

Use of Social and Educational Robots in Early Childhood and Elementary Education Teaching-Learning Processes: A Scoping Review

Maria Klara Brito Sena¹, Alan de Oliveira Santana^{1,2}, Raul Benites Paradedá^{1,2}

¹Laboratory of Learning Robotics – Department of Computer Science
State University of Rio Grande do Norte (UERN)
Av. Dr. João de Medeiros Filho – Natal – RN – Brazil – 59120-555

²Graduate Program in Computer Science (PPgCC)
State University of Rio Grande do Norte (UERN)
Mossoró – RN – Brazil

mariaklara@alu.uern.br, {alansantana, rauparadedá}@uern.br

Abstract. *This scoping review synthesizes robot-assisted education for children aged 3–12. Analyzing 27 empirical studies (2020–2025), we examine platforms, emotional effects, and learning domains. Findings show a reliance on costly commercial humanoids and high student motivation, but limited evidence for sustained cognitive gains. Research primarily targets STEM and literacy, varying between educational levels. However, the field suffers from small samples, brief interventions, and weak theoretical grounding. This review emphasizes the need for inclusive, methodologically robust, and pedagogically grounded research that prioritizes accessibility and equity.*

Keywords *Social robots, Educational robotics, Early childhood education, Affective engagement, Scoping review.*

1. Introduction

The integration of social robots into educational settings has emerged as a promising area within Human-Robot Interaction (HRI) research. Unlike traditional educational technologies, physically embodied robots offer unique affordances for learning through multimodal interaction, social presence, and affective engagement [Belpaeme et al. 2018]. Recent empirical studies have demonstrated the potential of these robots across diverse educational domains, from mathematics instruction [Vitale e Dello Iacono 2024, Ekström et al. 2024] and literacy development [Zhang et al. 2023] to second language acquisition [Tolksdorf et al. 2021], socioemotional support for children with Autism Spectrum Disorder (ASD) [van Otterdijk et al. 2020], and affective educational interactions [Souza et al. 2024], a direction that has also been pursued in the Brazilian context, as illustrated by the Roboldo system developed and presented at the SBC's Simposio Brasileiro de Robotica [Souza et al. 2024].

Previous reviews have synthesized evidence on social robots in education [Belpaeme et al. 2018, Woo et al. 2021], with particular focus on autism interventions [Kouroupa et al. 2022, Kewalramani et al. 2024] and individual learning differences [Konijn e Hoorn 2020]. However, rapid technological advances, such as the integration

of Large Language Models (LLMs) into robotic systems [Elgarf et al. 2024] and the diversification of hardware platforms, alongside the significant evolution in pedagogical applications, target populations, and reported outcomes, have created an urgent need for an updated synthesis that these prior reviews do not yet address.

To address these gaps, this scoping review maps and synthesizes the empirical literature on the use of social and educational robots in child teaching-learning processes, covering studies published between 2020 and 2025, with the following research questions:

1. What are the most frequently reported robot platforms (e.g., NAO, Jibo, QT) in the international literature for interventions with children?
2. What empirical evidence exists regarding the emotional impact of social robots in these contexts?
3. Regarding cognitive and learning impacts, which knowledge areas benefit the most?
4. How do learning objectives (application contexts) and types of robots used differ among the main age groups identified (e.g., preschool vs. elementary school)?

Following PRISMA-ScR guidelines [Tricco et al. 2018], a systematic search was conducted across six databases, including Brazilian repositories (BDTD and CAPES Portal), covering studies in English and Portuguese. After selection, 27 empirical studies were included in the final synthesis.

2. Related Work

Research on social and educational robots has grown considerably over the past decade, generating a body of systematic reviews that informed the design of this study. However, existing reviews present significant temporal, thematic, and methodological limitations that justify an updated synthesis focused on the 2020–2025 period and on age-specific patterns within early childhood and elementary education.

