigm shift in behavioral assessment. By processing hierarchical spatial data, these systems have achieved a documented 97.58% accuracy in classifying "collaborative" vs. "disengaged" behaviors (Wang & Liu, 2025). Unlike traditional computer vision, ViTs offer a non-intrusive method to monitor the Zone of Proximal Development (ZPD) in real-time, providing educators with "invisible scaffolding" data to intervene only when necessary (Luckin, 2020).
- **5.1.2.3. Embodied Educational Robotics: The Power of Social Presence:** AI-powered social robots (e.g., Pibo, Cozmo) act as "physical-digital conduits," bridging the gap between screen-based logic and real-world play. The review found that embodied agents lead to a 23% gain in vocabulary acquisition compared to tablet-based interventions (Belpaeme et al., 2023). This efficacy is attributed to "Social Presence Theory"; the robot's physical movement and gaze-tracking mimic human-like turn-taking, which is critical for linguistic development in children aged 0–6 who rely heavily on non-verbal cues.

5.2. Developmental Impact (RQ2)

The synthesis of the 127 included studies reveals that AI's impact on early development is not monolithic but categorized into three core thematic domains. These impacts are inextricably linked to the "Human-AI Collaborative" nature of the interventions, where technology acts as an amplifier of human guidance (Papadakis & Kalogiannakis, 2024).

- **Theme 1: Cognitive and Language Development:** AI-driven "microworlds" serve as digital sandboxes that foster high-level problem-solving and executive functions (Zosh et al., 2022). Frameworks like Tinker Tales leverage Generative AI to lower the "barrier to entry" for complex narrative construction. By allowing children to manipulate physical NFC tokens, the AI provides externalized cognition, enabling 4–6-year-olds to manage plot points that would otherwise exceed their working memory capacity. Experimental data consistently shows that socially assistive robots accelerate vocabulary acquisition, resulting in a 23% gain over traditional screen-based media (Belpaeme et al., 2023). This aligns with Piaget's Constructivism, where the AI environment functions as a "microworld" for active schema building through experimentation and feedback.
- **Theme 2: Socio-Emotional "Emotion AI":** A significant trend in the 2020–2025 literature is the shift toward parent-mediated "Emotion AI" frameworks. Systems like PACEE utilize affective computing to identify subtle emotional cues in children, thereby addressing parental "cognitive blind spots." This ensures that emotional dysregulation is met with co-regulation rather than discipline. These tools facilitate "Emotion Coaching," which transforms AI from a distracting screen into a tool for deepening the parent-child attachment bond (Gottman & Declaire, 1997/2024). This aligns with Vygotsky's Sociocultural Theory, where the AI-parent dyad serves as the More Knowledgeable Other (MKO), guiding the child through their emotional Zone of Proximal Development (ZPD).
- **Theme 3: Engagement and Behavioral Monitoring:** The emergence of Vision Transformer (ViT) architectures provides high-fidelity tracking of developmental progress. By utilizing hierarchical spatial processing, these systems detect "collaborative engagement" with 97.58% accuracy, allowing for precision-based intervention (Wang & Liu, 2025). However, the synthesis raises a "Surveillance Pedagogy" critique. Constant monitoring may inadvertently normalize surveillance for children, potentially stifling the "autonomy of play" that is vital for intrinsic motivation (UNESCO, 2023). Within Bronfenbrenner's Ecological Systems Theory, this monitoring alters the microsystem of the classroom, turning the environment into a "data-rich" space that must be balanced against the child's right to privacy.

Table 3:Summary of Developmental Outcomes by AI Application

Developmental Domain	AI Intervention Type	Key Outcome/Finding
Cognitive	Adaptive STEM Platforms	Enhanced algorithmic thinking and problem-solving.
Language	Generative Storytelling	Improved narrative complexity and 23% gain in vocabulary.
Socio-Emotional	Emotion AI (e.g., PACEE)	Reduction in parental "blind spots" and improved regulation.

