MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(3), em2793
ISSN:1305-8223 (online)
OPEN ACCESS Review Article https://doi.org/10.29333/ejmste/18068

Applications of generative AI in early childhood education: A systematic review

Yuxin Zhang [1], Siti Hajar Binti Halili [1*], Zamzami Zainuddin [2]

[1] Department of Curriculum and Instructional Technology, University of Malaya, Kuala Lumpur, MALAYSIA
[2] College of Education, Psychology and Social Work, Flinders University, Adelaide, SA, AUSTRALIA

Received 08 September 2025 ▪ Accepted 09 February 2026

Abstract
Generative artificial intelligence (Gen AI) has emerged as a topic of interest in education research. However, its applications in early childhood education (ECE) remain underexplored. This systematic review synthesizes empirical evidence on Gen AI in ECE, organizing findings by stakeholder perspective: young children, families, and teachers. Following PRISMA 2020 guidelines, 29 studies published between December 2022 and July 2025 were identified from eight databases. Results show that publications increased sharply from one in 2023 to 16 in 2024, with research concentrated in the USA and China. Gen AI supports ECE through enhanced learning outcomes, personalized learning, increased engagement, parent support, and teacher efficiency. However, these benefits were consistently dependent on active adult mediation. Challenges include technical limitations, content quality issues, and accuracy concerns. The findings suggest that Gen AI is best positioned as a complement to human guidance rather than a replacement. Future research should employ longitudinal designs and expand geographic diversity.

Keywords: generative AI, early childhood education, education technology, systematic review

INTRODUCTION

Contemporary education is undergoing rapid digital transformation. Among emerging technologies, generative artificial intelligence (Gen AI) has attracted particular attention. Gen AI uses deep learning to produce text, images, and interactive materials in response to user prompts (Yan et al., 2024). These tools can generate context-sensitive responses for planning, instruction, and assessment (Fuchs, 2023). As these technologies become more accessible, researchers have begun exploring their applications in early childhood education (ECE).

ECE has a rich tradition of child-centered pedagogy. Established approaches such as developmentally appropriate practice (Copple & Bredekamp, 2009), the Reggio Emilia approach (Edwards & Gandini, 2018), and Montessori education (Lillard, 2017) have shaped the field. These traditions share several commitments: respecting children as capable learners, tailoring experiences to individual needs, and using observation to inform teaching. These traditions have long supported differentiation, formative assessment, and responsive practice. The emergence of Gen AI raises new questions: How might these tools complement established approaches? What benefits and challenges arise when Gen AI enters settings built around relationships and play?

Systematic reviews have examined Gen AI integration in other educational sectors. Reviews of higher education and K-12 contexts report benefits such as enhanced engagement, personalized learning, and positive tutoring effects. However, these studies also identify persistent concerns about academic integrity and modest effect sizes in short-term interventions (Daniel et al., 2025; Létourneau et al., 2025; Merino-Campos, 2025; Sharma & Panja, 2025). Importantly, these reviews focus on older students who can read, type, and engage with text-based interfaces independently. ECE contexts differ substantially. Young children rely on non-verbal communication, learn primarily through play, and require adult mediation in technology use. Families also play a more central role in early learning.

A small number of reviews have begun to examine Gen AI in ECE, but none have systematically organized findings by stakeholder group. This distinction matters because Gen AI affects young children, families, and teachers in different ways. Primary research reveals both benefits and challenges across stakeholder groups. For

✉ 22104805@siswa.um.edu.my ✉ siti_hajar@um.edu.my **(*Correspondence)** ✉ zamzami.zainuddin@flinders.edu.au

Contribution to the literature

- This review proposes a stakeholder-centered analytical framework for early childhood education (ECE). It systematically examines generative AI applications, benefits, and challenges from the perspectives of young children, families, and teachers.
- Unlike previous reviews focused on older students, this study synthesizes empirical literature on generative AI in ECE from December 2022 to July 2025. It reveals that adult mediation is an essential mechanism through which generative AI benefits are realized in ECE contexts.
- This review identifies critical research gaps, including the absence of longitudinal designs, limited geographic diversity, and underdeveloped applications in STEM and creative domains. These findings offer directions for future research, technology development, and educational policy in ECE settings.

children, studies report enhanced language development, creative expression, and engagement (Lu et al., 2024; Vella et al., 2025; Zhou et al., 2024). For teachers, Gen AI shows promise in material creation and individualized planning (Rakap, 2024; Rakap & Balikci, 2024; Wong et al., 2024). However, research documents challenges, including technical limitations, content quality concerns, and safety considerations (Ho et al., 2025; J. Lee et al., 2024; Vella et al., 2025). A synthesis that distinguishes the perspectives of different stakeholders is needed.

This review addresses that gap by synthesizing empirical evidence on Gen AI applications in ECE and organizing findings by stakeholder perspective. Following international definitions, ECE encompasses settings for children from birth to age eight (United Nations Children’s Fund, 2023). The review examines three stakeholder groups: young children, families, and teachers. Three research questions (RQ) guide this review:

RQ1. What are the key trends in how generative AI is used in ECE?

RQ2. What benefits does generative AI offer to ECE?

RQ3. What challenges exist in applying generative AI in ECE?

By addressing these questions, the review identifies directions for future research and practice.

METHODOLOGY

Information Sources

This review followed the preferred reporting items for systematic reviews (PRISMA 2020) to ensure transparent and reproducible reporting (Page et al., 2021). Eight electronic databases were selected for comprehensive coverage. Web of Science and Scopus provided broad multidisciplinary coverage, while Springer Link and ACM Digital Library focused on educational technology and human-computer interaction. ERIC and PsycINFO captured specialized education and developmental psychology research. IEEE Xplore and PubMed contributed technology-focused studies and child health research, respectively.

Search Strategy

The literature search was conducted in two phases. Initial searches of Web of Science, Scopus, Springer Link, and ACM Digital Library were carried out on August 1, 2025. Supplementary searches of ERIC, PsycINFO, IEEE Xplore, and PubMed were conducted on December 28, 2025, to broaden source coverage. The search focused on publications from December 1, 2022, to July 31, 2025. It aligns with the public release and rapid uptake of Gen AI tools (Perifanou & Economides, 2025).

Keywords were organized into two clusters and combined using Boolean operators (Bramer et al., 2018). The first cluster addressed Gen AI. It includes terms such as “generative artificial intelligence”, “GenAI”, “generative AI”, “AI-generated content”, “ChatGPT”, and “AIGC”. The second cluster targeted ECE, including “preschool”, “early childhood”, and “kindergarten”. The search strings were adapted to suit the structure and indexing practices of each database. Filters were applied to limit results by language (English), document type (journal articles, conference papers, book chapters), and subject area (education or social sciences). **Table 1** presents detailed search strings and the number of records retrieved from each database.

Table 1. Search strings used in the process of finding relevant literature

Database	Search strings	Fields searched	Search date	Results
Web of Science	TS = (“Generative artificial intelligence*” OR “GenAI*” OR “Generative AI*” OR “AI-generated content*” OR “ChatGPT” OR “AIGC”) AND TS = (“Preschool” OR “early childhood*” OR “kindergarten”)	Topic	August 1, 2025	32
Scopus	TITLE-ABS-KEY (“Generative artificial intelligence*” OR “GenAI” OR “Generative AI*” OR “AI-generated content*” OR “CHATGPT” OR “AIGC”) AND TITLE-ABS-KEY (“PRESCHOOL” OR “EARLY CHILDHOOD*” OR “Kindergarten”)	Title, Abstract, Keywords	August 1, 2025	101

Table 1 (Continued). Search strings used in the process of finding relevant literature

Database	Search strings	Fields searched	Search date	Results
Springer Link	("Generative artificial intelligence*" OR "GenAI" OR "Generative AI*" OR "AI-generated content*" OR "CHATGPT" OR "AIGC") AND ("PRESCHOOL" OR "EARLY CHILDHOOD*" OR "Kindergarten")	Keywords	August 1, 2025	611
ACM Library	[[All: "generative artificial intelligence*"] OR [All: "genai"] OR [All: "generative ai*"] OR [All: "ai-generated content*"] OR [All: "chatgpt"] OR [All: "aigc"]] AND [[All: "preschool"] OR [All: "early childhood*"] OR [All: "kindergarten"]]	All Fields	August 1, 2025	354
ERIC	("Generative artificial intelligence*" OR "GenAI" OR "Generative AI*" OR "AI-generated content*" OR "CHATGPT" OR "AIGC") AND ("PRESCHOOL" OR "EARLY CHILDHOOD*" OR "Kindergarten")	All Fields	December 28, 2025	27
PsycINFO	Abstract: ("Generative artificial intelligence*" OR "GenAI" OR "Generative AI*" OR "AI-generated content*" OR "CHATGPT" OR "AIGC") AND (Abstract: "PRESCHOOL" OR Abstract: "EARLY CHILDHOOD*" OR Abstract: "Kindergarten")	Title, keywords, abstract	December 28, 2025	1
IEEE Xplore	("Document Title":("Generative artificial intelligence*" OR "Document Title":"GenAI" OR "Document Title":"Generative AI" OR "Document Title":"AI-generated content*" OR "Document Title":"CHATGPT" OR "Document Title":"AIGC")) AND ("Document Title":("PRESCHOOL" OR "Document Title":"EARLY CHILDHOOD*" OR "Document Title":"Kindergarten"))	Title	December 28, 2025	0
PubMed	(("Generative artificial intelligence*"[Title/Abstract] OR "GenAI"[Title/Abstract] OR "Generative AI*"[Title/Abstract] OR "AI-generated content*"[Title/Abstract] OR "CHATGPT"[Title/Abstract] OR "AIGC"[Title/Abstract])) AND (("PRESCHOOL"[Title/Abstract] OR "EARLY CHILDHOOD*"[Title/Abstract] OR "Kindergarten"[Title/Abstract]))	Title, abstract	December 28, 2025	3

Table 2. Inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
Period	Studies published from December 1, 2022 to July 31, 2025	Studies published before December 1, 2022 and after July 31, 2025
Language	English	Languages other than English
Accessibility	Open or institutional access to the full text	Not accessible
Research categories	Early childhood education & educational research	Other areas
Source	Peer-reviewed journal articles, conference papers, and book chapters	Literature other than peer-reviewed journal articles, conference papers, and book chapters
Study type	Empirical studies	Non-empirical studies
Participants	Preschool-aged children (3-8 years) or relevant stakeholders (teachers, parents, educational experts)	Participants not related to early childhood education
Topic	Focused on the use of GenAI; education-related	Not focused on the use of GenAI; not education-related

To further strengthen the comprehensiveness of the review, forward and backward citation searching was conducted on the included studies. This supplementary step aims to capture relevant research that may have been missed through the initial database searches. Retrieved articles were evaluated for quality and relevance against clearly defined inclusion and exclusion criteria (**Table 2**).