2.1. Systematic Reviews and Meta-Analyses

The foundational reference for the field remains Belpaeme et al. [Belpaeme et al. 2018], whose review established the conceptual vocabulary still widely used today: robots as tutors, peers, or tools; the role of social presence and embodiment in learning; and the tension between affective engagement and measurable cognitive outcomes. Their analysis covered studies from multiple educational domains and age groups, offering design principles that continue to guide research. However, this review predates the widespread integration of Large Language Models into robotic systems [Elgarf et al. 2024] and the significant diversification of hardware platforms that characterizes the post-2020 landscape.

Woo et al. [Woo et al. 2021] advanced the field by restricting their scope to fieldbased classroom studies, deliberately excluding laboratory experiments to prioritize ecological validity. This methodological choice yielded important insights into real-world implementation challenges, including infrastructure constraints, teacher roles, and classroom logistics. Yet, the review lacked systematic age-specific analysis within early childhood, treated robot platforms descriptively rather than comparatively, and did not examine emotional outcomes with the same rigor applied to behavioral and cognitive measures.

Two more recent reviews focused specifically on children with Autism Spectrum Disorder. Kouroupa et al. [Kouroupa et al. 2022] conducted a meta-analysis reporting moderate positive effects on social functioning ($g = 0.35$, 95% CI [0.09, 0.61]), representing one of the few quantitative effect size estimates available in this literature. Kewalramani et al. [Kewalramani et al. 2024] extended this line by examining socialemotional learning outcomes specifically, noting that robots function particularly well as non-threatening social agents for children who struggle with direct human interaction. Both reviews, however, are restricted to atypical populations and do not address typically developing children in mainstream educational settings.

Konijn and Hoorn [Konijn e Hoorn 2020] examined the role of robot communication styles in moderating learning outcomes, finding that individual differences such as prior knowledge and cognitive style significantly affect how children respond to robotic instruction. Their analysis highlighted that a uniform robot design cannot optimally serve all learners, a finding with direct implications for adaptive systems. However, their scope was limited to a single domain (mathematics) and a narrow age range, restricting the generalizability of their conclusions.

2.2. Educational Robotics in the Brazilian Context and Identified Gaps

Research on educational robotics in Brazil has grown notably in recent years. Sokolonski et al. [Sokolonski et al. 2020], in a systematic review, investigated the use of educational robotics as a facilitator of computational thinking across all educational levels, identifying positive effects on logical reasoning, problem solving, and creativity. Ferreira et al. [Ferreira e Lima 2023] extended this line in a broader review covering cognitive skill development through educational robotics, confirming that most studies report gains in problem solving and collaborative learning. These Brazilian reviews demonstrate a consolidated interest in the cognitive dimensions of educational robotics.

Also representative of recent Brazilian production is the Roboldo system [Souza et al. 2024], which explored affective and empathic interactions between children and a socially expressive robot. This work reflects a broader movement toward developing robots adapted to local educational realities, including public school constraints and socioeconomic variability that differ substantially from the North American and European contexts that dominate the international literature.

Beyond the studies discussed above, Brazilian research on robotics in education has also been disseminated through national conferences such as the Brazilian Symposium on Computers in Education (SBIE), the Brazilian Congress on Informatics in Education (CBIE), and the Workshop de Robótica na Educação (WRE). Nevertheless, much of this production focuses on educational robotics, computational thinking, or robot programming activities, rather than on socially interactive robots, which helps explain its limited representation within the scope of the present review.

Notably, despite the explicit inclusion of Brazilian repositories (BDTD and Capes Portal), no Brazilian empirical study met all predefined eligibility criteria for inclusion in the final corpus. This finding should not be interpreted as evidence of a lack of Brazilian research in educational robotics. Rather, it likely reflects the combined effect of the review scope, which focused exclusively on physically embodied social robots for children aged 3-12, the adopted search terms, and the indexing characteristics of the selected

databases. A substantial portion of Brazilian research has been disseminated through national venues such as SBIE, CBIE, and the Workshop de Robótica na Educação (WRE), frequently emphasizing educational robotics, computational thinking, or programmable robotic platforms rather than socially interactive robots. This observation reinforces the need for broader indexing of Brazilian research and highlights an opportunity for future reviews adopting complementary search strategies.