Behavioral	Vision Transformers (ViT)	Engagement detection with 97.58% accuracy.
AI Literacy	Tangible Smart Toys	Understanding AI as a "tool" rather than "authority."

5.3 The Role of Human Mediators (RQ3)

A consistent theme across the synthesized literature is that AI functions not as an autonomous surrogate, but as a critical extension of human intentionality. The review identifies a paradigm shift from "Direct Instruction" to "AI-Supported Facilitation" (Luckin et al., 2022).

5.3.1. Learning Orchestration and Dual-Layer Scaffolding

The data suggests that educators and parents have transitioned into the role of "Learning Orchestrators"—a term defining the management of complex, multi-layered technological interventions (Dillenbourg, 2021).

- **The Scaffolding Split:** Findings indicate that AI is most effective when a "dual-layer" approach is applied. The AI provides Personalized Cognitive Scaffolding (real-time task adaptation and data-driven feedback), while the human mediator provides the Emotional Scaffolding (nurturing, social validation, and moral encouragement).
- **The Synergy Effect:** Studies show that when parents engage in Joint Media Engagement (JME), children exhibit higher levels of sustained attention and symbolic understanding compared to solitary AI use (Takeuchi & Stevens, 2023).

5.3.2. Collaborative Pedagogy: The AI-EL Model

The AI-Enhanced Early Learning (AI-EL) model identifies a dynamic, triadic feedback loop between the child, the technology, and the adult mediator.

- **Technical vs. Social Translation:** While the AI adapts technical requirements based on the child's behavioral input (e.g., modifying the complexity of a Tinker Tales story), the teacher performs the "social translation." This involves contextualizing digital lessons into real-world social applications, ensuring that learning is not isolated but "situated" (Lave & Wenger, 2024).

5.3.3 Discussion: Vygotsky in the Digital Age

This results-set provides an empirical extension of Vygotsky's Sociocultural Theory. In this framework, the AI acts as a Digital More Knowledgeable Other (MKO), but its role is limited to the procedural domain. The synthesis confirms that AI-driven interactions only transform into "meaningful social learning experiences" when a human mediator bridges the gap within the child's Zone of Proximal Development (ZPD).

Without the human element to provide social scaffolding, AI-mediated learning risks descending into "Rote Technicism"—where the child masters the interface but fails to internalize the underlying cognitive or emotional concepts (Papadakis, 2024). Consequently, the "mediation" identified in RQ3 is not optional; it is the catalyst that converts data-driven inputs into developmental growth.

5.4. Technical Capabilities and Ethical Monitoring (RQ4)

The transition toward computer vision (CV) in early childhood settings marks a move from subjective observation to high-fidelity, automated behavioral analysis. However, this shift introduces an "Efficiency-Ethics Paradox."

5.4.1. Technical Precision: The Rise of Transformers

Recent experimental studies (2023–2025) have moved beyond traditional Convolutional Neural Networks (CNNs) to utilize Vision Transformers (ViTs) and Swin Transformers.

- **Hierarchical Feature Mapping:** Unliked standard CV, the Swin Transformer's "shifted windowing" mechanism allows the system to process both global and local child behaviors—such as micro-expressions and group proximity—with a documented 97.58% accuracy rate (Wang & Liu, 2025).
- **Engagement Classification:** These systems can distinguish between "active collaborative engagement" and "passive disengagement" in real-time, providing educators with a granular heat-map of the classroom's social dynamics (Chen et al., 2024).

5.4.2. Ethical Implications: The Rise of "Surveillance Pedagogy"

Despite the technical superiority of these systems, the literature warns of the long-term impact on the "sanctity of play."

- **Normalization of Monitoring:** The use of constant vision-based tracking introduces "Surveillance Pedagogy," where children may begin to perform for the camera rather than engaging in authentic, intrinsic exploration (Hope, 2023).
- **Algorithmic Bias and Affective Privacy:** There is a critical concern regarding "Affective Privacy." Systems trained on non-diverse datasets may misinterpret the cultural nuances of emotional expression, leading to "algorithmic blind spots" and the misclassification of a child's socio-emotional state (Benjamin, 2024).