Data Collection Process

The data collection process comprised four phases following the PRISMA 2020 framework. In the identification phase, searches across eight databases yielded 1,129 records based on the search strings detailed in **Table 1**. Database filters were then applied to remove records based on language, document type, and subject area, resulting in the exclusion of 847 records.

During screening, titles and abstracts of 282 records were reviewed, excluding 34 review or meta-analysis articles. Full texts were sought for the remaining 248 records. Four were inaccessible, and 244 were assessed for eligibility. Of these, 219 were excluded for not focusing on Gen AI (n = 87), not involving ECE (n = 64), or being other non-empirical studies (n = 68).

Five duplicates were removed, resulting in 20 studies from database searching. The relatively low duplicate

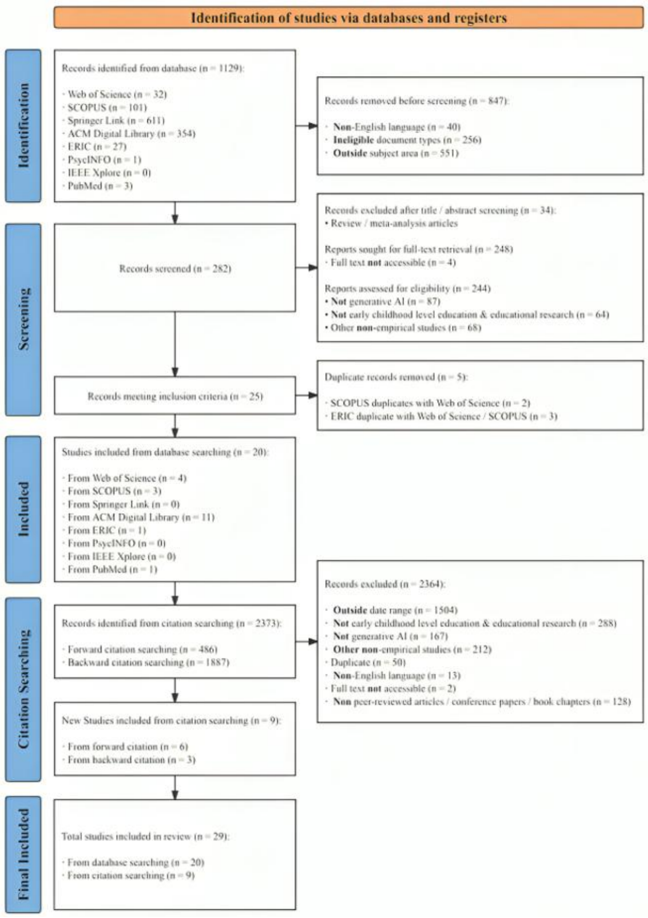

Figure 1. PRISMA flow diagram (Adapted from Page et al., 2021)

rate (20%) reflects the emerging nature of Gen AI research in ECE. It also reflects distinct coverage across education-focused and technology-focused databases. Citation searching of these 20 studies identified 2,373 additional records, from which 9 new studies were included. The final sample comprised 29 studies. The complete selection process is presented in **Figure 1**.

Study Selection and Quality Assessment

Study selection and data extraction were conducted independently by two reviewers following the PRISMA 2020 guidelines. Titles, abstracts, and full texts were screened based on predefined inclusion and exclusion criteria, as detailed in **Table 2**.

Data were extracted using a standardized form capturing publication year, country, participant characteristics, and methods (**Table 3**). Disagreements between reviewers were resolved through discussion, and a third reviewer was consulted as an arbitrator when consensus could not be reached.

The methodological quality of the included studies was evaluated using the mixed methods appraisal tool (MMAT). This framework is suitable for appraising quantitative, qualitative, and mixed-methods studies. Two reviewers independently performed the assessment; any disagreements were resolved through discussion. Quality assessment results are presented in

Table 3. Overview of the selected articles

Reference	Sample	Country	Methodology
Y. L. Zhang (2025)	433 preschool teachers	Malaysia	Quantitative survey
Su and Yang (2024)	10 kindergarten teachers	China, Hong Kong	Qualitative interviews
Seiradakis (2023)	6 early childhood special education experts	Greece	Qualitative expert interviews
M. Sun et al. (2025)	10 preschool teachers (from 5 provinces: Zhejiang, Heilongjiang, Jiangxi, Shandong, Shanghai)	China	Mixed methods (surveys, semi-structured interviews)
Lu et al. (2024)	60 preschoolers (5-6 years old)	China, Taiwan	Mixed methods (quasi-experimental [2×2 factorial design])
Zhao et al. (2024)	8 children (4-7 years old)	China	Mixed user study (pre-/post-test, observation, interview)
Wong et al. (2024)	97 in-service early childhood educators	China, Hong Kong	Mixed methods (questionnaire. open-ended question)
Rakap (2024)	22 novice special education teachers (working with preschool children with autism)	USA	RCT
Rakap and Balikci (2024)	30 special education teachers (working with preschool children with autism; ≥ 2 years experience)	USA/Turkey	RCT
Y. Sun et al. (2024)	17 parents of preschoolers (3-6 years old)	China/USA/ China, Hong Kong	Qualitative interviews
J. Lee et al. (2024)	9 families with preschool children (ages 50-71 months)	South Korea	Mixed methods (4-week deployment, voice collection, reading logs, surveys, semi-structured interviews)
Ho et al. (2025)	20 parent-child dyads (children aged 3-5)	USA	Mixed methods (interviews, surveys)
C. Zhang et al. (2024)	35 children (ages 4-8)	USA	Mixed methods (randomized controlled study, pre-/post-assessments, surveys)

Table 3 (Continued). Overview of the selected articles

Reference	Sample	Country	Methodology
Y. Liu et al. (2025)	24 parent-child pairs (children aged 6-8)	China/Canada	Mixed methods (between-subjects comparative study, pre-/post-assessments, questionnaires, interviews, video transcription)
D. Liu et al. (2024)	18 participants (7 families: 8 children aged 4-7, 10 parents) + 4 therapists	China	Qualitative empirical (semi-structured interviews, video observation, focus group)
He et al. (2025a)	5 early childhood teachers from predominantly Latine schools	USA	Qualitative case study
Lyu et al. (2024)	24 autistic children (M = 6.0 years) with their caregivers from two special education centers	Canada/China	RCT (between-subjects)
He et al. (2025b)	23 children (ages 4-7) and parents	USA	Mixed-methods user study (child questionnaires, parent semi-structured interviews)
Dietz Smith et al. (2024)	12 parent-child dyads (children aged 4-6); 4 educators; 8 families (iterative testing)	USA	Mixed-methods (qualitative user study, semi-structured interviews; Wilcoxon tests)
Y. Lee et al. (2023)	16 children (aged 5-7); 4 child education experts; crowd workers	South Korea	Mixed-methods (within-subjects experiment, semi-structured interviews, immediate assessment)
Chin et al. (2024)	8 child-parent dyads (children aged 5-6)	USA	Mixed-methods (within-subjects experiment, semi-structured interviews)
X. Liu et al. (2025)	73 children (aged 4-8)	USA/China	Mixed methods (between-subjects experiment)
Wang et al. (2025)	32 parent-child pairs (children aged 5-6)	China/USA	Mixed-methods (between-subjects experiment, semi-structured interviews)
Shen et al. (2025)	20 parent-child dyads (children aged 5-8)	USA	Mixed-methods (within-subjects experiment, semi-structured interviews)
S. Sun et al. (2024)	12 parent-child pairs (children aged 5-7)	China	Mixed-methods (single-group user study with scales and interviews)
Xu et al. (2025)	119 children aged 4-8	USA	RCT with questionnaire & behavioral coding
Li et al. (2024)	71 children aged 4-8 years	USA	RCT (between-subjects experiment with behavioral coding & questionnaire)
Zhou et al. (2024)	8 children aged 4-6 years	China	Qualitative case study
Vella et al. (2025)	15 children aged 3-5 years	Australia	Qualitative observational study

Table 4, using a traffic-light format (Y = yes, CT = can't tell, N = no).

