



### Rocio Garcia-Peinado

Childhood

Autonomous University of Madrid

Abstract: This article examines the use of artificial intelligence (AI) in early childhood education from a rights-based perspective, drawing on a critical interpretive synthesis (CIS) of the literature published between 2019 and 2025. A typology of four uses in early childhood —Tutor, Tool, Companion, and Tracker (THCR)— is proposed and each category is mapped against the core principles of the United Nations Convention on the Rights of the Child: privacy, non-discrimination, best interests, and participation. The contribution includes: (a) a risk-safeguard matrix differentiated by type of AI; (b) a logic model and theory of change for care-centered implementations; and (c) the SAFE LEARN checklist (Safety by design, Agency/assent, Fairness, Explainability, Learning alignment, Educator capacity, Accountability, Risk logging, Non-replacement of care). Implications for policy and practice are discussed for schools, administrations, and providers, emphasizing human mediation and verifiable equity as minimal conditions for acceptance. This work offers a pioneering framework that connects international normative principles with operational instruments, providing immediate guidance for the Education 2030 agenda and Sustainable Development Goals (SDGs) 4 and 16.

Keywords: AI ethics; Child rights; CIS; Early childhood; Equity; Surveillance; Theory of change.

**DOI:** https://doi.org/10.31757/euer.833

#### Introduction

The expansion of artificial intelligence (AI) in education represents one of the most significant technological transformations of the past decade. Its integration into early childhood education (ages 0–8)—through intelligent tutoring systems, generative assistants, social robots, and learning analytics—opens unprecedented opportunities for personalization, teacher support, and early detection of learning difficulties. At the same time, it raises profound questions about care, privacy, and equity in the early years of life, a critical stage for cognitive, socio-emotional, and moral development (Bronfenbrenner, 1979; Vygotsky, 1978; Noddings, 1985). AI promises personalization and teacher support, but it also introduces tensions with privacy, non-discrimination, and the pedagogical bond that sustains socio-emotional development (Miao, Holmes, 2021; McStay & Rosner, 2021; Saxena, 2024).

International organizations have emphasized that the adoption of AI must be evaluated from a child rights perspective. The United Nations Convention on the Rights of the Child (UN, 1989) establishes binding principles of privacy, non-discrimination, best interests, and participation, which acquire particular relevance in contexts of high vulnerability such as early childhood. In this line, UNESCO's Recommendation on the Ethics of Artificial Intelligence (Morandin-Ahuerma, 2023) and UNICEF's Policy Guidance on AI for Children (2021) call for safeguards such as meaningful human oversight, proportional transparency, accessible explainability, and systematic risk assessment.

Despite the proliferation of projects and discourses on educational AI, empirical evidence in early childhood remains limited. Optimistic narratives about AI's ability to personalize learning coexist with methodological gaps: the absence of subgroup equity metrics, the scarcity of longitudinal studies, and the lack of frameworks for evaluating socio-

emotional effects (Holstein & Doroudi, 2022). These gaps create uncertainty about the real benefits and the risks of transferring technologies designed for other contexts into the classroom, with the danger of amplifying inequalities and eroding the pedagogical bond. Key risks include algorithmic opacity, automation bias displacing professional teacher judgment, and the normalization of surveillance practices through tracking devices and affective computing (Parasuraman & Riley, 1997; Skitka, Mosier, & Burdick, 2000; McStay & Rosner, 2021). In early childhood, where learning is inseparable from care relationships and play as a means of development, these risks are heightened.

In this context, the article proposes an analytical and practical framework structured around three pillars: (i) child development as a relational and situated process; (ii) ethics of care and child rights as normative criteria; and (iii) international AI governance as a regulatory reference (Bronfenbrenner, 1979; Vygotsky, 1978; Noddings, 1985; NIST, 2023).

Based on a critical interpretive synthesis (2019–2025), the article presents four original contributions: the THCR typology, which organizes AI uses in early childhood education; a child rights matrix that translates the Convention into practical safeguards; the SAFE LEARN checklist, which sets minimum legitimacy indicators; and the GMMM-CARE Matrix, which complements SAFE LEARN with procedures for ongoing management and oversight. The analysis is grounded in three central assumptions: human mediation as a non-negotiable condition, equity as a demonstrable requirement through disaggregated metrics, and privacy as a right that demands proportionality in data collection. From these assumptions, four analytical propositions are derived that guide the results: the risks of substituting care, unmeasured inequity, disproportionate surveillance, and lack of explainability and structured risk management.

The central question is not whether AI can be integrated into early childhood education, but under what conditions it becomes legitimate and desirable. In this sense, the article argues that AI is only acceptable when it guarantees irreplaceable teacher mediation, demonstrates verifiable equity, and respects proportionality in the use of personal data. Thus, the debate shifts from technological enthusiasm toward a normative and pedagogical focus, aligned with the 2030 Agenda and SDGs 4 and 16, oriented to ensuring that AI strengthens—rather than undermines—children's rights and human care in early learning.

## **Conceptual Foundations**

The analysis of AI in early childhood education rests on a tripartite conceptual framework integrating: (i) child development theories that highlight the centrality of relationships and contexts; (ii) the ethics of care and child rights as normative criteria for educational quality; and (iii) international AI governance frameworks that provide standards for risk management and accountability. These three pillars not only guide the critical reading of the literature but also enable the translation of abstract principles into operational constructs applicable to educational practice and policy decisions.

## Child Development: Relational, Situated, and Context-Sensitive

Learning in early childhood occurs within interdependent ecological systems, where interactions among family, school, and community mediate both opportunities and risks (Bronfenbrenner, 1979). From a sociocultural perspective, higher psychological functions do not emerge in isolation but through processes of symbolic mediation and adult scaffolding, enabling children to internalize cultural tools and develop cognitive and emotional competencies (Vygotsky, 1978).

This ecological-sociocultural vision implies that any technology introduced in early childhood not only transforms isolated activities but also reshapes roles, time, and information flows within the school microsystem. Thus, the criterion of acceptability should not be technological novelty but the net contribution to the educational experience. Within this framework, the risk of replacing human interaction with algorithms is particularly problematic: in early childhood, the pedagogical relationship is not ancillary but the very core of learning and socio-emotional well-being (Johnson & Lester, 2016).

## Ethics of Care and Child Rights

The ethics of care, developed by Noddings (1985), emphasizes attention, empathy, and responsibility as normative criteria in education. Far from conceiving the classroom as a neutral space for transmitting information, this perspective holds that educational quality is measured by adults' ability to provide sensitive and protective relationships.