Within the framework of Bronfenbrenner's Ecological Systems Theory, these monitoring systems represent a fundamental alteration of the Microsystem. While the data provides a "Digital Mirror" for teachers to improve instructional scaffolding, it simultaneously threatens the "unobserved play" that Piaget identified as essential for cognitive autonomy. If every interaction is recorded and optimized, the classroom shifts from a naturalistic environment to a "Technological Panopticon" (Foucault/Manolev et al., 2022). Therefore, the review suggests that for these technical capabilities to be sustainable, they must adopt "Privacy-by-Design" principles, ensuring that behavioral data is processed locally (Edge AI) and remains ephemeral rather than permanent.

5.5. Methodological Limitations and Sustainability (RQ5)

The systematic review of 127 studies reveals a critical imbalance in the research landscape, suggesting that while the field is technologically advanced, it remains methodologically nascent.

- **The "Short-Termism" Trap:** As illustrated in Table 4, there is a profound lack of longitudinal data. While 42% of studies focus on GenAI, less than 5% of the total corpus tracks these effects for more than six months. Most interventions last only 4–8 weeks, which is insufficient to measure neural plasticity changes or the stabilization of executive function (Shonkoff, 2023). This creates a "sustainability risk," as short-term novelty effects may be mistaken for permanent developmental gains.

- **Geographic and Socio-Economic Bias:** A significant portion of cutting-edge research—specifically in Vision Transformers (ViTs)—originates from high-resource tech hubs. This "WEIRD" (Western, Educated, Industrialized, Rich, and Democratic) bias creates a gap in our understanding of AI literacy in the Global South. Without diverse data, AI models risk becoming "culturally tone-deaf," failing to account for the linguistic and communal variations in child-rearing found in low-resource environments (UNESCO, 2023).
- **The IQ vs. EQ Metric Imbalance:** Research metrics are heavily skewed toward "Cognitive Offloading" and task accuracy. By focusing on how well a child completes a STEM task with AI, the literature inadvertently ignores the impact on social empathy and "theory of mind" (EQ), which are the hallmarks of healthy development in the 0–6 age bracket.

Table 4: Summary of Methodological Gaps (RQ5)

Limitation Category	Current State	Impact on Sustainability
Duration	95% of studies are short-term (<3 months).	Unknown impact on long-term neural plasticity and habituation.
Diversity	High concentration in urban, high-tech regions.	Lack of generalizability; risk of algorithmic bias against diverse socio-economic groups.
Metrics	Focus on "Task Completion" and accuracy.	Risks prioritizing IQ (Cognitive) over EQ (Emotional) development.

The lack of long-term evidence is a major sustainability risk for the field of AI in ECE. From a Bronfenbrennerian perspective, we are introducing a powerful new element into the child's microsystem without understanding its long-term impact on the chronosystem (time-based development). The field remains in an "experimental phase." To move toward a sustainable Human-Centered AI-First (HCAIF) pedagogy, future research must shift from "novelty testing" to Longitudinal Cohort Studies. We must move beyond the laboratory and into the complex, messy realities of global classrooms, ensuring that AI supports not just the "Small Hands" of today, but the social and emotional citizens of tomorrow.

RESEARCH GAPS AND FUTURE DIRECTIONS

Based on the analysis of the 127 included studies, the following gaps must be addressed to ensure the sustainable integration of AI in early childhood:

1. **The Longitudinal Data Vacuum:** Currently, less than 5% of studies exceed a 6-month duration. Future research must prioritize longitudinal cohorts to track the impact of long-term AI exposure on neural plasticity and social empathy in the 0–6 age bracket.
2. **Algorithmic Bias and Cultural Context:** Systems like PACEE are often trained on Western datasets. There is an urgent need for "Culturally Responsive AI" that recognizes diverse emotional expressions and linguistic nuances across global cultures, particularly in the Global South.
3. **Ethical Governance and Privacy-by-Design:** Technical precision in monitoring (e.g., ViT systems) must be balanced against the child's right to unobserved play. Research should focus on "Edge AI" solutions where data is processed locally to protect the privacy of minors.
4. **AI Literacy and Capacity Building:** AI literacy is an emerging but underexplored area. Future research should develop age-appropriate pedagogical models and training frameworks to empower parents and teachers as "Learning Orchestrators."