RESULTS

Current Trends of Generative AI in ECE

This review analyzes 29 articles published between December 1, 2022, and July 31, 2025. **Table 3** provides an overview of the 29 included studies. Child participants ranged from 3 to 8 years of age. Adult participants included preschool teachers, early childhood educators, special education teachers, and families with young children. Only one article appeared in 2023. The number rose sharply in 2024 to 16. By July 2025, 12 articles had been published, which signals sustained activity and suggests further growth by year-end. This pattern aligns

Table 4. Quality assessment results

Study	Design	S1	S2	Q1	Q2	Q3	Q4	Q5	Rating
Y. L. Zhang (2025)	QD	Y	Y	Y	Y	Y	CT	Y	Medium
Su and Yang (2024)	QL	Y	Y	Y	Y	Y	Y	Y	High
Seiradakis (2023)	QL	Y	Y	Y	Y	Y	Y	Y	High
M. Sun et al. (2025)	MM	Y	Y	Y	Y	Y	CT	Y	Medium
Lu et al. (2024)	MM	Y	Y	CT	Y	Y	CT	Y	Medium
Zhao et al. (2024)	MM	Y	Y	Y	CT	Y	CT	CT	Medium
Wong et al. (2024)	MM	Y	Y	Y	CT	Y	CT	Y	Medium

Table 4 (Continued). Quality assessment results

Study	Design	S1	S2	Q1	Q2	Q3	Q4	Q5	Rating
Rakap (2024)	RCT	Y	Y	CT	Y	Y	Y	Y	Medium
Rakap and Balikci (2024)	RCT	Y	Y	CT	Y	Y	Y	Y	Medium
Y. Sun et al. (2024)	QL	Y	Y	Y	Y	Y	Y	Y	High
J. Lee et al. (2024)	MM	Y	Y	Y	Y	Y	Y	Y	High
Ho et al. (2025)	MM	Y	Y	Y	Y	Y	CT	Y	Medium
C. Zhang et al. (2024)	MM	Y	Y	Y	Y	Y	Y	Y	High
Y. Liu et al. (2025)	MM	Y	Y	Y	Y	Y	Y	Y	High
D. Liu et al. (2024)	QL	Y	Y	Y	Y	Y	Y	Y	High
He et al. (2025a)	QL	Y	Y	Y	Y	Y	Y	Y	High
Lyu et al. (2024)	RCT	Y	Y	CT	CT	Y	N	Y	Low
He et al. (2025b)	MM	Y	Y	Y	CT	Y	Y	Y	Medium
Dietz Smith et al. (2024)	MM	Y	Y	Y	Y	Y	Y	Y	High
Y. Lee et al. (2023)	MM	Y	Y	Y	Y	Y	Y	Y	High
Chin et al. (2024)	MM	Y	Y	Y	Y	Y	Y	CT	Medium
X. Liu et al. (2025)	MM	Y	Y	Y	Y	Y	Y	Y	High
Wang et al. (2025)	MM	Y	Y	Y	Y	Y	Y	Y	High
Shen et al. (2025)	MM	Y	Y	Y	Y	Y	Y	Y	High
S. Sun et al. (2024)	MM	Y	Y	Y	Y	Y	Y	CT	Medium
Xu et al. (2025)	RCT	Y	Y	Y	Y	Y	CT	Y	Medium
Li et al. (2024)	RCT	Y	Y	Y	Y	Y	CT	Y	Medium
Zhou et al. (2024)	QL	Y	Y	Y	CT	CT	CT	CT	Medium
Vella et al. (2025)	QL	Y	Y	Y	Y	Y	Y	Y	High

Note. Legend: Y = Yes, CT = Can't tell, N = No (high risk); Design: QL = Qualitative; QD = Quantitative descriptive; QN = Quantitative non-randomized; MM = Mixed methods; S1 = Clear research questions; S2 = Data allow to address research questions; Q1-Q5 correspond to methodological criteria specific to each study design:
Qualitative (1.1-1.5): Approach appropriate, data collection adequate, findings derived, interpretation substantiated, coherence; RCT (2.1-2.5): Randomization, baseline comparability, complete outcome data, blinding, adherence; Non-randomized (3.1-3.5): Representativeness, measurements appropriate, complete data, confounders, intervention as intended; Descriptive (4.1-4.5): Sampling strategy, sample representative, measurements appropriate, nonresponse bias low, statistical analysis; & Mixed methods (5.1-5.5): Rationale, integration effective, outputs adequate, divergences addressed, component quality

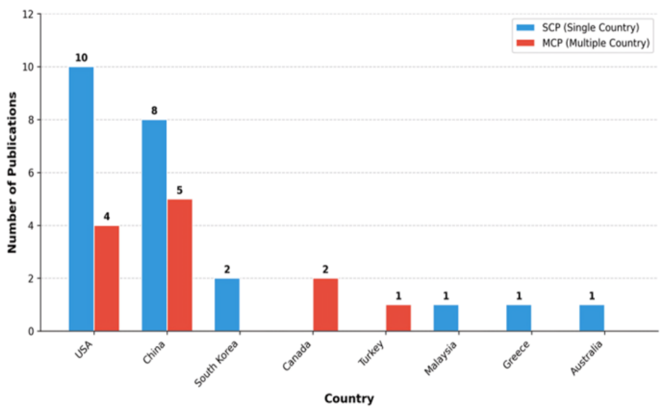

Figure 2. Geographical distribution of articles by country and type of collaboration (SCP vs. MCP) (Source: Authors' own elaboration)

with the rapid diffusion of Gen AI tools and increasing research attention in ECE settings.

As shown in **Figure 2**, the included studies originate from eight countries across Asia, Europe, North America, and Oceania. Single-country publications predominate (n = 23, 79.3%), with the remaining six studies involving multinational collaboration. By country contribution, the USA was involved in 14 studies (48.3%), followed by China (n = 13, 44.8%), South Korea (n = 2, 6.9%), and Canada (n = 2, 6.9%). Australia, Greece, Malaysia, and Turkey each contributed one study.

Six studies (20.7%) involved multi-country collaborations. These collaborations were primarily between the USA and China (n = 4), followed by China-Canada (n = 2) and USA-Turkey (n = 1). The concentration of research in the USA and China may reflect the early adoption and availability of Gen AI technologies in these regions, though caution is warranted given the small sample size.

To identify thematic clusters, an author keyword co-occurrence network was generated using the Bibliometrix R package. The analysis used author keywords with 50 words, minimum cluster frequency of 3, and Walktrap clustering algorithm. A synonym list was applied to standardize terms, and irrelevant terms were removed. As shown in **Figure 3**, three thematic clusters emerged. The red cluster centers on teachers and technology integration, including "teacher perceptions," "early childhood education," and "technology integration." The blue cluster focuses on children's learning interactions, featuring "children," "storytelling," "human-AI interaction," and "agent system." The green cluster addresses special education applications, encompassing "preschool children with

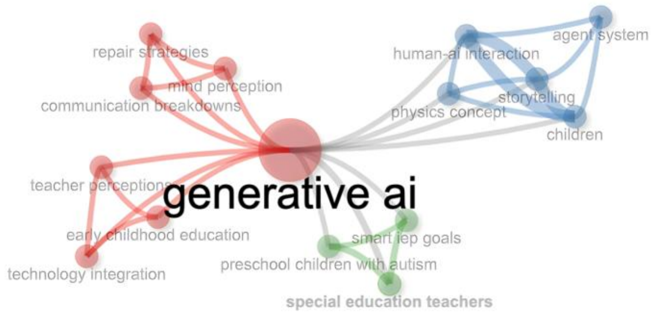

Figure 3. Author keywords co-occurrence network (Source: Authors' own elaboration)

autism," "special education teachers," and "SMART IEP goals." "Generative AI" serves as the central node connecting all three clusters, reflecting its role as the core topic across diverse research themes.

Benefits of Generative AI in ECE

Analysis of the 29 included studies identified five thematic categories of Gen AI benefits in ECE:

(1) enhanced learning outcomes,

(2) personalized and adaptive learning,

(3) enhanced engagement,

(4) parent support and interaction, and

(5) teacher efficiency.

Table A1 in **Appendix A** presents these themes and subcategories with study characteristics, methodological quality ratings, and effect sizes where reported. The following sections synthesize these findings by stakeholder perspective: young children, parents and families, and teachers.

Benefits for young children

The results showed that Gen AI applications in ECE were associated with benefits for young children in three areas: learning outcomes, personalized learning experiences, and engagement.

Learning outcomes were examined across multiple developmental domains. Language development received the most attention. Gen AI-powered storytelling systems were associated with improvements in narrative abilities such as comprehension, oral expression, and story creation (Lu et al., 2024). Y. Liu et al. (2025) focused on spatial language and found greater gains among children using an LLM-based tangible play system. Personalization appeared to strengthen these effects. When AI-generated stories aligned with children's interests, vocabulary acquisition increased compared to generic content (J. Lee et al., 2024). Conversational agents were also used to support cognitive development. These tools were used to teach physics (Zhao et al., 2024), math (C. Zhang et al., 2024), and critical thinking (Wang et al., 2025). Notably, Y. Lee et al. (2023) found no significant difference between Gen AI and human reading partners. This suggests a complementary rather than superior role for Gen AI. By contrast, social-emotional learning and creativity received less attention. Lyu et al. (2024) found improved emotion identification in autistic children. Shen et al. (2025) reported increased use of affect words during Gen AI-supported reading. Lu et al. (2024) and Su and Yang (2024) documented benefits for creative expression, though empirical evidence in these areas remains limited.

A key mechanism underlying these benefits was Gen AI's capacity to adapt to individual children. This adaptation occurred at multiple levels. At the content level, AI-generated stories could match children's interests (J. Lee et al., 2024) or adjust based on their responses during play (Y. Liu et al., 2025). At the interaction level, developmentally appropriate responses reduced distraction and increased willingness to continue learning (Y. Lee et al., 2023). At the support level, adjustable difficulty levels helped autistic children learn emotion identification (Lyu et al., 2024). Gen AI tools also accommodated diverse abilities in inclusive settings (He et al., 2025b; Seiradakis, 2023). Meanwhile, Rakap and Balikci (2024) found that such adaptation improved individualized education program (IEP) quality. However, these studies relied on short-term interventions. Whether these benefits persist over time remains unclear.