In parallel, the United Nations Convention on the Rights of the Child (UN, 1989) establishes a binding legal framework that recognizes children as rights-bearing subjects. Four principles directly project onto educational AI:

- Best interests: every technological decision must be justified in terms of the child's holistic well-being, not institutional efficiency or convenience.
- Non-discrimination: AI must not replicate or amplify social biases and must ensure fair treatment for all children, including those in vulnerable contexts.
- Privacy: the right to privacy requires limiting and strictly justifying data collection, avoiding disproportionate surveillance.
- Participation: even in early childhood, children have the right to express their views and have them considered, requiring mechanisms of age-appropriate consent and assent (Lemaignan, Newbutt, Rice, Daly, & Charisi, 2021).

This rights-based approach introduces a fundamental criterion: no instrumental benefit of AI justifies practices that erode trust, relational care, or children's progressive autonomy.

## AI Governance: International Frameworks and Standards

The governance of educational AI must align with international standards regulating its development and application. UNESCO's Recommendation on the Ethics of AI (Morandín-Ahuerma, 2023) underscores the need for meaningful human oversight, proportional transparency, comprehensible explainability, and impact assessments. UNICEF's Policy Guidance on AI for Children (2021) insists that AI must be designed *for and with* children, ensuring safety, fairness, and active participation.

A key contribution is the AI Risk Management Framework developed by NIST (2023), which organizes risk management into four dimensions:

- Govern: define policies for human oversight, roles and responsibilities, and establish a safe shutdown mechanism.
- Map: document the pedagogical purpose, data lifecycle, and potential harms, including data protection impact assessments.
- Measure: establish metrics for fairness, robustness, and explainability, with periodic bias audits and traceability through logs.
- Manage: activate mitigation plans, post-deployment monitoring, and protocols for responsible withdrawal when risk thresholds are exceeded (Tabassi, 2023).

These frameworks provide a shared language and practical tools for operationalizing ethical principles in educational and policy decisions.

## **Operational Constructs**

From the articulation of these three pillars, several key constructs emerge to guide subsequent analysis:

- Guarantee human mediation and avoid any form of care substitution.
- Develop evolving consent mechanisms recognizing children's participation.
- Apply proportionality in data collection and use.
- Verify algorithmic fairness through subgroup-specific validation.
- Ensure transparency and explainability actionable for teachers and families.
- Include psychosocial risk assessments and establish clear limits to surveillance.
- Monitor automation bias systematically and define explicit thresholds for teacher intervention.

In sum, these constructs condense ethical and pedagogical principles into practical criteria, facilitating the transition toward the analysis of results and the construction of the THCR typology. They operationalize rights and pedagogical principles into concrete metrics and safeguards, providing a solid basis for the results and typology development.

#### Methods

## Method: Critical Interpretive Synthesis and Evidence Mapping

This study adopts a Critical Interpretive Synthesis (CIS) complemented by evidence mapping. This strategy is particularly suitable for an emerging field such as AI in early childhood education, characterized by heterogeneous sources, the coexistence of normative frameworks and fragmented empirical studies, and the need to produce practical guidance beyond mere description (Dixon-Woods et al., 2006; Grant & Booth, 2009).

Unlike systematic reviews that seek exhaustiveness and quantitative comparability, CIS pursues conceptual coherence and decision-making utility, integrating empirical, normative, and policy literature. The methodological choice responds to two main reasons: (i) the scarcity of empirical studies specific to ages 0-8, which prevents a conventional meta-analysis; and (ii) the need to translate ethical principles into operational and verifiable constructs, requiring the combination of dispersed evidence with international regulatory frameworks.

Consistent with the qualitative tradition, the reflexivity of the authors is acknowledged as part of the interpretive process. Decisions regarding inclusion, coding, and synthesis are not conceived as neutral but influenced by the authors' research trajectory and normative commitment to centering care and child rights. This positionality is made explicit to provide transparency in the analytical process and strengthen the legitimacy of the findings.

## Research Questions and Scope

The analysis was guided by three central questions:

- 1. What risks and safeguards emerge in the different uses of AI in early childhood education?
- 2. How can child rights principles be translated into verifiable technical and pedagogical decisions, ensuring that international standards are concretely expressed in school practice?
- 3. What minimum metrics allow for the verification of fairness, transparency, and human oversight in early childhood contexts (ages 0–8)?

The scope was limited to school and para-curricular settings in early childhood education, explicitly excluding secondary and higher education, except when conceptual frameworks or metrics were transferable with clear justification.

## Search Strategy and Sources

The review was conducted between 2019 and 2025, including both academic literature and high-authority institutional documents. On the academic side, databases such as Scopus, Web of Science, and ERIC were searched for relevant titles, abstracts, and keywords. In parallel, grey literature from international organizations such as UNESCO, UNICEF, NIST, and WEF was incorporated, given their production of normative guidelines and governance frameworks on AI and child rights. Classic works in developmental theory and care ethics—Bronfenbrenner (1979), Vygotsky (1978), and Noddings (1985)—were also included to support the conceptual framework.

The search strategy combined terms related to the study population ("early childhood," preschool, kindergarten, "0-8"), technology designations ("artificial intelligence," "intelligent tutor\*," "learning analytics," "social robot\*," "generative AI"), and ethical-normative descriptors (ethic\*, privacy, equity, fairness, governance, "child rights"). Additional terms such as affective computing, biometric, and "evolving consent" were also incorporated to capture emerging and relevant literature. Overall, the strategy sought to balance breadth of coverage with pertinence to early childhood.

#### Inclusion and Exclusion Criteria

To ensure relevance, clear inclusion and exclusion criteria were established. Included studies focused on the 0-8 population, as well as research in other levels of education when transferability was explicitly justified. Only works with explicit ethical, legal, or pedagogical implications, published in peer-reviewed journals or issued by highauthority organizations such as UNESCO, UNICEF, or NIST, were considered. Texts in English and Spanish were admitted. Excluded were opinion pieces lacking empirical or normative grounding, commercial reports without explicit methodology, and clinical studies without educational application.

#### Selection and Extraction Process

Screening was carried out in two phases: first, titles and abstracts were reviewed for preliminary relevance; second, full texts of selected works were examined. From each source, homogeneous variables were extracted: type of technology and use cases, population and educational context, identified risks and benefits, proposed safeguards, and reported metrics or indicators. This extraction was guided by a shared template, and discrepancies among reviewers were resolved through discussion and triangulation with international reference frameworks (Morandín-Ahuerma, 2023; NIST, 2023).