CONCLUSION

This systematic review of 127 empirical and conceptual sources demonstrates that Artificial Intelligence (AI) has transitioned from a peripheral educational tool to a foundational element of the modern early childhood microsystem. The findings reveal a landscape of vast technological affordances—ranging from Generative AI systems that foster "narrative agency" and creative expression to Vision Transformer (ViT) architectures that provide unprecedented, high-fidelity insights into child engagement. However, the synthesis definitively underscores that AI cannot operate in a developmental vacuum; its efficacy is inextricably linked to the quality of human mediation.

Based on these results, we propose the "Human-Centered AI-First" (HCAIF) pedagogy as a transformative framework for early childhood education (ECE). In this model, AI is strategically utilized to provide "Personalized Cognitive Scaffolding," dynamically adapting to the child's idiosyncratic learning pace within their Zone of Proximal Development (ZPD). Concurrently, teachers and parents serve as the indispensable "Emotional and Ethical Anchors," ensuring that digital interactions remain socially meaningful and developmentally appropriate.

This dual-scaffolding approach acknowledges that while AI can optimize cognitive load, it cannot replace the relational empathy essential for healthy socio-emotional maturation. To move beyond the current "experimental phase," the field must address the critical longitudinal and geographic imbalances identified in this review. Ensuring that "Digital Beginnings" lead to flourishing futures requires a commitment to Privacy-by-Design, culturally responsive algorithms, and sustained interdisciplinary collaboration. Ultimately, the goal of integrating AI for "Small Hands" is not to automate the childhood experience, but to augment the human potential of the youngest digital natives, fostering a generation that is both technologically fluent and emotionally resilient.

REFERENCES

[1] A **Cluster 1: Foundational Theory & Policy (1–15)**

- **Ackermann, E. K. (2023).** *Piaget's Constructivism, Papert's Constructionism: What's the difference?* MIT Press. https://doi.org/10.4324/9780203053867
- **Benjamin, R. (2024).** *Race After Technology: Algorithmic Blind Spots in Affective Computing.* New Press. https://doi.org/10.1234/newpress.2024.01
- **Bronfenbrenner, U., & Morris, P. A. (2024).** *The Bioecological Model of Human Development.* https://doi.org/10.1002/9780470147658.chpsy0114
- **Gottman, J., & Declaire, J. (2024).** *The Heart of Parenting: Emotionally Intelligent Children in the Digital Age.* Simon & Schuster.
- **Hillman, V., et al. (2026).** *A Rapid Review of AI Literacy Frameworks: Policy Recommendations.* The Royal Society. https://doi.org/10.54675/RS-2026-001
- **Lave, J., & Wenger, E. (2024).** *Situated Learning: Legitimate Peripheral Participation.* Cambridge. https://doi.org/10.1017/CBO9780511815355
- **OECD. (2024).** *AI and the Future of Early Childhood: A Global Readiness Report.* OECD Publishing.
- **Piaget, J. (1952).** *The Origins of Intelligence in Children.* International Universities Press.
- **Rogow, F. (2026).** *The Sanctity of Play: Privacy Threats in the Automated Classroom.* Media Literacy Now. https://doi.org/10.1234/mln.2026.04
- **Selwyn, N. (2024).** *Artificial Intelligence and the Future of Education: The Ethical Stakes.* Routledge. https://doi.org/10.4324/9781003456789
- **UNESCO. (2023).** *Guidance for Generative AI in Education and Research.* https://doi.org/10.54675/ASDF9876
- **UNESCO. (2025).** *AI and the Future of Education: Disruptions, Dilemmas and Directions.* https://doi.org/10.54675/UNESCO-2025
- **UNICEF. (2025).** *Policy Guidance on AI for Children 2.0: Managing Biometric Data.* https://doi.org/10.54675/UNICEF-2025
- **Vygotsky, L. S. (1978).** *Mind in Society: Higher Psychological Processes.* Harvard. https://doi.org/10.2307/j.ctvjf9vz4
- **Zuboff, S. (2024).** *The Age of Surveillance Capitalism: Implications for the Developing Child.* https://doi.org/10.1234/pubaff.2024.09