Taken together, the existing body of reviews leaves four interconnected gaps unaddressed. First, most comprehensive syntheses predate 2022 and do not capture the integration of generative AI and LLMs into robotic systems [Elgarf et al. 2024], which has meaningfully expanded interaction capabilities. Second, existing reviews either focus narrowly on specific populations such as children with ASD, or treat all ages uniformly, without systematic developmental analysis comparing preschool and elementary school contexts. Third, no prior review has comprehensively mapped the landscape of deployed robot platforms across recent empirical work, making it difficult to assess trends in morphology, cost, and accessibility. Fourth, even Brazilian reviews with broader scope [Sokolonski et al. 2020, Ferreira e Lima 2023] prioritize computational thinking and cognition over affective outcomes and social robot platforms, leaving the emotional dimension of robot-mediated learning underexplored in the national literature. This scoping review addresses all four gaps by covering 2020–2025 literature, systematically mapping platforms, roles, and populations, jointly examining emotional and cognitive outcomes, and situating findings within a discussion of accessibility, equity, and pedagogical relevance for the Brazilian educational context.

3. Methodology

This section details the methodological approach adopted to conduct this scoping review. The process was structured to ensure transparency, rigor, and replicability, following the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guidelines [Tricco et al. 2018] and the methodological framework for scoping reviews proposed by Peters et al. [Peters et al. 2021].

3.1. Research Design and Rationale

A scoping review approach was selected to systematically map the extent, scope, and nature of existing literature on social and educational robots in early childhood and elementary education. This approach is particularly suited for identifying concepts, platforms, educational contexts, and research gaps in rapidly evolving fields, rather than estimating effect sizes [Peters et al. 2021]. The four research questions presented in Section 1 guided this mapping across robot platform diversity, emotional impacts, cognitive and learning outcomes, and age-specific deployment patterns.

3.2. Databases

Six databases were searched between October 2 and 14, 2025, using institutional credentials: BDTD (<https://bdttd.ibict.br>), CAPES Portal (<https://www.periodicos.capes.gov.br>), IEEE Xplore (<http://ieeexplore.ieee.org/Xplore>), Scopus (<https://www.scopus.com>), Springer Nature (<https://link.springer.com>), and Web of Science (<https://www.webofscience.com>). The Brazilian repositories (BDTD and CAPES

Portal) were included to capture national production in Portuguese, which is systematically absent from most international reviews in this area.

3.3. Search Terms and Boolean Strings

The search strategy combined descriptors from three dimensions: (1) Social/Educational Robots, (2) Educational Context, and (3) Target Population, using Boolean operators (AND, OR). The core search string was:

Primary Search String: ("social robot*" OR "educational robot*" OR "robotic tutor*") AND ("learning" OR "education*" OR "classroom" OR "teaching") AND ("child*" OR "preschool*" OR "elementary*" OR "early childhood")

Secondary Search String (Expanded for Scopus/WoS):
(TITLE-ABS-KEY("social robot*" OR "educational robot*" OR "robotic tutor*") AND TITLE-ABS-KEY("learning outcome*" OR "cognitive development" OR "emotional engagement" OR "literacy" OR "mathematics") AND TITLE-ABS-KEY("child*" OR "preschool*" OR "elementary school"))

Database-specific syntax adaptations (e.g., TITLE-ABS-KEY for Scopus) were applied. Results were filtered for the 2020–2025 publication period and English or Portuguese languages. These language restrictions were consistently applied across all databases and throughout the study selection process.

3.4. Selection Criteria

Studies were selected based on predefined inclusion and exclusion criteria established a priori, as detailed in Table 1.

Table 1. Inclusion and Exclusion Criteria

ID	Inclusion Criteria (IC)	Exclusion Criteria (EC)
C1	Open Access publications	Paywalled or restricted access publications
C2	Published between 2020 and 2025	Studies published before 2020 or after 2025
C3	Written in English or Portuguese	Publications in other languages
C4	Primary empirical research articles	Reviews, meta-analyses, theoretical papers, or grey literature
C5	Focus on physically embodied social/educational robots for children aged 3–12 years, reporting learning/behavioral outcomes	Studies focusing exclusively on virtual (screen-based) agents, outside the target age group, or unrelated topics

3.5. Study Selection Process and Data Extraction

The study selection followed five sequential steps: (1) initial database searches; (2) deduplication via Mendeley; (3) independent title and abstract screening by two reviewers, with disagreements resolved by a third; (4) full-text assessment against the inclusion criteria, with exclusion reasons documented; and (5) final inclusion of eligible studies. Detailed results, including record counts at each stage and the PRISMA flow diagram, are reported in Section 4.