Alongside these outcomes, studies reported high engagement with Gen AI tools. Children sustained attention and expressed positive attitudes toward Gen AI as a learning companion (C. Zhang et al., 2024; He et al., 2025b; Wong et al., 2024). Y. Lee et al. (2023) found reduced distraction and greater willingness to reuse the system. Parents perceived similar effects. Y. Sun et al. (2024) reported that caregivers noticed improved focus, though this relied on parent perceptions. However, the quality of engagement warrants scrutiny. Chin et al. (2024) found that ChatGPT provided more positive feedback than parents during shared reading. Yet parents generated more conversational turns. This raises questions about whether Gen AI-facilitated engagement reflects the reciprocal interactions emphasized in early childhood pedagogy.

Benefits for parents and family

Studies examined how Gen AI tools supported parents in home-based learning contexts. Findings clustered around four areas: reduced burden, role transformation, enhanced interaction, and parent capability.

A common finding was that Gen AI reduced parental burden during shared learning activities. Tools provided scaffolding that parents typically struggle to offer, such as guidance resources and structured prompts (Wang et al., 2025). Some systems enabled asynchronous engagement, allowing parents to participate in their child's learning without co-viewing videos (Shen et al., 2025). Others facilitated real-time interaction, increasing conversational turns during shared reading (Dietz Smith et al., 2024; Y. Lee et al., 2023). These findings suggest perceived reductions in effort, though whether this reflects genuine support or cognitive offloading warrants consideration.

This reduced burden appeared to transform parental roles. When Gen AI mediated conversations, parents shifted from "leaders" to "observing supporters" (Dietz Smith et al., 2024; Wang et al., 2025). This shift may reduce parental pressure. However, it raises questions about whether parents become less actively involved in their children's learning.

Studies also reported changes in parent-child interaction patterns. Gen AI facilitated new forms of collaboration, such as family co-creation of storybooks (D. Liu et al., 2024) and parent-child play (Y. Liu et al., 2025). Parents described these tools as a "bridge" for communication. However, these findings relied primarily on parent perceptions. How such interactions compare to unmediated exchanges remains unclear.

Studies suggested that Gen AI may build parent capability. Parents reported learning new pedagogical strategies and social-emotional terminology (Ho et al., 2025; Shen et al., 2025). They also gained confidence in understanding their child's progress (Y. Liu et al., 2025). While promising, these findings reflect values of efficiency and optimization. Such values may conflict with ECE perspectives emphasizing sustained, reciprocal relationships.

Benefits for teachers

Studies examined how Gen AI tools supported ECE teachers in their professional tasks. Findings clustered around content creation, time savings, documentation quality, and administrative efficiency.

Teachers reported using Gen AI to develop learning materials and activity plans (Wong et al., 2024). This included culturally responsive content for diverse learners (He et al., 2025a). Teachers also experienced time savings. Rakap (2024) found that teachers using ChatGPT spent less time developing IEP goals. Importantly, reduced time did not compromise quality. Two RCT studies found that Gen AI-assisted IEPs received higher quality scores than those produced without AI support (Rakap, 2024; Rakap & Balikci, 2024).

Teachers also perceived administrative benefits. Early childhood special education experts viewed Gen AI as useful for streamlining routine tasks (M. Sun et al., 2025; Seiradakis, 2023). However, evidence in this area relied on self-report measures. Whether perceived efficiency translates to improved teaching practice or child outcomes remains unclear.

Challenges of Gen AI in ECE

Analysis of the included studies identified challenges across five areas:

(1) technical limitations,

(2) content quality,

(3) accuracy and bias,

(4) contextual constraints, and

(5) safety and implementation concerns.

Table B1 in **Appendix B** presents these themes with study characteristics and methodological quality ratings.

Technical limitations posed fundamental barriers to Gen AI use with young children. Speech recognition emerged as a primary concern. Systems struggled to interpret young children's developing pronunciation and speech patterns (Li et al., 2024; Vella et al., 2025). This led to communication breakdowns that disrupted learning activities. Beyond speech recognition, response latency affected interaction flow (Zhou et al., 2024), and system stability issues interrupted ongoing activities (C. Zhang et al., 2024; Vella et al., 2025). Network problems created additional barriers in some contexts (Su & Yang, 2024). These technical constraints highlight that current Gen AI systems were not designed with young children's developmental characteristics in mind.

Even when systems functioned properly, content quality remained problematic. The core issue was a mismatch between Gen AI output and young children's developmental levels. Generated content often exceeded children's comprehension (Chin et al., 2024; Shen et al., 2025). Meanwhile, systems produced inappropriate or inconsistent material that required adult screening (Dietz Smith et al., 2024; M. Sun et al., 2025; Su & Yang, 2024; Wang et al., 2025). This unpredictability made reliability a persistent concern (C. Zhang et al., 2024; Dietz Smith et al., 2024; Y. Lee et al., 2023).

More fundamentally, accuracy and bias raised concerns about educational integrity. Gen AI systems generated plausible but incorrect information, a phenomenon known as hallucination (C. Zhang et al., 2024; Y. Lee et al., 2023). Young children are particularly vulnerable to such errors because they may not distinguish accurately from inaccurate content (He et al., 2025a; Seiradakis, 2023). Compounding this risk, systems could perpetuate stereotypes embedded in training data (He et al., 2025a; Y. Lee et al., 2023), and errors could accumulate across multi-turn interactions (Y. Lee et al., 2023).

Contextual constraints further limited Gen AI's suitability for young children. A fundamental limitation is that current text-based systems cannot perceive non-verbal communication. They miss facial expressions, gestures, and actions central to how young children communicate (Chin et al., 2024; Dietz Smith et al., 2024; Wang et al., 2025; Xu et al., 2025). Systems also struggled to maintain coherence across conversations and often lacked personalization features (Dietz Smith et al., 2024; Shen et al., 2025; Vella et al., 2025; Y. Lee et al., 2023). These limitations conflict with ECE principles emphasizing responsive, individualized interactions.

Finally, safety and implementation challenges shaped how Gen AI could be deployed. A consistent finding was that adult involvement remained essential. Parents and teachers needed to monitor content safety and address privacy concerns (C. Zhang et al., 2024; He et al., 2025a; J. Lee et al., 2024; Shen et al., 2025; Vella et al., 2025). They also needed to intervene during activities to support children's interactions (D. Liu et al., 2024; Dietz Smith et al., 2024; He et al., 2025a; Ho et al., 2025; Y. Liu et al., 2025). Additionally, adults managed cognitive load for children unfamiliar with the technology (M. Sun et al., 2025; Wang et al., 2025; Y. Sun et al., 2024). Teachers reported needing more training and resources to fulfill these roles effectively (Seiradakis, 2023; Su & Yang, 2024; Wong et al., 2024). These findings suggest that Gen AI in ECE requires substantial scaffolding and cannot replace human guidance.

DISCUSSION

This review examined 29 empirical studies on Gen AI applications in ECE. The analysis addressed research trends (**RQ1**), benefits (**RQ2**), and challenges (**RQ3**). The findings reveal an emerging field responding rapidly to technological developments. Publications surged from one study in 2023 to 16 in 2024. This growth demonstrates researchers' swift attention to Gen AI's potential in ECE following ChatGPT's public release in late 2022 (Seiradakis, 2023). Research is geographically concentrated in the USA and China, reflecting early technology adoption and research capacity in these regions. However, this concentration limits understanding of how Gen AI integration may vary across diverse cultural and educational contexts. Research also concentrates on language and storytelling applications, with limited attention to STEM domains. This pattern may reflect early research priorities rather than inherent limitations of the technology. Rather than treating benefits and challenges as separate categories, the following discussion examines how they interact. This integrated analysis draws on established theoretical perspectives to interpret findings. These perspectives include developmentally appropriate practice (Copple & Bredekamp, 2009) and sociocultural theory (Vygotsky, 1978).

A critical finding is that benefits and challenges are deeply interconnected. Reported benefits included learning gains, personalization, and engagement. Yet these benefits were consistently dependent on conditions that the identified challenges directly threaten. Personalization emerged as a key mechanism supporting learning outcomes. AI-generated stories matching children's interests enhanced vocabulary acquisition (J. Lee et al., 2024). Adaptive difficulty levels supported autistic children's emotion identification (Lyu et al., 2024). However, these benefits occurred within carefully structured interventions. Adults actively mediated children's interactions in these studies. When technical challenges such as speech recognition failures disrupted this mediation, learning activities broke down (Li et al., 2024; Vella et al., 2025).

Adult mediation emerges as the central theme connecting benefits and challenges. Every category of benefits involved substantial adult involvement, including children's learning outcomes, parental support, and teacher efficiency. Meanwhile, every category of challenges requires adult intervention to address technical limitations, content quality issues, accuracy concerns, and safety considerations. For parents, they needed to monitor content safety and manage cognitive load (J. Lee et al., 2024; Wang et al., 2025). For teachers, they needed to screen inappropriate material and intervene during activities (He et al., 2025a; M. Sun et al., 2025). This finding aligns with established theoretical perspectives in ECE. Vygotsky's (1978) sociocultural theory positions learning as fundamentally social. Specifically, learning occurs through interaction with more capable others who provide scaffolding within the zone of proximal development. From this perspective, Gen AI in some studies cannot independently support young children's learning. It lacks the responsive, contingent interaction that effective scaffolding requires. Building on this, adults remain essential as mediators. In this case, they need to interpret children's needs, adjust support accordingly, and repair breakdowns when they occur.