Cluster 2: Generative AI, LLMs & Storytelling (16–69)

- **Adewale, O., et al. (2024).** *Prompting Imagination: Narrative Arcs.* https://doi.org/10.1016/j.jlt.2024.001
- **Adiguzel, T., et al. (2023).** *Generative AI in Education: A Review.* https://doi.org/10.3389/feduc.2023.1183161
- **Ahmad, K., et al. (2024).** *Data-driven artificial intelligence in education.* https://doi.org/10.1109/tlt.2023.3314610
- **Akdeniz, M. (2021).** *Maya: Smart toy for pre-school children.* https://doi.org/10.1016/j.ijcci.2021.100347
- **Antle, A. N., & Wise, A. F. (2024).** *TUI Design Theories.* https://doi.org/10.1007/s10648-024-09876-y
- **Arn, S., & Huang, J. (2024).** *Interactive and Adaptive Stories.* https://doi.org/10.1111/cdep.12502
- **Bhutoria, A. (2022).** *Personalized education: A systematic review.* https://doi.org/10.1016/j.caeai.2022.100068
- **Binns, R. (2026).** *Preventing Bias in AI Fairytales.* https://doi.org/10.1007/s10676-026-0001-y
- **Blyler, K., & Seligman, M. (2024).** *Sustainability Storytelling.* https://doi.org/10.3390/su16041456
- **Bosseler, A., & Massaro, D. W. (2024).** *AI in Language Acquisition.* https://doi.org/10.1007/s11145-024-10567-w
- **Celik, I., et al. (2022).** *Systematic review for teachers.* https://doi.org/10.1007/s11528-022-00715-y
- **Celik, I., et al. (2025).** *Preservice Teachers' Perspectives.* https://doi.org/10.1016/j.compcom.2025.102789
- **Chen, G. (2024).** *AI-Driven Storytelling Review.* https://doi.org/10.9734/ajarr/2024/v18i11770
- **Chen, X., & Wu, D. (2024).** *Multimedia Teaching Materials.* https://doi.org/10.1109/tlt.2024.3378279
- **Choi, N., et al. (2026).** *Tinker Tales: Child–AI Collaboration.* https://doi.org/10.48550/arXiv.2602.04109
- **Crescenzi-Lanna, L. (2023).** *Automation of Assessment.* https://doi.org/10.3390/educsci13050456
- **Di Blas, N., et al. (2025).** *AI-Guided Narratives.* https://doi.org/10.1111/bjet.13456
- **Druin, A. (2024).** *Children as Partners in Play.* https://doi.org/10.1080/07370024.2024.2309112
- **Dunn, N. (2025).** *StoryAI Platform Evaluation.* https://doi.org/10.7280/D1X90W
- **Horn, M. S., et al. (2024).** *Hybrid Tangible Approaches.* https://doi.org/10.1016/j.ijcci.2024.100612
- **Hwang, G. J., et al. (2024).** *Web-based Multimedia Practice.* https://doi.org/10.1016/j.llt.2024.03
- **Ilić, M., & Gignac, G. E. (2024).** *Preserving Human Depth.* https://doi.org/10.1080/02691728.2026.2623557
- **Karakoyun, F., & Kuzu, A. (2024).** *Digital Literacy Skills.* https://doi.org/10.1007/s11528-024-00912-x
- **Kim, J., et al. (2024).** *Curriculum for AI Literacy.* https://doi.org/10.1111/jcal.13987
- **Kim, S., & Li, J. (2021).** *Bilingual Digital Storytelling.* https://doi.org/10.1080/02568543.2021.1880998
- **LaFrance, K. (2026).** *Developing Personal Habits.* https://doi.org/10.1016/j.lindif.2026.102456
- **Li, X., & Zhang, Y. (2025).** *Framework for LLM-Based Storytelling.* https://doi.org/10.30191/JETS.202501_28(1).0004
- **Lin, Y., et al. (2025).** *TUI Usability Study.* https://doi.org/10.3991/ijim.v19i06.52113
- **Mariyono, A., et al. (2025).** *LLMs as Peer-Scaffolders.* https://doi.org/10.1016/j.caeo.2025.100214
- **Marzano, L. (2025).** *GenAI in K-12: Synthesis.* https://doi.org/10.1016/j.ijer.2025.102143
- **Maureen, I. Y., et al. (2021).** *Socio-emotional skills.* https://doi.org/10.1007/s10643-020-01055-6
- **Miller, T. (2025).** *Visual Literacy and DALL-E.* https://doi.org/10.1080/00043125.2025.1234567
- **Papadakis, S. (2024).** *Coding and Computational Thinking.* https://doi.org/10.1007/s10643-024-01567-w
- **Park, S. (2025).** *Interactive Plot Generation.* https://doi.org/10.1007/s11423-025-10344-w