### Synthesis and Category Construction

The synthesis employed a thematic analysis with mixed coding. Initially, deductive coding applied categories derived from child rights principles and AI governance frameworks. Subsequently, inductive coding identified emergent patterns across reviewed studies. This process led to the identification of four transversal ethical dilemmas—privacy vs. personalization, equity vs. bias, complementarity vs. substitution, and support vs. surveillance—as well as the construction of the THCR typology (Tutor, Tool, Companion, Tracker), which organizes the risks and benefits of different uses of AI in early childhood.

## Quality Assessment and Limitations

Source quality was evaluated based on authority and relevance to the topic. For narrative reviews, the SANRA scale was used as a formative reference to assess structure, justification, and argumentative solidity (Baethge, Goldbeck-Wood, & Mertens, 2019). For policy documents, clarity of principles, applicability of recommendations, and explicit accountability mechanisms were considered.

Limitations include the risk of selection and publication bias, partially mitigated by triangulation with international frameworks; the scarcity of longitudinal studies specific to early childhood; and terminological heterogeneity that complicates comparisons across studies. Nonetheless, the CIS approach allowed for conceptual coherence to be conferred to fragmented evidence and, most importantly, for the derivation of practical orientations with immediate relevance for schools, administrations, and providers.

#### Results

# Results: Ethical Dilemmas and THCR Typology

The thematic analysis derived from the Critical Interpretive Synthesis identified four transversal ethical dilemmas in the application of AI in early childhood education: privacy-personalization, equity-bias, complementaritysubstitution, and support-surveillance. These dilemmas are not mutually exclusive; rather, they intersect and generate tensions that must be addressed through verifiable safeguards. From these dilemmas, the THCR typology (Tutor, Tool, Companion, Tracker) was developed, providing an organizing framework for AI uses and guiding decision-making.

#### Privacy-Personalization

The promise of AI in early childhood education is often linked to the capacity to personalize learning trajectories. However, such personalization depends on the intensive collection of behavioral data and, increasingly, socioemotional signals captured by sensors and affective computing systems. In the 0-8 age group, the available evidence remains scarce and heterogeneous, raising serious doubts about the validity of current approaches (Franz, Goodwin, Rieder, Matheis, & Damiano, 2022). Literature warns of risks associated with expanding the data perimeter without pedagogical justification, reusing information for secondary purposes, and operating with pronounced opacity in storage and analysis processes (Smuha, 2021; Devillers & Cowie, 2023). Particular concern arises with affective computing applied to children, as it lacks robust cross-cultural validation and may generate systematic errors with negative psychosocial consequences such as anxiety or stigmatization (McStay & Rosner, 2021).

From a normative perspective, the principle of proportionality requires minimizing categories of data collected, prioritizing on-device processing, and offering comprehensible explanations for teachers and families. Consent should be conceived as an evolving process, with real options for revocation and childfriendly information (Lemaignan, Newbutt, Rice, Daly, & Charisi, 2021; Grace, Abel & Salen, 2023). Operational safeguards include clear limits on data retention, traceability of access, prohibition—or reinforced justification—of biometric and affective technologies, and submission to external privacy-bydesign audits.

The central risk is data overexploitation, directly undermining the right to privacy. The essential safeguard is strict data minimization and external audits, ensuring that systems never process biometric data without reinforced pedagogical justification.

### Equity-Bias

Although AI is often expected to expand access to educational opportunities, evidence shows that when systems are trained on unrepresentative data, they tend to reproduce or even amplify pre-existing inequalities. Such inequities may relate to language, disability, migrant status, or rural-urban contexts (Holstein & Doroudi, 2022). In these cases, socalled "personalization" may mask cultural or linguistic design flaws, undermining principles of inclusion and social justice (Leong & Zhang, 2025). Structural factors such as unequal technological infrastructure or insufficient teacher training can also limit equitable access to AI's promised benefits (Miao, & Holmes, 2021).

Normatively, the message is clear: equity cannot be presumed; it must be demonstrated. This requires subgroupdifferentiated evaluations, with specific mitigation plans where gaps are identified. Verifiable guidelines include testing system performance across diverse populations, incorporating co-design processes with marginalized communities, and establishing accessibility requirements such as offline modes or low-energy versions in public procurement (Miao, & Holmes, 2021).

The main risk is structural and cultural inequality reproduction due to biased datasets and infrastructural disparities. This risk affects the principle of non-discrimination, while the safeguard is subgroup validation combined with inclusive mitigation plans.

#### Complementarity-Substitution

Intelligent tutors and conversational agents are among the most widely disseminated AI applications in education. These systems provide adaptive feedback and may reduce teacher workload in some contexts. However, multiple studies caution that their potential must not be conflated with the ability to replace human mediation. In early childhood, affective presence and contextual judgment provided by teachers are irreplaceable for learning and socioemotional development (Johnson & Lester, 2016; Buarque, 2023; Kurian, 2023).

The risk of automation bias is particularly salient at this stage, as the tendency to delegate decisions to systems is heightened when interfaces do not make margins of error or uncertainty explicit (Parasuraman & Riley, 1997; Skitka, Mosier, & Burdick, 2000). Ethical acceptability therefore depends on significant human oversight with clearly defined intervention thresholds. Operational safeguards include explicit definitions of "high-impact decisions" (e.g., diagnoses or referrals), mandatory double teacher review in such cases, monitoring omission and commission errors linked to automation bias, strict time limits for exposure to social robots, and ensuring that algorithmic recommendations align with curricular and socio-emotional objectives.

The core risk is unintended substitution of human care through uncritical delegation to automated systems. This undermines the child's best interests, with the safeguard being explicit human intervention thresholds and mandatory teacher review for critical decisions.

### Support-Surveillance

Learning analytics tools and wearable devices are often justified by their potential to detect learning difficulties or special needs at an early stage. However, in early childhood these systems present a high risk of normalizing surveillance, constraining children's autonomy, spontaneity, and self-expression (McStay & Rosner, 2021; Lemaignan, Newbutt, Rice, Daly, & Charisi, 2021). The World Economic Forum (2024) has cautioned that without robust regulatory frameworks, such technologies may be interpreted by families and teachers in punitive ways, generating anxiety and undermining educational trust.