Furthermore, developmentally appropriate practice similarly emphasizes relationships. Young children learn through interactions with responsive adults who observe, interpret, and respond to children's cues (Copple & Bredekamp, 2009). Current Gen AI systems cannot fulfill this responsive role. The benefits observed reflect thoughtful integration of Gen AI into adult-mediated learning experiences. They do not reflect what Gen AI can achieve independently. This creates a practical tension. Gen AI is often positioned as reducing the adult workload. Yet effective implementation appears to require substantial adult engagement. Whether Gen AI ultimately reduces or redistributes adult demands remains an open question.

Additionally, the reviewed studies frequently emphasize efficiency. For instance, teachers saved time on IEP development (Rakap, 2024). Parents experienced reduced burden during shared activities (Wang et al., 2025). Content creation for learning materials became faster (Wong et al., 2024). While these outcomes may address genuine pressures facing educators and families, they warrant critical examination. Efficiency and optimization are not core values in ECE traditions.

The Reggio Emilia approach emphasizes slow, emergent learning through sustained relationships (Edwards & Gandini, 2018). Therefore, the process of learning holds as much value as outcomes.

Meanwhile, Montessori education values children's self-paced discovery over adult-directed efficiency (Lillard, 2017). Developmentally appropriate practice foregrounds responsive relationships and play-based learning rather than productivity metrics (Copple & Bredekamp, 2009). Studies report that Gen AI shifts parental roles from "leaders" to "observing supporters" (Dietz Smith et al., 2024; Wang et al., 2025). This transformation merits careful consideration. Reduced parental burden may genuinely support overstretched families. Alternatively, it may represent cognitive offloading that diminishes meaningful engagement. The reviewed studies do not provide sufficient evidence to distinguish between these interpretations.

Similarly, teachers reported time savings on administrative tasks. However, this evidence relied primarily on self-report measures (M. Sun et al., 2025; Seiradakis, 2023). Whether perceived efficiency translates to improved teaching practice or child outcomes remains unclear. If time saved on documentation allows teachers to invest more in direct interactions with children, efficiency gains may serve ECE values well. However, if efficiency becomes the primary criterion for evaluating Gen AI in ECE, this presents a risk. Such an approach may import values that conflict with the relational foundations of quality early childhood practice.

The findings collectively support positioning Gen AI as a complement to human guidance rather than a replacement. C. Zhang et al. (2024) found no significant difference between Gen AI and human reading partners in supporting children's learning. This finding suggests an appropriate role for Gen AI. It can provide support comparable to human partners in specific, structured contexts. However, it does not surpass human capabilities. This complementary positioning aligns with the challenges identified across studies. Current Gen AI systems cannot perceive non-verbal communication, central to how young children express themselves (Chin et al., 2024; Dietz Smith et al., 2024). They cannot maintain the coherent, personalized content that responsive teaching requires (Vella et al., 2025; Y. Lee et al., 2023). They also produce inaccuracies that young children cannot identify (He et al., 2025a).

However, these limitations do not preclude valuable applications. The strongest evidence emerged in teacher-facing applications. Two RCT studies demonstrated quality improvements alongside time savings in IEP development (Rakap, 2024; Rakap & Balikci, 2024). Administrative and planning tasks may represent the most appropriate current applications. In these contexts, Gen AI supports adult professionals rather than directly interacting with children. For child-facing applications, the evidence suggests that Gen AI works best as one element within adult-orchestrated learning experiences.

Situating these findings within the broader educational technology literature reveals important distinctions. Reviews of Gen AI in higher education emphasize academic integrity concerns (Bittle & El-Gayar, 2025; Yusuf et al., 2024). Students may use Gen AI to complete assignments dishonestly. This concern is largely absent from ECE research. Young children do not produce written work vulnerable to AI-generated plagiarism. Instead, ECE research foregrounds developmental appropriateness and safety. The technical challenges identified in this review reflect that current Gen AI systems were largely designed for adult users. Speech recognition failures with young children's developing pronunciation illustrate this mismatch (Li et al., 2024; Vella et al., 2025). Content quality challenges similarly reflect systems trained on adult-oriented text (M. Sun et al., 2025; Su & Yang, 2024). The essential role of adult mediation also distinguishes ECE from other contexts. In K-12 and higher education, students can interact with Gen AI independently. They can read and type without adult assistance. Young children, by contrast, rely on non-verbal communication. They learn primarily through play and physical exploration (Copple & Bredekamp, 2009). They require adult support for virtually all technological interactions. Findings from Gen AI research with older students, therefore, cannot be directly applied to ECE settings.

The findings suggest several implications for practice. For teachers, Gen AI may be most valuable for administrative and documentation tasks. The evidence supporting Gen AI-assisted IEP development suggests potential for reducing documentation burden while maintaining quality (Rakap, 2024; Rakap & Balikci, 2024). However, teachers should approach Gen AI-generated content with professional judgment. They need to screen for developmental appropriateness, accuracy, and alignment with individual children's needs. For families, Gen AI tools may support home-based learning when parents remain actively engaged. The benefits observed occurred with parental involvement. Outcomes without such involvement remain unknown. Parents should view Gen AI as a tool for enhancing their interactions with children, not replacing them. For policymakers and technology developers, the persistent technical challenges highlight a need. Gen AI systems should be designed specifically for young children's developmental characteristics. Current systems largely require children to adapt to adult-oriented interfaces. Developmentally appropriate design would reverse this relationship.

CONCLUSION

This systematic review analyzed the use of Gen AI in ECE based on 29 empirical studies. The findings reveal an emerging field with demonstrated benefits alongside substantial challenges.

This study aimed to identify the current state of Gen AI research in ECE. It also examined the benefits and challenges associated with using Gen AI with young children, families, and teachers. The first question of the study focused on the current trends of Gen AI in ECE. The results showed that publications increased sharply from one study in 2023 to 16 in 2024. This suggests Gen AI will be one of the most important research topics in ECE in the coming years. Research is currently concentrated in the USA and China, with limited representation from other regions. This work is among the first to systematically analyze Gen AI in ECE from stakeholders' perspectives. Therefore, the findings can contribute to the literature and provide insights for researchers and practitioners.

The findings show that Gen AI supports ECE across multiple areas. These include enhanced learning outcomes, personalized and adaptive learning experiences, increased engagement, parent support and interaction, and teacher efficiency in administrative tasks. However, these benefits were consistently dependent on active adult mediation. This finding aligns with sociocultural theory and developmentally appropriate practice, which emphasize the central role of responsive adults in young children's learning.

The findings also identified several challenges. Technical limitations, particularly speech recognition failures with young children's developing pronunciation, posed fundamental barriers. Content quality issues arose from mismatches between Gen AI output and children's developmental levels. Accuracy and bias concerns also raised questions about educational integrity. The findings suggest that Gen AI is best positioned as a complement to, rather than a replacement for human guidance in ECE.

The findings also reveal several gaps in the current literature that warrant attention in future research. First, most studies employed small samples, short-term interventions, and self-report measures. Longitudinal studies with larger samples and more rigorous designs are needed to determine whether observed benefits persist over time. Second, the geographic concentration of research in the USA and China limits cross-cultural generalizability. Future studies should examine how cultural and contextual factors shape Gen AI integration in diverse educational settings. Third, applications in STEM, arts, and creativity domains remain underexplored and merit systematic investigation. Fourth, future studies should examine the impact of specific pedagogical approaches and teaching strategies on Gen AI effectiveness.

Finally, this review highlights the need for developmentally appropriate Gen AI systems designed specifically for young children. Current systems require children to adapt to adult-oriented interfaces. Future development should reverse this relationship. Additionally, teachers and parents require training and resources to integrate Gen AI effectively while maintaining the relational foundations of quality early childhood practice. Future research should focus not only on efficiency outcomes but also on how Gen AI integration affects the quality of adult-child relationships.

Limitations

The included studies offer valuable initial insights into Gen AI applications in ECE. They span multiple countries and involve diverse stakeholder groups, including children, teachers, parents, and therapists. Nevertheless, several methodological limitations warrant careful consideration. First, the majority of studies employed relatively small sample sizes. Approximately 20 of the 29 studies involved 35 or fewer participants. The smallest samples comprised only 5 teachers, 6 experts, or 8 children, thereby limiting statistical power and generalizability. Second, a substantial reliance on self-reported data was observed. Twelve studies (43%) used exclusively self-report measures, while only 2 studies (7%) employed purely objective assessments. This dependence may introduce response biases and compromise the objectivity of findings. Third, most studies utilized single-session or short-term interventions, ranging from 30 minutes to several weeks. Such brief exposure raises concerns about novelty effects. Observed benefits may reflect initial enthusiasm rather than sustained educational impact. The absence of long-term follow-up also means that the durability of reported outcomes remains uncertain. Finally, only 5 studies (18%) employed randomized controlled trials (RCT), while 8 studies (29%) adopted qualitative approaches that preclude causal inferences. Several studies were also conducted in controlled laboratory settings, potentially limiting ecological validity.

Beyond individual study limitations, several notable gaps persist in the current research landscape. Most critically, longitudinal research is virtually absent. Nearly all included studies employed cross-sectional or short-term designs. This precludes assessing whether observed benefits endure over time or how children's interactions with Gen AI evolve developmentally. Geographic and cultural diversity also remains limited. Studies are concentrated predominantly in the USA and China, with minimal representation from other global regions. This geographic concentration limits cross-cultural generalizability and raises questions about how contextual factors shape Gen AI integration in diverse educational settings. Research contexts were also unevenly distributed. More studies were conducted in libraries and home environments than in formal preschool or kindergarten classrooms. Special education settings were also underrepresented. Finally, coverage across educational domains remains uneven. Literacy and language development received considerable attention, while applications in STEM learning, arts, and creativity remain underexplored.