A standardized data extraction form guided systematic information collection. Extracted data included study characteristics (authors, year, country, design, sample, setting), population

characteristics (age, grade, special needs), robot characteristics (platform, features, interaction modalities), pedagogical characteristics (robot role, learning domain, intervention duration), measured outcomes (cognitive, emotional, behavioral), and key findings (results, effect sizes, significance, author-noted limitations).

3.6. Data Synthesis and Analysis

Given the heterogeneity of the included studies, a narrative synthesis was adopted rather than a quantitative meta-analysis. The synthesis was organized thematically around each research question, using descriptive statistics for RQ1 and thematic synthesis for RQ2 through RQ4, to identify patterns, trends, and gaps across the literature.

4. Results

This section presents the findings of the scoping review organized in two parts. The first describes the study selection process following PRISMA-ScR guidelines, including record counts at each stage. The second addresses each of the four research questions.

4.1. Study Selection Process: PRISMA Flow

The systematic search across the six selected databases initially identified 2,063 records. As illustrated in Figure 1, applying the database filtering criteria (C1 to C5) excluded 1,816 records, resulting in 247 records assessed for eligibility. Following full-text assessment, 212 records were excluded, leaving 35 pre-selected eligible records. After removing 8 duplicate records, a final corpus of 27 empirical studies was included in the scoping review. Table 2 summarizes their key characteristics, with detailed information available in [Sena e Paradedá 2025].

4.2. Results by Research Question

RQ1 (Robot Platforms): Twenty-one distinct robot platforms were identified across the 27 studies, comprising four platforms with multiple occurrences and 17 platforms appearing in only one study each (see Table 3). NAO was the most frequently used ($n=9$), attributed to its established research track record and robust SDKs [Almoussa e Alghowinem 2022]. Pepper ($n=4$) was often employed in collaborative scenarios for older children [Ekström et al. 2024]. Furhat ($n=3$) and Jibo ($n=2$) were also notable, with smaller platforms such as Jibo preferred for younger children due to cost and portability [Zhang et al. 2023]. A consistent trend in 2024–2025 studies is the integration of Large Language Models to enhance conversational naturalness [Elgarf et al. 2024]. Notably, all 27 included studies originate from North America, Europe, or East Asia; no study conducted in Brazil or Latin America met the inclusion criteria, despite the explicit search of BDTD and CAPES Portal.

RQ2 (Emotional Impact): A majority of studies (23/27, 85.2%) reported at least one affective outcome. Among these 23 studies, engagement was the most common finding ($n=19$, 82.6%), followed by increased motivation ($n=14$, 60.9%) and enjoyment ($n=12$, 52.2%). Seven studies (30.4%) reported anxiety reduction, especially when robots acted as non-judgmental peers [Cagiltay et al. 2023, Tolksdorf et al. 2021]. Results are summarized in Table 4. A notable methodological limitation across these studies was the heterogeneity in measurement approaches, ranging from direct observation and self-report questionnaires to automated facial expression analysis, which hinders direct comparisons across studies.