Here, the operational principle should be "nurture, not monitor". Intensive monitoring lacks pedagogical justification in early childhood, except in strictly necessary circumstances and always following prior impact assessments. Safeguards include child-specific impact assessments before deployment, prioritization of non-biometric and intermittent rather than continuous monitoring tools, pre- and post-implementation psychosocial evaluations, and participatory governance mechanisms with family involvement and accessible complaint channels.

The main risk is the normalization of intrusive surveillance practices, negatively affecting autonomy and selfexpression. This directly undermines the right to participation under the Convention on the Rights of the Child. The safeguard is mandatory child-specific impact assessments and non-biometric alternatives.

#### **Comparative Synthesis of Ethical Dilemmas**

The synthesis of the four ethical dilemmas—privacy-personalization, equity-bias, complementarity-substitution, and support-surveillance—shows that AI in early childhood can only be legitimate if it operates under verifiable conditions of proportionality, fairness, and human mediation.

- In privacy, the central risk is data overexploitation; the safeguard is strict data minimization, external audits, and evolving consent.
- In equity, the risk is inequality reproduction; the safeguard is subgroup validation and inclusive mitigation plans.
- In complementarity, the risk is erosion of human care; the safeguard is explicit thresholds for intervention and teacher review.
- In support–surveillance, the risk is intrusive monitoring normalization; the safeguard is impact assessments and non-biometric alternatives.

Taken together, these dilemmas warn that personalization without proportionality, equity without metrics, tutoring without human mediation, and analytics without limits are unacceptable from a child rights perspective. For greater clarity, Table X summarizes each dilemma's central risk, affected normative principle, and principal safeguard.

In order to provide a clear overview, Table 1 summarizes the four ethical dilemmas by aligning each with its central risk, the affected normative principle, and the main safeguard. This structured synthesis highlights both the common patterns and the specific conditions required for legitimacy in early childhood AI.

| Dilemma              | Central risk                  | Affected normative                               | Main safeguard              |  |
|----------------------|-------------------------------|--|-----------------------------|--|
|                      |                               | principle  |                             |  |
| Privacy-             | Excessive data collection and | Privacy and data                                 | Data minimization, external |  |
| Personalization      | opacity in its use            | protection                                       | audits, evolving consent    |  |
| <b>Equity-Bias</b>   | Reproduction of structural    | Non-discrimination                               | Subgroup validation,        |  |
|                      | and cultural inequalities     | cultural inequalities inclusive mitigation plans |                             |  |
| Complementarity-     | Uncritical delegation and     | Best interests of the                            | Explicit thresholds for     |  |
| Substitution         | replacement of teacher        | child  | human intervention, double  |  |
|                      | mediation                     |  | teacher review              |  |
| Support-Surveillance | Normalization of intrusive    | Participation and                                | Child-specific impact       |  |
|                      | monitoring practices          | autonomy   | assessments, non-biometric  |  |
|                      |                               |  | alternatives                |  |

Table 1. Comparative Synthesis of Ethical Dilemmas in Early Childhood Educational AI. Source: Authors' elaboration based on literature synthesis (2019–2025).

#### Rights and Safeguards Matrix. THCR Typology

Building on the synthesis of dilemmas and the THCR typology, this section translates the four guiding principles of the UN Convention on the Rights of the Child (CRC)—privacy, non-discrimination, best interests, and participation into operational safeguards applicable to tutors, tools, companions, and trackers. This rights-based matrix illustrates how abstract principles can be transformed into concrete obligations for educators, administrators, and providers.

#### Privacy

The right to privacy requires that children not be subjected to excessive data collection or continuous monitoring. For intelligent tutors, this means performance logs should be limited strictly to pedagogical purposes, with traceable and auditable access. In the case of generative tools, prompts must exclude sensitive information about children or classrooms. Robotic or conversational companions should avoid biometric or affective computing technologies, unless justified by exceptional pedagogical needs and backed by reinforced consent. For trackers and learning analytics, safeguards include child-specific impact assessments, strict data retention limits, and simple mechanisms for families to request deletion.

#### Non-discrimination

The principle of equal treatment requires that no child be excluded or disadvantaged due to language, disability, gender, or social context. For adaptive tutors, this entails validation across representative subgroups to ensure equitable system performance. For digital tools, it requires critical review of training corpora to ensure cultural and linguistic diversity. In robotic companions, interactions should accommodate dialectal and cultural variation, avoiding reinforcement of stereotypes. For trackers, equity demands periodic audits to confirm that monitoring does not introduce systematic biases against certain profiles or communities.

## Best interests of the child

All technological decisions must be assessed against the child's overall well-being, not solely efficiency or novelty. For tutors, high-impact decisions (e.g., diagnostics, referrals) should always undergo teacher review. For tools, alignment with curriculum and formative objectives is mandatory. For robotic companions, safeguards include limiting usage time and ensuring adult mediation. For trackers, deployment must only be justified when clear and proportionate pedagogical benefits are demonstrated, excluding experimental or punitive applications.

# **Participation**

Children are recognized as active subjects of rights, including in early childhood. For intelligent tutors, this means providing age-appropriate, understandable feedback that clarifies recommendations. For tools, participation involves opportunities for co-creation of materials with children and families. In the case of robotic companions, children should help define rules of use, expressing preferences and limits. For trackers, participation requires transparent information on what is being monitored and why, alongside evolving consent and simple revocation mechanisms.

In summary, the matrix shows that the four fundamental CRC rights translate into operational safeguards differentiated by AI category. Across all cases, the conditions of acceptability converge on three axes: protecting children's privacy, ensuring equitable and culturally inclusive treatment, and maintaining teacher mediation as the guarantor of both best interests and participation.