Several limitations define the scope of this systematic review. The search focused exclusively on English-language publications, which introduces a potential language bias. This scope may exclude relevant evidence from other linguistic contexts such as Chinese, Korean, or Japanese. The study is also restricted to the time window from December 1, 2022, to July 31, 2025. While this timeframe captures the emergence of Gen AI following the release of ChatGPT, it may not reflect the most recent developments in this rapidly evolving field. Furthermore, significant measurement heterogeneity exists across the 29 included studies. The use of diverse assessment instruments and outcome measures prevents direct comparability and precludes a meta-analysis. The evidence base also comprises varied research designs and includes articles from conference proceedings. These sources may not always match the rigorous peer-review standards of journal articles. Consequently, the generalization of the identified benefits and challenges requires caution. These findings should be interpreted within the context of these non-comparable methodologies and proceedings.

Author contributions: YZ: conceptualization, methodology, formal analysis, data curation, writing – original draft, visualization, writing – review & editing; **SHBH:** supervision, validation, writing – review & editing; **ZZ:** supervision, validation, writing – review & editing. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study did not require ethical approval from an ethics committee, as it is a systematic review focusing exclusively on previously published research articles and involved no data collection from human or animal subjects.

AI statement: The authors stated they used Grammarly to assist with grammar and spelling checks.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information, 16*(4), Article 296. https://doi.org/10.3390/info16040296

Bramer, W. M., De Jonge, G. B., Rethlefsen, M. L., Mast, F., & Kleijnen, J. (2018). A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association, 106*(4), 531-541. https://doi.org/10.5195/jmla.2018.283

Chin, J. H., Lee, S., Ashraf, M., Zago, M., Xie, Y., Wolfgram, E. A., Yeh, T., & Kim, P. (2024). Young children's creative storytelling with ChatGPT vs. parent: Comparing interactive styles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-7). ACM. https://doi.org/10.1145/3613905.3650770

Copple, C., & Bredekamp, S. (Eds). (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (3rd ed.). National Association for the Education of Young Children.

Daniel, K., Msambwa, M. M., & Wen, Z. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education, 60*(1), Article e70036. https://doi.org/10.1111/ejed.70036

Dietz Smith, G., Prasad, S., Davidson, M. J., Findlater, L., & Shapiro, R. B. (2024). Contextq: Generated questions to support meaningful parent-child dialogue while co-reading. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (pp. 408-423). ACM. https://doi.org/10.1145/3628516.3655809

Edwards, C. P., & Gandini, L. (2018). The Reggio Emilia approach to early childhood education. In J. L. Roopnarine, J. E. Johnson, S. Flannery Quinn, & M. M. Patte (Eds.), *Handbook of international perspectives on early childhood education* (pp. 365-378). Routledge. https://doi.org/10.4324/9781315562193-26

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is Chat GPT a blessing or a curse? *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1166682

He, K., Liu, X., Xu, Y., Bustamante, A. S., & Warschauer, M. (2025a). "Carlitos the curious caterpillar": Exploring teacher-AI co-creation of culturally responsive educational materials for young learners. In *Proceedings of the 24th Interaction Design and Children* (pp. 236-254). ACM. https://doi.org/10.1145/3713043.3727056

He, K., Gastón-Panthaki, A., Zhuo, D., Munsey, J., Zhang, M., & Warschauer, M. (2025b). StoryPal: Supporting Young children's dialogic reading with large language models. In *Proceedings of the 24th Interaction Design and Children* (pp. 494-511). https://doi.org/10.1145/3713043.3728867

Ho, H.-R., Kargeti, N., Liu, Z., & Mutlu, B. (2025). SET-PAiREd: Designing for parental involvement in learning with an AI-assisted educational robot. In

Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (pp. 1-20). ACM. https://doi.org/10.1145/3706598.3713330

Lee, J., Yoon, S., Lee, K., Jeong, E., Cho, J.-E., Park, W., Yim, D., & Hwang, I. (2024). Open sesame? Open salami! Personalizing vocabulary assessment-intervention for children via pervasive profiling and bespoke storybook generation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-32). ACM. https://doi.org/10.1145/3613904.3642580

Lee, Y., Kim, T. S., Kim, S., Yun, Y., & Kim, J. (2023). Dapie: Interactive step-by-step explanatory dialogues to answer children's why and how questions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-22). ACM. https://doi.org/10.1145/3544548.3581369

Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. *npj Science of Learning*, *10*, Article 29. https://doi.org/10.1038/s41539-025-00320-7

Li, Z., Thomas, T., Yu, C.-L., & Xu, Y. (2024). 'I said knight, not night!': Children's communication breakdowns and repairs with AI versus human partners. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (pp. 781-788). ACM. https://doi.org/10.1145/3628516.3659394

Lillard, A. S. (2017). *Montessori: The science behind the genius* (3rd ed.). Oxford University Press.

Liu, D., Zhou, H., & An, P. (2024). 'When he feels cold, he goes to the seahorse'—Blending generative AI into multimaterial storymaking for family expressive arts therapy. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-21). ACM. https://doi.org/10.1145/3613904.3642852

Liu, X., Xing, Y., Chen, Y., Pei, W., Choi, J., & Xu, Y. (2025). Children's narrative participation in story co-creation with AI vs. Human partners: Response styles and narrative features. In *Proceedings of the 24th Interaction Design and Children* (pp. 803-809). ACM. https://doi.org/10.1145/3713043.3731486

Liu, Y., Zha, S., Zhang, Y., Wang, Y., Zhang, Y., Xin, Q., Nie, L. Y., Zhang, C., & Xu, Y. (2025). BrickSmart: Leveraging generative AI to support children's spatial language learning in family block play. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-19). ACM. https://doi.org/10.1145/3706598.3714212

Lu, R.-S., Lin, H.-C. K., Yang, Y.-C., & Chen, Y.-P. (2024). Integrating urban mining concepts through ai-generated storytelling and visuals: Advancing sustainability education in early childhood. *Sustainability, 16*(24), Article 11304. https://doi.org/10.3390/su162411304

Lyu, Y., Liu, D., An, P., Tong, X., Zhang, H., Katsuragawa, K., & Zhao, J. (2024). EMooly: Supporting autistic children in collaborative social-emotional learning with caregiver participation through interactive AI-infused and AR activities. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (pp. 1-36). ACM. https://doi.org/10.1145/3699738

Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. *Trends in Higher Education, 4*(2), Article 17. https://doi.org/10.3390/higheredu4020017

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Perifanou, M., & Economides, A. A. (2025). Collaborative uses of genAI tools in project-based learning. *Education Sciences, 15*(3), Article 354. https://doi.org/10.3390/educsci15030354

Rakap, S. (2024). Chatting with GPT: Enhancing individualized education program goal development for novice special education teachers. *Journal of Special Education Technology, 39*(3), 339-348. https://doi.org/10.1177/01626434231211295

Rakap, S., & Balikci, S. (2024). Enhancing IEP Goal development for preschoolers with autism: A preliminary study on ChatGPT integration. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06343-0

Seiradakis, E. V. (2023). Unpacking experts' opinions on ChatGPT potential assistive roles and risks in early childhood special education. In *Proceedings of the International Conference on New Media Pedagogy* (pp. 380-392). Springer. https://doi.org/10.1007/978-3-031-63235-8_25

Sharma, R. C., & Panja, S. K. (2025). Addressing academic dishonesty in higher education: A systematic review of generative AI's impact. *Open Praxis, 17*(2), 251-269. https://doi.org/10.55982/openpraxis.17.2.820

Shen, J. J., King Chen, J., Findlater, L., & Dietz Smith, G. (2025). eaSEL: Promoting social-emotional learning and parent-child interaction through AI-mediated content consumption. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-18). ACM. https://doi.org/10.1145/3706598.3713405

Su, J., & Yang, W. (2024). Powerful or mediocre? Kindergarten teachers' perspectives on using ChatGPT in early childhood education. *Interactive Learning Environments, 32*(10), 6496-6508. https://doi.org/10.1080/10494820.2023.2266490

Sun, M., Yan, R., & Wen, R. (2025). Generative AI in Chinese early childhood education: Teachers' usage patterns, perceptions, and factors influencing pedagogical applications. *International Journal of Teacher Education and Professional Development, 8*(1), 1-23. IEEE. https://doi.org/10.4018/IJTEPD.382379

Sun, S., Wang, Z., Zhang, J., Zhao, X., & Qiao, X. (2024). StoryChat: An interactive storytelling system for engaging with agent characters in a story. In *Proceedings of the 2024 17th International Symposium on Computational Intelligence and Design* (pp. 213-217). https://doi.org/10.1109/ISCID63852.2024.00055

Sun, Y., Chen, J., Yao, B., Liu, J., Wang, D., Ma, X., Lu, Y., Xu, Y., & He, L. (2024). Exploring parent's needs for children-centered AI to support preschoolers' interactive storytelling and reading activities. In *Proceedings of the ACM on Human-Computer Interaction* (pp. 1-25). ACM. https://doi.org/10.1145/3687035

United Nations Children's Fund. (2023). *Early childhood development: UNICEF vision for every child*. UNICEF. https://doi.org/10.18356/9789213585429

Vella, K., Dobson, M., & Brereton, M. (2025). 'Hello, Mr tree': Toying with playful conversational AI in the early years. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-6). ACM. https://doi.org/10.1145/3706599.3719878