RQ3 (Cognitive Outcomes): Mathematics/STEM was the largest cluster (11 studies, 37.9%), with evidence for gains in concept acquisition, particularly in geometry [Ekström et al. 2024], and collaborative problem-solving. Literacy/Language (8 studies,

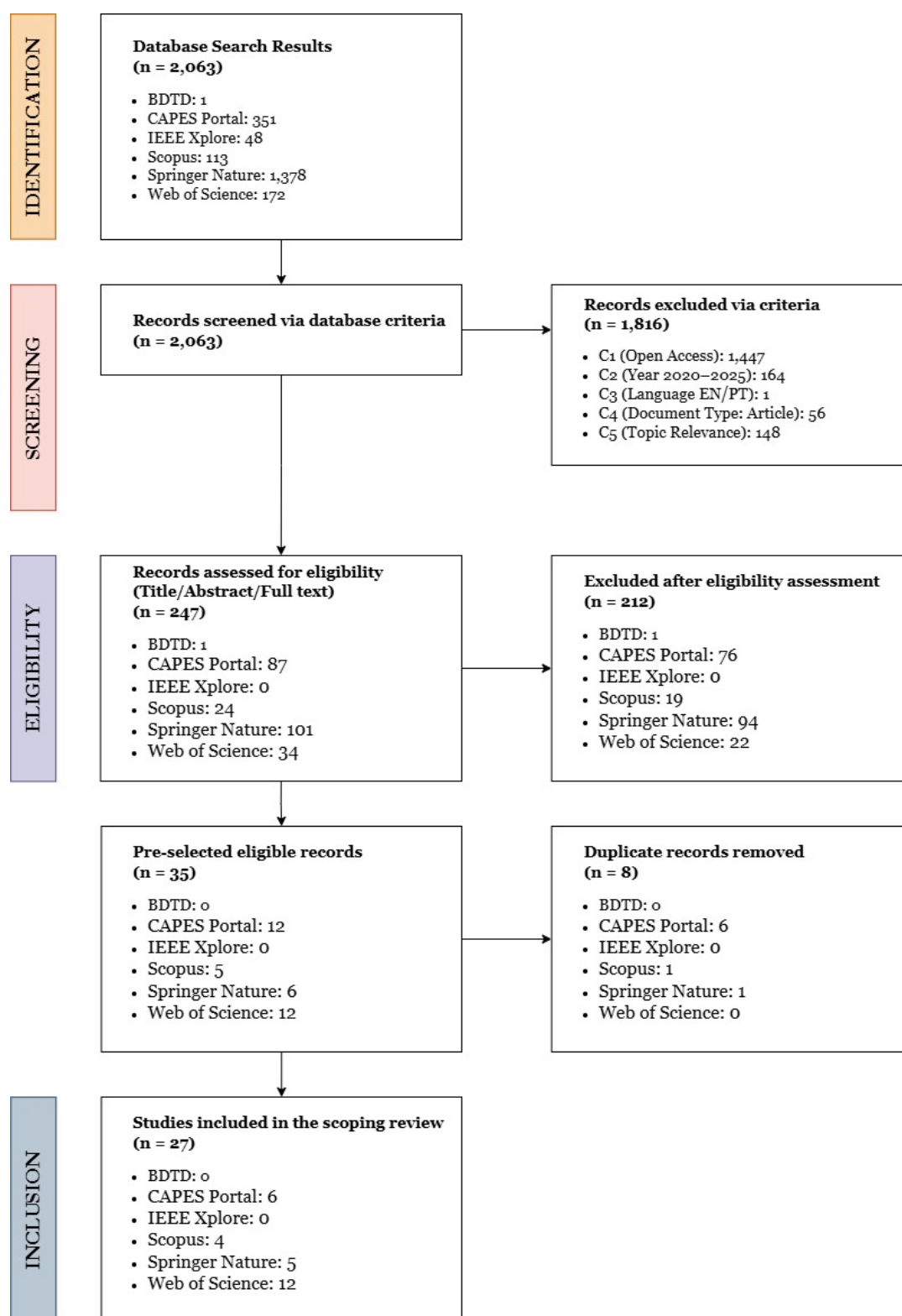


Figure 1. PRISMA flow diagram illustrating the systematic study selection process. The diagram shows the number of records identified, screened, assessed for eligibility, and ultimately included in the scoping review.

Table 2. Overview of included studies (N=27) by key characteristics. Some studies were classified into more than one category (e.g., robot platform, learning domain, or study design); therefore, category totals may exceed number of included studies.