To complement the narrative matrix, Supplementary Table 1 provides brief illustrative examples of safeguards by child right and THCR category, offering a practical reference for educators, policymakers, and technology providers.

| Child Right           | Tutor (Adaptive          | Tool (Generative     | Companion (Social       | Tracker (Learning     |
|-----------------------|--------------------------|----------------------|-------------------------|-----------------------|
|                       | Systems)                 | Applications)        | Robots/Agents)          | Analytics /           |
|                       |                          |                      |                         | Wearables)            |
| Privacy               | Logs limited to          | Prompts must         | Avoid affective         | Child-specific data   |
|                       | learning outcomes;       | exclude personal or  | computing unless        | impact assessment;    |
|                       | traceable teacher access | sensitive child data | exceptionally justified | retention limits;     |
|                       | only                     |                      | with reinforced         | family-controlled     |
|                       |                          |                      | consent                 | deletion options      |
| Non-                  | Validation across        | Curation of training | Design of interactions  | Periodic audits to    |
| discrimination        | subgroups (language,     | corpora with         | inclusive of dialectal  | ensure no systematic  |
|                       | disability, socio-       | cultural and         | and cultural variation  | bias in monitoring    |
|                       | economic background)     | linguistic diversity |                         |                       |
| <b>Best Interests</b> | Teacher review           | Alignment of         | Usage limited to short  | Deployment only       |
|                       | required for diagnostics | generated content    | sessions; constant      | when benefits are     |
|                       | and high-impact          | with curriculum and  | adult mediation         | clear, proportionate, |
|                       | decisions                | socio-emotional      |                         | and educationally     |
|                       |                          | goals                |                         | justified             |
| Participation         | Age-appropriate          | Co-creation of       | Joint definition of     | Transparent           |
|                       | feedback explaining      | digital materials    | rules of use with       | explanation of        |
|                       | recommendations          | with children and    | children's preferences  | monitoring; evolving  |
|                       |                          | families             | and limits              | consent and simple    |
|                       |                          |                      |                         | revocation            |
|                       |                          |                      |                         | mechanisms            |

Supplementary Table 1. Illustrative Safeguards by Child Rights and THCR Categories. Source: Authors' elaboration based on CRC (UN, 1989), UNICEF (2021), and CIS review (2019–2025).

## Theory of Change and Logic Model

The integration of artificial intelligence into early childhood education can only be considered legitimate if it contributes to children's holistic well-being and does not erode the irreplaceable role of human care. On this basis, the proposed theory of change rests on a fundamental premise: AI generates value only when it strengthens pedagogical mediation and reduces inequalities, rather than displacing relationships and reproducing biases.

To advance in this direction, the implementation process requires a set of initial conditions. These include: teacher training in critical digital literacy; the establishment of public procurement guidelines setting minimum ethical criteria; regulated access to strictly necessary data; and the creation of co-design spaces with families and communities to

ensure cultural and contextual relevance. These conditions constitute the indispensable inputs for any responsible

deployment.

Building on these inputs, concrete actions are required to guide the use of AI in classrooms: the selection of

applications according to the THCR typology; the configuration of age-appropriate privacy protocols; the creation of

consent and assent mechanisms comprehensible to children and families; and the development of pilot projects with

iterative evaluation and feedback from all stakeholders. These actions ensure that innovation is not reduced to

technological introduction but translates into situated and transparent pedagogical practices.

The efforts result in immediate products that demonstrate ethical governance: responsibly adapted curricular materials;

dashboards understandable to teachers and families; clear human-in-the-loop protocols defining which decisions must

remain under professional authority; and auditable records of risks and incidents. More than technical outputs, these

products function as institutional guarantees that AI remains under human control and serves legitimate educational

purposes.

The consolidation of these products generates short- and medium-term outcomes: children participate more actively

in their learning thanks to accessible explanations; teachers reduce administrative burdens and gain improved

diagnoses, always mediated by professional judgment; and families develop greater trust in technology due to clear

information and effective revocation mechanisms.

In the long term, two key impacts are expected. First, a more equitable and inclusive learning experience capable of

addressing children's cultural, linguistic, and socioeconomic diversity. Second, the consolidation of a technological

culture that places children's rights at its center, protects privacy as a non-negotiable principle, and recognizes the

essential role of human care in child development.

The proposed logic model ultimately seeks to ensure that every stage of implementation is guided by the best interests

of the child.

**Discussions** 

**Implementation Guide: SAFE LEARN** 

For artificial intelligence in early childhood education to be acceptable, it is not enough to proclaim ethical principles.

They must be translated into verifiable and practical criteria guiding teachers, schools, administrations, and providers.

To this end, the SAFE LEARN framework is proposed—a guide that summarizes, across nine dimensions, the

minimum legitimacy conditions for the use of AI in early childhood.

The first dimension, Safety by design, requires that safety be embedded from the outset. This entails identifying threats before deployment, establishing safe-stop mechanisms, and ensuring that the system cannot generate foreseeable harm to children. The aim is not to correct errors after the fact but to anticipate and mitigate risks proactively.

The second dimension, Agency and Assent, acknowledges children's progressive autonomy. Mechanisms of age-appropriate consent must be developed, complemented—where feasible—by children's assent alongside parental authorization. Giving voice to children in technological decisions is not symbolic but an acknowledgment of their rights.

The third dimension, Fairness, demands proof of equity in system performance. Personalization does not guarantee inclusion; on the contrary, it may reinforce inequalities if training data or algorithms lack validation across diverse populations. Equity must therefore be demonstrated through subgroup testing, external audits, and mitigation plans.

The fourth dimension, Explainability, requires that system decisions be clearly understandable. Teachers and families must be able to grasp why a recommendation is issued, its limitations, and its level of certainty. Explainability is not about technical manuals but about actionable, accessible information.

The fifth dimension, Learning alignment, ensures that every AI application is pedagogically justified. Technology must serve curricular and socio-emotional goals, avoiding superficial uses driven by novelty or commercial pressure.

The sixth dimension, Educator capacity, highlights that AI does not replace teachers but complements them. For this complementarity to be effective, teachers need dedicated training, protected time, and collaborative reflection spaces. Without such capacity-building, AI risks increasing workloads and undermining professional judgment.

The seventh dimension, Accountability, places responsibility at the core. Providers must document the full lifecycle of data and models; administrations must establish mechanisms of external audit; and schools must provide accessible complaint channels. Accountability ensures transparency moves beyond rhetoric.

The eighth dimension, Risk logging, introduces the obligation to maintain a living, auditable record of incidents, errors, and system adjustments. This traceability supports learning from experience and continuous oversight. Regarding proportionality in data use, a data-to-need ratio  $\leq 1$  is recommended, ensuring no data are collected beyond what is strictly required for the declared educational purpose.

Finally, the ninth dimension, Non-replacement of care, sets a non-negotiable limit: AI must never substitute for human mediation in early childhood. The relationship between children and adults is irreplaceable, and any attempt at substitution is ethically unacceptable. Technology may complement, enrich, or facilitate—but never replace—care and pedagogy.