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, Z., Pan, J., Jin, D., Zhang, J., Cao, J., Zhang, C., Li, Z., Hansen, P., Zhao, Y., Sun, S., & Qiao, X. (2025). CharacterCritique: Supporting children's development of critical thinking through multi-agent interaction in story reading. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-21). ACM. https://doi.org/10.1145/3706598.3713602

Wong, J. M. S., Liu, P. H. C., & Yip, C. W. H. (2024). Navigating the digital shift: Early childhood educators' attitudes towards generative artificial intelligence and emerging technologies. In *Proceedings of the Asia Education Technology Symposium* (pp. 87-97). Springer. https://doi.org/10.1007/978-981-96-4952-5_6

Xu, Y., Thomas, T., Yu, C.-L., & Pan, E. Z. (2025). What makes children perceive or not perceive minds in generative AI? *Computers in Human Behavior: Artificial Humans, 4*, Article 100135. https://doi.org/10.1016/j.chbah.2025.100135

Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour, 8*(10), 1839-1850. https://doi.org/10.1038/s41562-024-02004-5

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*, Article 21. https://doi.org/10.1186/s41239-024-00453-6

Zhang, C., Liu, X., Ziska, K., Jeon, S., Yu, C.-L., & Xu, Y. (2024). Mathemyths: Leveraging large language models to teach mathematical language through Child-AI co-creative storytelling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1-23). ACM. https://doi.org/10.1145/3613904.3642647

Zhang, Y. L. (2025). Predicting teachers' intentions for AIGC integration in preschool education: A hybrid SEM-ANN approach. *Journal of Information Technology Education: Research, 24*, Article 016. https://doi.org/10.28945/5502

Zhao, Y., Li, J., & Shu, S. (2024). DailyPhysics: Fostering physics concept exploration in children through a tangible AI storytelling approach. In *Proceedings of the 2024 17th International Symposium on Computational Intelligence and Design* (pp. 231-234). IEEE. https://doi.org/10.1109/ISCID63852.2024.00059

Zhou, H., Nikolova, A., & An, P. (2024). 'My lollipop dropped ...'–Probing design opportunities for SEL agents through children's peer co-creation of social-emotional stories. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-8). ACM. https://doi.org/10.1145/3613905.3651867

APPENDIX A

Table A1. Benefits of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
Enhanced learning outcomes	Language & expression	Lu et al. (2024)	J	Mixed methods	N=60 children (5-6 yrs)	Taiwan, kindergarten	Self-report	Medium	Cohen's d=0.52-0.91 (4 outcomes: comprehension, oral expression, picture description, creation)
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002
		J. Lee et al. (2024)	C	Mixed methods	N=9 families (ages 50-71 months)	South Korea, home-based	Mixed	High	Interest correlations: ρ=.69-.96; Interest-acquisition: ρ=.42; Word-type correlations: r=.49-.59; Acquisition: 64% vs. 39%
	Comprehension & cognition	Zhao et al. (2024)	C	Mixed methods	N=8 children (4-7 yrs)	China, library	Mixed	Medium	Physics concept accuracy improvement 27.5% (pre-post)
		Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1, 31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1,18)=8.308, p=.010 (AI advantage); AI vs. human F(1, 31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs. Human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
	Critical thinking	Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking: t=2.058-4.897, p=.000-.048 (11/14 subskills significant). Giggle Gauge: M=3.50-4.63. Interaction experience: M=3.75-4.50. Conversation: Exp vs. control duration 26.31/9.38 min; rounds 21.81/30.63; topics 8.06/3.25.
	Social-emotional	Lyu et al.(2024)	C	RCT	N=24 autistic children (M=6)	China, special ed centers	Mixed	Low	Emotion recognition: EMooly (Pre M=4.83, Post M=6.33) vs. Baseline (Pre M=5.41, Post M=5.00), t=2.634, p=.015
		Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
		S. Sun et al. (2024)	C	Mixed methods	N=12 dyads (5-7 yrs)	China, library	Mixed	Medium	Not reported
	Creativity	Lu et al. (2024)	J	Mixed methods	N=60 children (5-6 yrs)	Taiwan, kindergarten	Self-report	Medium	Cohen's d=0.52-0.91 (4 outcomes: comprehension, oral expression, picture description, creation)
		Su and Yang (2024)	J	Qualitative interviews	N=10 teachers	Hong Kong, kindergartens	Self-report	High	N/A (qual)
Personalized & adaptive learning	Personalized content	J. Lee et al. (2024)	C	Mixed methods	N=9 families (ages 50-71 months)	South Korea, home-based	Mixed	High	Interest correlations: ρ=.69-.96; Interest-acquisition: ρ=.42; Word-type correlations: r=.49-.59; Acquisition: 64% vs. 39%
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002

Table A1 (Continued). Benefits of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
	Adaptive scaffolding	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote, home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002
		Lyu et al. (2024)	C	RCT	N=24 autistic children (M=6)	China, special ed centers	Mixed	Low	Emotion recognition: EMooly (Pre M=4.83, Post M=6.33) vs. Baseline (Pre M=5.41, Post M=5.00), t=2.634, p=.015
		He et al. (2025b)	C	Mixed methods	N=23 children (4-7 yrs) and parents	USA, community center + lab	Mixed	Medium	Not reported
	Differentiated support	Seiradakis (2023)	C	Qualitative interviews	N=6 ECSE experts	Greece, inclusive kindergarten	Self-report	High	N/A (qual)
		Rakap and Balikci (2024)	J	RCT	N=30 teachers	USA, public schools	Objective	Medium	IEP quality: 9.3 vs. 7.12 (control), p<.001
		He et al. (2025b)	C	Mixed methods	N=23 children (4-7 yrs) and parents	USA, community center + lab	Mixed	Medium	Not reported
Enhanced engagement	High engagement	He et al. (2025b)	C	Mixed methods	N=23 children (4-7 yrs) and parents	USA, community center + lab	Mixed	Medium	Not reported
		C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1, 31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1, 18)=8.308, p=.010 (AI advantage); AI vs. Human F(1, 31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs. Human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
		Wong et al. (2024)	BC	Mixed methods	N=97 ECE practitioners	Hong Kong, ECE program	Self-report	Medium	Not reported
	Interest & enjoyment	Chin et al. (2024)	C	Mixed methods	N=8 dyads (children aged 5-6)	USA, lab-based	Mixed	Medium	Conversational turn: Parent M=69.50 vs ChatGPT M=10.88; Positive feedback ratio: Parent M=.07 vs. ChatGPT M=.91
		Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote, home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		S. Sun et al. (2024)	C	Mixed methods	N=12 dyads (5-7 yrs)	China, library	Mixed	Medium	Not reported
	Attention & focus	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote, home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		Y. Sun et al. (2024)	C	Qualitative interviews	N=17 parents	China, home-based	Self-report	High	N/A (qual)
Parent support & interaction	Reduced burden	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote, home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85

Table A1 (Continued). Benefits of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
		Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote, home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
		Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking: t=2.058-4.897, p=.000-.048 (11/14 subskills significant). Giggle Gauge: M=3.50-4.63. Interaction experience: M=3.75-4.50. Conversation: Exp vs. Control duration 26.31/9.38 min; rounds 21.81/30.63; topics 8.06/3.25.
	Role transformation	Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking: t=2.058-4.897, p=.000-.048 (11/14 subskills significant). Giggle Gauge: M=3.50-4.63. Interaction experience: M=3.75-4.50. Conversation: Exp vs. Control duration 26.31/9.38 min; rounds 21.81/30.63; topics 8.06/3.25.
		Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
	Enhanced interaction	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote, home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		D. Liu et al. (2024)	C	Qualitative empirical	N=18 (7 families: 8 children aged 4-7, 10 parents) + 4 therapists	China, therapy setting	Self-report	High	N/A (qual)
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002
	Parent capability	Ho et al. (2025)	C	Mixed methods	N=20 dyads (3-5 yrs)	USA, home-based	Self-report	Medium	Pre vs. Post parent perception (7-pt scale): Literacy L1 (M=3.51 vs. 3.59), L2 (M=2.75 vs. 3.12*), L3 (M=4.08 vs. 4.30), L4 (M=3.75 vs. 4.17); Math L1 (M=4.46 vs. 4.62), L2 (M=4.08 vs. 4.26), L3 (M=3.33 vs. 3.68**), L4 (M=2.23 vs. 3.30**)
		Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002
Teacher efficiency	Content creation	Wong et al. (2024)	BC	Mixed methods	N=97 ECE practitioners	Hong Kong, ECE program	Self-report	Medium	Not reported
		He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (qual)

Table A1 (Continued). Benefits of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
	Time savings	Rakap (2024)	J	RCT	N=22 teachers	USA, public schools	Objective	Medium	Group effect (ChatGPT × training): η^2=.80; IEP quality: 9.56 vs. 6.78; Goal Development time: 15.6 vs. 25.5 min; r=-.53 (time-quality)
		He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (Qual)
	IEP quality	Rakap (2024)	J	RCT	N=22 teachers	USA, public schools	Objective	Medium	Group effect (ChatGPT × training): η^2=.80; IEP quality: 9.56 vs. 6.78; Goal Development time: 15.6 vs. 25.5 min; r=-.53 (time-quality)
		Rakap and Balikci (2024)	J	RCT	N=30 teachers	USA, public schools	Objective	Medium	IEP quality: 9.3 vs. 7.12 (control) , p<.001
	Cultural responsiveness	He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (qual)
	Admin efficiency	Seiradakis (2023)	C	Qualitative interviews	N=6 ECSE experts	Greece, inclusive kindergarten	Self-report	High	N/A (qual)
		M. Sun et al. (2025)	J	Mixed methods	N=10 teachers	China (5 provinces), kindergarten	Self-report	Medium	Not reported