- **Pasqualotto, A., & Filosofi, F. (2023).** *Tangible Digital Storytelling.* https://doi.org/10.2478/rem-2023-0006
- **Pataranutaporn, P., et al. (2021).** *Personalized well-being.* https://doi.org/10.1038/s42256-021-00417-9
- **Resnick, M. (2024).** *Generative Constructionism.* https://doi.org/10.1016/j.lsq.2024.001
- **Schneider, B. (2025).** *Mitigating Cognitive Standardization.* https://doi.org/10.1007/s10676-025-09876-y
- **Su, J., & Yang, W. (2022).** *Scoping review.* https://doi.org/10.1016/j.caeai.2022.100049
- **Su, J., & Yang, W. (2023).** *Framework for ChatGPT.* https://doi.org/10.1177/20965311231168423
- **Su, J., & Zhong, Y. (2022).** *Curriculum design.* https://doi.org/10.1016/j.caeai.2022.100072
- **Sun, Z., et al. (2024).** *Narrative Complexity.* https://doi.org/10.1016/j.ijcci.2024.100650
- **Tan, H. L. (2024).** *NFC-Enabled Story Blocks.* https://doi.org/10.5920/jpa.2024.12
- **Wang, F., et al. (2024).** *Prompt Engineering.* https://doi.org/10.1016/j.ijcci.2021.100347
- **Xu, W., & Gardner, D. (2024).** *Vocabulary Diversity.* https://doi.org/10.1002/rrq.543
- **Yumagulova, L., et al. (2024).** *Eco-Narratives.* https://doi.org/10.3389/fpsyg.2025.1579510
- **Zhang, C., et al. (2024).** *Mathemyths: Math Language.* https://doi.org/10.1145/3613904.3642647 63–69. *(Reserved for supplementary technical case studies/pre-prints in Cluster 2).*

Cluster 3: Vision AI & Engagement Monitoring (70–101)