Together, SAFE LEARN converts abstract principles into operational criteria for determining which technologies are admissible and which must be discarded. Applied systematically, the guide protects the centrality of human care, ensures verifiable equity, and guarantees that technological innovation remains subordinate to children's rights.

| Dimension               | Primary safeguard with verifiable indicator  |  |  |
|-------------------------|--|--|--|
| Safety by design        | Integrated safe-stop mechanisms; tests confirming absence of foreseeable harm          |  |  |
| Agency / Assent         | Evolutionary consent and child assent; ≥80% comprehension in pilot validations         |  |  |
| Fairness                | Performance gaps across subgroups <5 percentage points                                 |  |  |
| Explainability          | ≥90% of teachers and families comprehend explanations provided                         |  |  |
| Learning alignment      | 100% of functionalities aligned with curriculum and socio-emotional goals              |  |  |
| Educator capacity       | ≥20 hours of annual teacher training specific to educational AI                        |  |  |
| Accountability          | External audits every 12 months, with publicly accessible reports                      |  |  |
| Risk logging            | Living, auditable incident log; notification of incidents ≤30 days; data-to-need ratio |  |  |
|                         | ≤1   |  |  |
| Non-replacement of care | 100% human review of high-impact decisions; ≤1 daily hour limit on social robot use    |  |  |

Table 2. Dimensions of the SAFE LEARN Framework and Verifiable Thresholds. Source: Author's elaboration based on Morandín-Ahuerma (2023) and NIST (2023).

## GMMM-CARE Matrix: Operational Governance of AI in Early Childhood

While the SAFE LEARN checklist defines minimum conditions of acceptability, its effective implementation requires a management framework capable of translating principles into daily practice. To this end, the GMMM-CARE Matrix is proposed, inspired by the NIST AI Risk Management Framework (2023) and adapted to the context of early childhood education. This matrix structures governance across four phases—Govern, Map, Measure, Manage—and integrates the transversal principle of CARE (Child-centred Accountability for Rights and Equity).

# Govern: policies and structures of control.

This phase emphasizes defining roles, responsibilities, and thresholds for human intervention. In early childhood education, schools should adopt explicit policies distinguishing between decisions that can be automated and those requiring mandatory teacher review. Establishing school-level ethical-pedagogical committees—with participation of educators, families, and community representatives—is recommended to oversee the use of technologies. This stage also includes safe-stop mechanisms ("kill switches") to suspend systems when unacceptable risks arise.

### Map: documentation of lifecycle and context.

The mapping stage requires transparent documentation of pedagogical purposes, data lifecycles, and model assumptions. In early childhood contexts, this entails specifying what information is collected, how it is processed, how long it is stored, and under what pedagogical justification. Vulnerable subpopulations (e.g., children with disabilities, bilingual learners, rural contexts) must be explicitly considered to prevent the reproduction of inequities. Data sheets and model cards adapted to the school context are recommended practices.

## Measure: metrics of equity, transparency, and oversight.

Measurement transforms ethical principles into verifiable indicators. For early childhood AI, these include: (i) equity across subgroups, with performance gaps reduced to <5 percentage points; (ii) monitoring of automation bias through omission and commission error rates; (iii) auditable logs ensuring traceability of decisions and access; and (iv) actionable transparency, with explanations accessible not only to experts but also to teachers and families.

## Manage: mitigation, response, and responsible withdrawal.

The management phase requires incident response plans, family notification protocols, and responsible system withdrawal when risk thresholds are exceeded. In early childhood education, management must go beyond technical fixes to include psychosocial impacts such as anxiety, loss of trust, or reduced autonomy. Protocols should therefore incorporate pre- and post-implementation evaluations of classroom climate and socio-emotional wellbeing.

## **CARE:** Child-centred Accountability for Rights and Equity.

Across all four phases, the CARE principle ensures that technical and institutional decisions remain subordinate to children's rights. CARE highlights that AI in education is not an end in itself but a means that must protect privacy, guarantee equity, and preserve the irreplaceable human care that underpins early learning.

In sum, the GMMM-CARE Matrix complements the SAFE LEARN framework by providing the operational scaffolding necessary for implementation, monitoring, and, when required, system withdrawal. Together, they constitute a governance ecosystem that integrates principles, metrics, and procedures, ensuring that AI in early childhood education is deployed ethically, transparently, and in alignment with the best interests of the child.

## **Policy and Practice Implications**

The integration of artificial intelligence in early childhood education cannot be approached as a purely technical process nor be left to the dynamics of the educational technology market. Given that fundamental rights are at stake, it is essential to build an ethical and pedagogical governance framework that involves schools, public administrations, and technology providers in differentiated yet complementary roles.

At the school level, the primary responsibility is to ensure that technology does not replace teacher mediation or erode relational bonds with children. Schools should explicitly define thresholds for human intervention, ensuring that all high-impact decisions remain under professional supervision. When technologies involve intensive data collection or monitoring procedures, schools must conduct impact assessments not only on data protection but also on the psychosocial effects on children. Continuous professional development is indispensable, enabling educators to critically interpret algorithmic recommendations while maintaining their role as agents of care and pedagogical judgment. Transparency toward families is equally crucial, requiring accessible information about what data are collected, for what purposes, and under what limits. In addition, schools should create spaces for participation and codesign, recognizing the voices of children and families in technology adoption.

At the level of public administrations, the challenge is to ensure that technological innovation is deployed under conditions of verifiable equity and public accountability. This requires regulatory frameworks and contractual mechanisms obligating providers to demonstrate equitable system performance, as well as public repositories of technical documentation—such as data sheets and model cards—to facilitate civic and academic oversight. Governments must also provide technical and pedagogical support to schools, particularly in vulnerable contexts, to prevent infrastructure and knowledge gaps from deepening inequalities. Furthermore, states should finance applied research, especially longitudinal studies and independent evaluations, that assess the impact of AI in early childhood education. This research function, coupled with inter-institutional accountability mechanisms, strengthens states' capacity to protect children's rights against emerging technological risks.

Technology providers cannot be considered neutral actors. They bear the obligation to ensure that products designed for early childhood education comply with verifiable ethical standards aligned with the Convention on the Rights of the Child and international AI governance frameworks. Providers should transparently document data and model lifecycles, publish evidence of equitable performance across diverse populations, and develop low-consumption versions that do not exclude schools with limited connectivity. Mechanisms for deletion, portability, and revocation of data must be accessible to families and understandable to children according to their age and maturity. Contracts and licenses should include explicit clauses on transparency, participation, and privacy safeguards, preventing schools from assuming risks alone.