Note. PT: Publication type; C: Conference; J: Journal; BC: Book chapter

APPENDIX B

Table B1. Challenges of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
Technical limitations	Speech recognition	Vella et al. (2025)	C	Qualitative observational study	N=15 children (3-5 yrs)	Australia, early learning center	Self-report	High	N/A (qual)
		Li et al. (2024)	C	RCT	N=71 children (4-8 yrs)	USA, libraries	Mixed	Medium	GLMM: Breakdowns β=-0.90, p<.001 (condition), β=-0.40, p=.002 (age); Repairs β=3.02, p<.001 (condition), β=0.10, p=.028 (mind perception); no interaction effect
	Response latency	Zhou et al. (2024)	C	Qualitative case study	N=8 children (4-6 yrs)	China, kindergarten	Self-report	Medium	N/A (qual)
	System stability	C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1, 31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1, 18)=8.308, p=.010 (AI advantage); AI vs. Human F(1,31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs. Human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
		Vella et al. (2025)	C	Qualitative observational study	N=15 children (3-5 yrs)	Australia, early learning centre	Self-report	High	N/A (Qual)
	Communication breakdown	Li et al. (2024)	C	RCT	N=71 children (4-8 yrs)	USA, libraries	Mixed	Medium	GLMM: Breakdowns β=-0.90, p < .001 (condition), β=-0.40, p=.002 (age); Repairs β=3.02, p<.001 (condition), β=0.10, p=.028 (mind perception); no interaction effect
	Network issues	Su and Yang (2024)	J	Qualitative interviews	N=10 teachers	Hong Kong, kindergartens	Self-report	High	N/A (qual)
Content quality	Language complexity	Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
		Chin et al. (2024)	C	Mixed methods	N=8 dyads (children aged 5-6)	USA, lab-based	Mixed	Medium	Conversational turns: Parent M=69.50 vs. ChatGPT M=10.88; Positive feedback ratio: Parent M=.07 vs. ChatGPT M=.91
	Inappropriate content	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking: t=2.058-4.897, p=.000-.048 (11/14 subskills significant). Giggle Gauge: M=3.50-4.63. Interaction experience: M=3.75-4.50. Conversation: Exp vs. control duration 26.31/9.38 min; rounds 21.81/30.63; topics 8.06/3.25.
		M. Sun et al. (2025)	J	Mixed methods	N=10 teachers	China (5 provinces), kindergartens	Self-report	Medium	Not reported
		Su and Yang (2024)	J	Qualitative interviews	N=10 teachers	Hong Kong, kindergartens	Self-report	High	N/A (qual)

Table B1 (Continued). Challenges of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
	Output issues	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1,31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1,18)=8.308, p=.010 (AI advantage); AI vs. Human F(1,31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs. Human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
Accuracy & bias	Hallucination	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1,31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1,18)=8.308, p=.010 (AI advantage); AI vs. Human F(1,31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs. Human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
	Accuracy issues	He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (qual)
		Seiradakis (2023)	C	Qualitative interviews	N=6 ECSE experts	Greece, inclusive kindergarten	Self-report	High	N/A (qual)
	Bias propagation	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (Qual)
	Error propagation	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
Context & multimodal	No visual input	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		Chin et al. (2024)	C	Mixed methods	N=8 dyads (children aged 5-6)	USA, lab-based	Mixed	Medium	Conversational turns: Parent M=69.50 vs. ChatGPT M=10.88; Positive feedback ratio: Parent M=.07 vs. ChatGPT M=.91

Table B1 (Continued). Challenges of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
	Non-verbal missing	Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking t-tests: t=2.058-4.897, p=.000-.048 for significant comparisons (11/14 subskills). Giggle Gauge: M=3.50-4.63 (5-point scale). Interaction experience: M=3.75-4.50. Conversation duration: Exp M=26.31 min vs. Control M=9.38 min. Conversation rounds: Exp M=21.81 min vs. Control M =30.63 min. Topics discussed: Exp M=8.06 vs. Control M=3.25
		Xu et al. (2025)	J	RCT	N=119 children (4-8 yrs)	USA, libraries	Mixed	Medium	Partial η^2: Agency .16***, Experience .60***, Response rate .07*, Substantive rate .11**. Regression b: Present Human→Agency .74***, →Experience 1.94***; Hidden Human→Experience .37*. R^2 = .23-.64. Pearson r: readability-experience .43***, articulation-experience -.45***. Mediation β: articulation .16-.22**, readability .25*.
	Context understanding	Y. Lee et al. (2023)	C	Mixed methods	N=16 children (5-7 yrs); 4 child education experts; crowd workers	South Korea, remote/home	Mixed	High	DAPIE vs. Baseline (M): 7.43/5.13 (understanding); 11.9%/29% (distraction); 3.38/2.31 (reuse); 3.69/2.75 (better teacher)
		Vella et al. (2025)	C	Qualitative observational study	N=15 children (3-5 yrs)	Australia, early learning centre	Self-report	High	N/A (Qual)
	Lack of personalization	Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
Safety & implementation	Content safety	Shen et al. (2025)	C	Mixed methods	N=20 dyads (5-8 yrs)	USA, remote	Mixed	High	Affect words: Cliff's d=0.27 (small), p=.02; eaSEL M=0.058 vs. control M=0.040
		J. Lee et al. (2024)	C	Mixed methods	N=9 families (ages 50-71 months)	South Korea, home-based	Mixed	High	Interest correlations: ρ=.69-.96; Interest-acquisition: ρ=.42; Word-type correlations: r=.49-.59; Acquisition: 64% vs. 39%
	Privacy & ethics	J. Lee et al. (2024)	C	Mixed methods	N=9 families (ages 50-71 months)	South Korea, home-based	Mixed	High	Interest correlations: ρ=.69-.96; Interest-acquisition: ρ=.42; Word-type correlations: r=.49-.59; Acquisition: 64% vs. 39%
		C. Zhang et al. (2024)	C	Mixed methods	N=35 children (4-8 yrs)	USA, libraries	Mixed	High	Learning: Pre-post F(1, 31)=17.009, p<.001 (M: 16.69→18.97); Definition subscale F(1, 18)=8.308, p=.010 (AI advantage); AI vs. Human F(1, 31)=0.012, p=.912. Story creation: Ideas p=.128; Elaborations p=.461. System: 87% recognition, 80% sensible; AI M=2.59 vs.

Table B1 (Continued). Challenges of Gen AI in ECE

Theme	Sub-theme	Studies	PT	Design	Sample	Context	Measure	MMAT	Effect size
									human M=1.94. Enjoyment: p=.184-.743 (4 measures: enjoyment, comfort, reading partner, interest)
		Vella et al. (2025)	C	Qualitative observational study	N=15 children (3-5 yrs)	Australia, early learning centre	Self-report	High	N/A (qual)
		He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (qual)
	Training & resources	Wong et al. (2024)	BC	Mixed methods	N=97 ECE practitioners	Hong Kong, ECE program	Self-report	Medium	Not reported
		Seiradakis (2023)	C	Qualitative interviews	N=6 ECSE experts	Greece, inclusive kindergartens	Self-report	High	N/A (qual)
		Su and Yang (2024)	J	Qualitative interviews	N=10 teachers	Hong Kong, kindergartens	Self-report	High	N/A (qual)
	Adult intervention	Y. Sun et al. (2024)	C	Qualitative interviews	N=17 parents	China, home-based	Self-report	High	N/A (Qual)
		Ho et al. (2025)	C	Mixed methods	N=20 dyads (3-5 yrs)	USA, home-based	Self-report	Medium	Pre vs. Post parent perception (7-pt scale): Literacy L1 (M=3.51 vs. 3.59), L2 (M=2.75 vs. 3.12*), L3 (M=4.08 vs. 4.30), L4 (M=3.75 vs. 4.17); Math L1 (M=4.46 vs. 4.62), L2 (M=4.08 vs. 4.26), L3 (M=3.33 vs. 3.68**), L4 (M=2.23 vs. 3.30**)
		Dietz Smith et al. (2024)	C	Mixed methods	N=12 dyads (4-6 yrs); 4 educators; 8 families (iterative testing)	USA, remote/home	Mixed	High	Parent turns: M=30.8 vs. 13.4, p=.002, r=0.88; Child turns: M=25.1 vs. 10.6, p=.004, r=0.85
		D. Liu et al. (2024)	C	Qualitative empirical	N=18 (7 families: 8 children aged 4-7, 10 parents) + 4 therapists	China, therapy setting	Self-report	High	N/A (Qual)
	Cognitive load	Wang et al. (2025)	C	Mixed methods	N=32 dyads (5-6 yrs)	China, library	Mixed	High	Critical thinking: t=2.058-4.897, p=.000-.048 (11/14 subskills significant). Giggle Gauge: M=3.50-4.63. Interaction experience: M=3.75-4.50. Conversation: Exp vs. Control duration 26.31/9.38 min; rounds 21.81/30.63; topics 8.06/3.25
		Y. Liu et al. (2025)	C	Mixed methods	N=24 dyads (6-8 yrs)	China, lab-based	Mixed	High	Spatial language improvement: 49.0% vs. 18.23%, U=127.0, p=.002
		He et al. (2025a)	C	Qualitative case study	N=5 teachers	USA, Latine school district	Self-report	High	N/A (qual)
		M. Sun et al. (2025)	J	Mixed methods	N=10 teachers	China (5 provinces), kindergartens	Self-report	Medium	Not reported

Note. PT: Publication type; C: Conference; J: Journal; BC: Book chapter