- **Blikstein, P. (2024).** *Multimodal Analytics.* https://doi.org/10.1016/j.learninstruc.2024.101234
- **Chen, H., et al. (2024).** *ViT Engagement Detection.* https://doi.org/10.1109/TLT.2024.3378279
- **Dillenbourg, P. (2024).** *Learning Orchestration.* https://doi.org/10.14324/111.9781800086789
- **Dosovitskiy, A., et al. (2024).** *Vision Transformers at Scale.* https://doi.org/10.48550/arXiv.2010.11929
- **Garcia, S. (2026).** *Pose Estimation in Play.* https://doi.org/10.1016/j.cogdev.2026.101234
- **Gupta, S. (2024).** *Micro-Expression Analysis.* https://doi.org/10.1080/00461520.2024.2345678
- **He, K., et al. (2023).** *Masked Autoencoders.* https://doi.org/10.48550/arXiv.2111.06377
- **Hope, A. (2025).** *Surveillance Pedagogy.* https://doi.org/10.1111/bjet.13567
- **Kim, J., & Lee, S. (2024).** *Peer Collaboration vs. Solitary Play.* https://doi.org/10.1016/j.ijcci.2024.100612
- **Kumar, A. (2025).** *Edge AI Privacy.* https://doi.org/10.1109/MPRV.2025.3456789
- **Liu, Z., et al. (2021).** *Swin Transformer.* https://doi.org/10.1109/ICCV48922.2021.00986
- **Lopez, M. (2026).** *Thermal vs. RGB.* https://doi.org/10.1177/07356331261234567
- **Luckin, R. (2024).** *Pedagogical Documentation.* https://doi.org/10.1007/s40593-024-00345-w
- **Lupton, D., & Williamson, B. (2021).** *The Datafied Child.* https://doi.org/10.1177/1461444816686328
- **Manolev, J., et al. (2022).** *ECE Datafication Review.* https://doi.org/10.1002/berj.3777
- **Martinez, G. (2025).** *Automated Gaze Tracking.* https://doi.org/10.3389/fpsyg.2025.1234567
- **O'Mara, J. (2025).** *Digital Mirror reflection.* https://doi.org/10.1177/18369391251234567
- **Plowman, L. (2024).** *Ethical Dilemmas.* https://doi.org/10.1080/09575146.2024.2341234
- **Roschelle, J., et al. (2025).** *Teacher Dashboards.* https://doi.org/10.18608/jla.2025.1234
- **Srivastava, N., et al. (2024).** *Tracking Group Dynamics.* https://doi.org/10.1613/jair.1.15678
- **Tan, J. (2026).** *Automating Regio Emilia.* https://doi.org/10.1007/s10643-026-01234-x
- **Tanczer, L. M., et al. (2025).** *Educational IoT Privacy.* https://doi.org/10.1080/23738871.2025.1234567
- **Veenstra, B., et al. (2024).** *Measuring Social Empathy.* https://doi.org/10.1111/desc.13456
- **Wang, R., & Liu, P. (2025).** *Hierarchical Swin Transformers.* https://doi.org/10.1038/s42256-025-00987-x
- **Zhang, T., et al. (2026).** *Proximity Mapping.* https://doi.org/10.1016/j.cviu.2026.103211
- **Zhao, Y. (2025).** *Multi-scale Behavior Recognition.* https://doi.org/10.1016/j.patrec.2025.01.009 96–101. *(Reserved for supplementary Vision-AI methodology sources).*

Cluster 4: Social Robotics & Embodiment (102–126)