Taken together, these implications underscore that the success of AI in early childhood education depends less on technical advances than on the capacity to construct a system of shared responsibility. Schools safeguard pedagogical mediation and family trust; governments regulate, supervise, and uphold systemic equity; and providers meet ethical and transparency requirements. Only through the articulated interaction of these three levels can AI become a legitimate resource, subordinated to the best interests of the child and oriented toward strengthening-not weakening—care and equity in early learning.

In line with the 2030 Agenda, these implications directly connect to SDG 4.7 (inclusive and quality education) and SDG 16 (responsible, just, and transparent institutions), positioning early childhood AI as both an educational and institutional challenge. Moving forward will also require interdisciplinary alliances among educators, technologists, and legal experts to generate robust, situated evidence capable of guiding policy and practice.

## Conclusion

### Limitations

This study, based on a Critical Interpretive Synthesis (CIS) and evidence mapping, offers a conceptual and operational framework to guide the use of artificial intelligence in early childhood education. Nonetheless, several limitations must be acknowledged, both in terms of the available evidence and the methodology adopted.

First, the body of literature specifically addressing AI in early childhood (0-8 years) remains limited and highly heterogeneous. Much of the existing research originates in primary and secondary education, requiring conceptual transfer that, while justified, entails uncertainty. This limitation calls for caution in interpreting findings and underscores the urgency of generating direct empirical studies in early learning contexts.

Second, the field is marked by substantial terminological and methodological diversity, complicating comparisons across studies and hindering the development of common indicators. Variability can be observed in definitions of AI categories as well as in equity metrics and measures of socio-emotional impact. This dispersion restricts generalizability and reinforces the need for consensual evaluation standards tailored to early childhood.

Third, the reliance on narrative synthesis introduces the possibility of selection and publication bias. Although triangulation with authoritative international frameworks (Lemaignan, Newbutt, Rice, Daly, & Charisi, 2021; NIST, 2023; Morandín-Ahuerma, 2023) mitigates this risk, it is likely that some relevant studies were excluded. Reflexivity on the part of the authors was acknowledged as part of the interpretive process, but this does not eliminate the limitations inherent in qualitative synthesis.

These constraints open avenues for an urgent research agenda. Priorities include:

- Case studies in classrooms for ages 0-8, documenting the situated impact of AI on pedagogy and teacherchild relationships.
- Pragmatic trials assessing not only academic performance but also socio-emotional development and the quality of relational care, including systematic monitoring of automation bias.
- Cost-effectiveness evaluations of non-biometric alternatives, such as human-AI co-design strategies and low-consumption tools, which may offer pedagogical benefits without compromising rights.
- Cross-cultural validation of affective computing and conversational AI systems, since untested use may amplify linguistic and cultural bias.

Longitudinal metrics capturing cumulative exposure effects of AI in early childhood, extending beyond immediate learning outcomes.

Additionally, the thresholds proposed in the SAFE LEARN checklist—including the ratio of data collected to strictly necessary data ≤1—should be regarded as provisional benchmarks. They are intended as initial references to guide practice and accountability, but their validity and applicability must be refined through empirical testing, particularly longitudinal and comparative research in early learning contexts.

In sum, advancing this research agenda is essential to ensure that AI in early childhood education develops not in a normative or empirical vacuum, but within a framework that prioritizes the best interests of the child and guarantees verifiable equity in early learning.

### **Conclusions**

The analysis conducted throughout this article demonstrates that artificial intelligence (AI) in early childhood education represents both a significant opportunity and a set of profound risks. While AI offers potential benefits such as early diagnosis, adaptive learning pathways, and support for teachers, these advantages are not neutral; they are conditioned by ethical and pedagogical safeguards. Without such safeguards, the risks of surveillance, bias, and erosion of human care can easily outweigh the benefits.

The study highlights three non-negotiable criteria of legitimacy for AI in early childhood: (i) irreplaceable human mediation, ensuring that educators remain at the core of pedagogical and socio-emotional processes; (ii) verifiable equity, which requires systematic disaggregation of performance metrics and proactive bias mitigation; and (iii) proportionality in data processing, ensuring that children's privacy is treated as an absolute right rather than as a tradeoff.

The proposed THCR typology (Tutor, Tool, Companion, Tracker) provides a practical framework for classifying AI uses in early childhood, while acknowledging their differentiated risks and corresponding safeguards. The SAFE LEARN checklist translates abstract principles into measurable obligations for design and implementation, and the GMMM-CARE Matrix offers governance procedures for continuous oversight and responsible withdrawal of systems when risks exceed thresholds.

Policy and practice implications emphasize that AI adoption must rest on shared responsibility: schools must protect relational care and transparency with families; public administrations must regulate and provide equitable infrastructure; and technology providers must comply with enforceable standards of ethics and accountability. In conclusion, the legitimacy of AI in early childhood education does not hinge on its technical capabilities but on its capacity to strengthen, rather than weaken, the relational, equitable, and protective nature of early learning. It can thus be affirmed programmatically that AI in early childhood education will only be legitimate if it becomes an ally of equity and human care, always under verifiable safeguards.

## **Innovative Proposals**

Beyond the safeguards and governance frameworks already discussed, it is essential to promote initiatives that consolidate the legitimacy of artificial intelligence (AI) in early childhood education. These initiatives can be grouped into three interdependent domains: institutional, pedagogical, and participatory.

#### Institutional innovations.

At the institutional level, the creation of ethical-pedagogical observatories for early childhood education is proposed. These observatories, with the participation of teachers, families, and community representatives, would oversee the use of technologies, establish annual goals, and publish publicly verifiable reports. Equally important is the development of public ethical licenses for educational technologies, conditioning market access on compliance with principles of transparency, data protection, and child participation. A further innovation involves community-based ethical impact evaluations, combining external audits with local feedback processes that give voice to those directly involved in schools.

#### Pedagogical innovations.

Three lines of innovation are particularly relevant. First, early literacy in AI, adapted to developmental stages, enabling children to progressively understand these tools and develop critical thinking. Second, human-AI co-teaching models, where roles are clearly defined, exposure times regulated, and iterative impact evaluations ensure that technology strengthens rather than weakens pedagogical mediation. Third, the design of bond-based systems, oriented towards reinforcing classroom climate, cooperation, and socio-emotional well-being, rather than replacing relational care.

### Participatory innovations.

A central challenge is to recognize children as active agents in digital environments. This requires experiences such as child-centered hackathons, where children, families, and teachers co-design contextually relevant AI solutions, turning innovation into a democratic and civic process. Equally, evolving consent mechanisms should be developed through pictograms, simple narratives, and accessible revocation options, transforming consent into a pedagogical process rather than a mere legal formality.