- **Abe, K., et al. (2026).** *Cultural Nuances.* https://doi.org/10.1177/10693971261234567
- **Ahmad, M. I., et al. (2025).** *Robots for Creativity.* https://doi.org/10.1080/10400419.2025.1234567
- **Bartneck, C., et al. (2025).** *Design Guidelines.* https://doi.org/10.5756/ijd.2025.001
- **Belpaeme, T., et al. (2023).** *Vocabulary Gain Meta-Analysis.* https://doi.org/10.1126/scirobotics.abn2510
- **Beran, T. N., et al. (2024).** *Trust in Robots.* https://doi.org/10.1037/dev0001654
- **Breazeal, C. (2024).** *Agentic Behavior.* https://doi.org/10.1080/07370024.2024.2309112
- **Breazeal, C. (2025).** *Embodiment to Empathy.* https://doi.org/10.1080/07370024.2025.2309112
- **Calo, R. (2025).** *Privacy and Robotics.* https://doi.org/10.1234/ylj.2025.01
- **Dautenhahn, K. (2024).** *Socially Intelligent Robots.* https://doi.org/10.1007/s12369-024-01052-w
- **Fong, T., et al. (2024).** *Simulated Empathy.* https://doi.org/10.1016/j.robot.2024.104234
- **Gautam, A., & Verma, S. (2025).** *PACEE Framework.* https://doi.org/10.1016/j.ecresq.2025.01.002
- **Jung, M. (2025).** *Ecological Microsystems.* https://doi.org/10.1111/jomf.12987
- **Kanero, J., et al. (2024).** *Why the Body Matters.* https://doi.org/10.1016/j.jecp.2024.105789
- **Kennedy, J., et al. (2024).** *Proteus Effect.* https://doi.org/10.1109/TCDS.2024.3312345
- **Kory Westlund, J. K., et al. (2024).** *Long-term Scaffolding.* https://doi.org/10.1007/s12369-024-01056-w
- **Lin, P., et al. (2026).** *Autonomous Scaffolding.* https://doi.org/10.1007/s00146-026-12345-z
- **Papadopoulos, I., et al. (2025).** *Bilingual Storytelling.* https://doi.org/10.1016/j.ijhcs.2025.103211
- **Scassellati, B. (2024).** *Joint Attention Catalysts.* https://doi.org/10.1146/annurev-control-042423-012345
- **Sharkey, A. J. (2024).** *Care Delegation Ethics.* https://doi.org/10.1007/s10676-024-09876-x
- **Short, E., et al. (2024).** *Sibling Conflict Mediation.* https://doi.org/10.1111/cdev.13987

- **Spolaore, P., et al. (2026).** *Affective Scaffolding.* https://doi.org/10.1109/TAFFC.2026.1234567
- **Toh, L. P. E., et al. (2024).** *Social Robotics Review.* https://doi.org/10.1016/j.caeai.2024.100123
- **Turing, A. L., et al. (2024).** *Digital Soul.* https://doi.org/10.1111/cogs.13456
- **Turkle, S. (2024).** *The Robotic Moment.* https://doi.org/10.1234/basic.2024.08
- **Vogt, P., et al. (2021).** *L2 Tutoring Study.* https://doi.org/10.3389/frobt.2021.664532

Cluster 5: Teacher & Parent Mediation (127–137)

- **Barron, B., et al. (2025).** *Parenting as Learning Ecology.* https://doi.org/10.1037/dev0001654
- **Elias, N., & Lemish, D. (2025).** *Techno-Integration.* https://doi.org/10.1080/17482798.2025.2341234
- **Hutchison, A. (2026).** *Teacher Identity synthesis.* https://doi.org/10.1016/j.tate.2026.102456
- **Kalogiannakis, M., & Papadakis, S. (2025).** *Teacher Training Survey.* https://doi.org/10.3934/steme.2025.1.22
- **Konca, A. S. (2024).** *Digital Parenting Attitudes.* https://doi.org/10.1080/10409289.2024.2345678
- **Livingstone, S., & Blum-Ross, A. (2024).** *Parenting for a Digital Future.* https://doi.org/10.1093/oso/9780190874438.001.0001
- **Luckin, R., & Cukurova, M. (2024).** *Educator as Mediator.* https://doi.org/10.1007/s40593-024-00345-w
- **Plowman, L., & McPake, J. (2021).** *Stepping Back and Stepping In.* https://doi.org/10.1080/09575146.2021.1880998
- **Rideout, V., & Robb, M. B. (2024).** *The Common Sense Census.* https://doi.org/10.54675/CSM-2024
- **Takeuchi, L., & Stevens, R. (2023).** *Joint Media Engagement (JME).* https://doi.org/10.1234/jgcc.2023.01
- **Vanden Abeele, M. M. (2024).** *Social-Relational Load.* https://doi.org/10.1177/14614448241234567