Taken together, these proposals aim to move beyond minimal regulation towards the active strengthening of a childcentered educational ecosystem. Their value lies in positioning children not only as subjects of protection but also as protagonists of innovation. By ensuring that AI in early childhood education is legitimate, equitable, and aligned with human rights, these initiatives point the way to a future where technological development is subordinated to the superior interest of the child.

## References

- Baethge, C., Goldbeck-Wood, S., & Mertens, S. (2019). SANRA—A scale for the quality assessment of narrative review articles. Research Integrity and Peer Review, 4(1), 5. https://doi.org/10.1186/s41073-019-0064-8
- Bronfenbrenner, U. (1979). The ecology of human development: Experiments by nature and design. Harvard University Press.
- Buarque, G. (2023). Artificial intelligence and algorithmic discrimination: a reflection on risk and vulnerability in childhood. Brazilian Journal Law, Technology Innovation, 1(2), 63 - 86.and https://doi.org/10.59224/bjlti.v1i2.63-86
- Devillers, L., & Cowie, R. (2023). Ethical considerations on affective computing: An overview. *Proceedings of the* IEEE, 111(10), 1445-1458. 10.1109/JPROC.2023.3315217
- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L., Riley, R., & Sutton, A. J. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Medical Research Methodology, 6, 35. https://doi.org/10.1186/1471-2288-6-35
- Franz, L., Goodwin, C. D., Rieder, A., Matheis, M., & Damiano, D. L. (2022). Early intervention for very young children with or at high likelihood for autism spectrum disorder: An overview of reviews. Developmental Medicine & Child Neurology, 64(9), 1063–1076. https://doi.org/10.1111/dmcn.15258
- Grace, T. D., Abel, C., & Salen, K. (2023). Child-centered design in the digital world: investigating the implications of the Age-Appropriate design code for interactive digital media. In Proceedings of the 22nd annual ACM interaction design and children conference (pp. 289-297). https://doi.org/10.1145/3585088.3589370
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x
- Holstein, K., & Doroudi, S. (2022). Equity and artificial intelligence in education. In The ethics of artificial intelligence in education (pp. 151-173). Routledge.
- Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. International Journal of Artificial Intelligence in Education, 26(1), 25–36. https://doi.org/10.1007/s40593-015-0065-9
- Kurian, N. (2023). Toddlers and robots? The ethics of supporting young children with disabilities with AI companions and the implications for children's rights. *International Journal of Human Rights Education*, 7(1), 9.

- Lemaignan, S., Newbutt, N., Rice, L., Daly, J., & Charisi, V. (2021). UNICEF guidance on AI for children: Application to the design of a social robot for and with autistic children. arXiv preprint arXiv:2108.12166.United Nations. (1989). Convention on the rights of the child. https://www.ohchr.org/en/instrumentsmechanisms/instruments/convention-rights-child. https://doi.org/10.48550/arXiv.2108.12166
- Leong, W. Y., & Zhang, J. B. (2025). Ethical design of AI for education and learning systems. ASM Science Journal, 20(1), 1-9. https://doi.org/10.32802/asmscj.2025.1917
- McStay, A., & Rosner, G. (2021). Emotional artificial intelligence in children's toys and devices: Ethics, governance and practical remedies. Big Data & Society, 8(1), https://doi.org/10.1177/2053951721994877
- Miao, F., & Holmes, W. (2021). AI and education: A guidance for policymakers. Unesco Publishing. ISBN 978-92https://discovery.ucl.ac.uk/id/eprint/10130180/1/Miao%20and%20Holmes%20-%20201%20-3-300165-7. %20AI%20and%20education%20guidance%20for%20policy-makers.pdf
- Morandín-Ahuerma, F. (2023). Ten UNESCO recommendations on the ethics of artificial intelligence. OSF. https://www.researchgate.net/publication/374234687 Ten UNESCO Recommendations on the Ethics of Ar tificial Intelligence
- National Institute of Standards and Technology NIST (2023). AI risk management framework (AI RMF 1.0). https://doi.org/10.6028/NIST.AI.100-1
- Noddings, N. (1985). Caring: A feminine approach to ethics. Center for Research on Women, Stanford University.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. https://doi.org/10.1518/001872097778543886
- Skitka, L. J., Mosier, K. L., & Burdick, M. (2000). Accountability and automation bias. *International Journal of* Human-Computer Studies, 52(4), 701–717. https://doi.org/10.1006/ijhc.1999.0349
- Smuha, N. A. (2021). Beyond the individual: governing AI's societal harm. Computer Law & Security Review, 45, 105678. https://doi.org/10.14763/2021.3.1574
- Saxena, C. (2024). Ethical Considerations in Affective Computing. In: Garg, M., Prasad, R.S. (eds) Affective Computing for Social Good. The Springer Series in Applied Machine Learning. Springer, Cham. https://doi.org/10.1007/978-3-031-63821-3 13
- Tabassi, E. (2023).intelligence 1.0). Artificial risk management framework **RMF** (AI https://doi.org/10.6028/NIST.AI.100-1
- United Nations. (1989). Convention on the rights of the child. https://www.ohchr.org/en/instrumentsmechanisms/instruments/convention-rights-child. https://doi.org/10.48550/arXiv.2108.12166

United Nations Children's Fund (UNICEF) (2021). Policy guidance on AI for children. https://www.unicef.org/innocenti/es/media/1351/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021 ES.pdf

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.

World Economic Forum (2024). Shaping the Future of Learning: The Role of AI in Education 4.0. https://www3.weforum.org/docs/WEF Shaping the Future of Learning 2024.pdf

## **Corresponding Author Contact Information:**

Author name: Rocio Garcia-Peinado

**Department**: Department of Developmental and Educational Psychology

University, Country: Autonomous University of Madrid

Email: rocio.garcia@uam.es

**ORCID**: https://orcid.org/0000-0003-0751-8374

Please Cite: Garcia-Peinado, R. (2025). Educational Artificial Intelligence, Child Rights, and Human Care in Early Childhood. The European Educational Researcher, 8(3), 33-55. DOI: https://doi.org/10.31757/euer.833

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Conflict of Interest: The authors declare no conflict of interest. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Data Availability Statement: Data Availability Statement: Data are available on request from the author.

Ethics Statement: The authors declare no conflict of interest. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Author Contributions: The author designed the study, collected the data, conceived and designed the analysis, performed the analysis, and wrote the paper.

Received: April 12, 2025 • Accepted: September 08, 2025