Characteristic	Category	n	%
Publication Year	2020	3	11.1
	2021	7	25.9
	2022	4	14.8
	2023	4	14.8
	2024-2025	9	33.3
Robot Platform	NAO	9	25.7
	Pepper	4	11.4
	Jibo	2	5.7
	Furhat	3	8.6
	Other (n=1 each)	17	48.6
Age Group	Preschool (3-5 years)	10	37.0
	Elementary (6-12 years)	17	63.0
Learning Domain	Mathematics/STEM	11	37.9
	Literacy/Language	8	27.6
	Social-Emotional Learning	6	20.7
	Second Language (L2)	3	10.3
	Multiple Domains	1	3.4
Robot Role	Tutor	16	55.2
	Peer	9	31.0
	Tool/Programmable	4	13.8
Study Design	Experimental/RCT	12	41.4
	Quasi-experimental	10	34.5
	Case study/Field study	7	24.1

Table 3. Comparison of robot platforms used in educational interventions

Platform	n	Height	Key Features	Modalities
NAO	9	58 cm	Humanoid, 25 DOF	Speech, gesture, touch
Pepper	4	120 cm	Humanoid, tablet	Speech, gesture, screen
Jibo	2	28 cm	Social robot, 3 DOF	Speech, gaze, animation
Furhat	3	41 cm	Head-only, expressive	Speech, gaze, facial exp.
Other (n=1 each)	17	Varies	Diverse	Varies

Table 4. Emotional impacts reported in included studies

Emotional Outcome	n	%	Measurement Method
Engagement/Interest	19	82.6	Observation, self-report, video coding
Motivation	14	60.9	Questionnaires, task persistence
Enjoyment/Affect	12	52.2	Facial expression analysis, surveys
Anxiety Reduction	7	30.4	Self-report, physiological measures
Social Connection	9	39.1	Interaction patterns, interviews

27.6%) showed moderate effects, with robots serving as reading partners[Zhang et al. 2023] and adaptive tutors capable of adjusting difficulty based on children’s attention levels [Almoussa e Alghowinem 2022]. Social-Emotional Learning (6 studies, 20.7%) showed promise, especially for children with ASD, though skill generalization to human interactions remained variable across studies.

RQ4 (Age-Specific Patterns): Clear developmental differentiation was found across age groups, as detailed in Table 5. Preschool interventions (37.0%) favored smaller platforms such as NAO and Jibo, predominantly in tutor roles (80%) to scaffold foundational skills, with shorter sessions averaging 10–15 minutes and outcome measures focused on engagement and affect. Elementary school interventions (63.0%) more often employed taller, peer-like robots such as Pepper in peer roles (59%) for collaborative work on more abstract content, with sessions averaging 20–30 minutes and a stronger emphasis on cognitive learning gains. It should be noted that studies spanning a broad age range were categorized into both groups for specific dimensions, which may result in occurrence counts exceeding the number of studies per age category in those dimensions.

Although elementary education studies more frequently reported positive learning outcomes (82%), these outcomes predominantly referred to immediate post-intervention measures rather than long-term retention or transfer. Therefore, the higher frequency of reported learning gains should not be interpreted as conclusive evidence of sustained cognitive benefits, but rather as an indication of promising short-term effects that require further longitudinal validation.

Table 5. Comparison of robot interventions: Preschool (3-5 years) vs. Elementary (6-12 years)

Dimension	Preschool (n=10)	Elementary (n=17)
Predominant Platforms	NAO (70%), Jibo (20%); <i>Smaller, less intimidating</i>	Pepper (29%), NAO (29%); <i>Taller, more peer-like stature</i>
Primary Robot Roles	Tutor (80%), Companion (20%); <i>Expert scaffolding preferred</i>	Peer (59%), Tutor (35%); <i>Collaborative learning emphasized</i>
Learning Domains	Literacy (60%), Basic math (30%); <i>Foundational skills focus</i>	STEM/Math (47%), L2 (24%); <i>Advanced reasoning and content</i>
Session Duration	Shorter (10-15 min average); <i>Limited attention spans</i>	Longer (20-30 min average); <i>Extended engagement capacity</i>
Interaction Modality	Touch + speech + gesture; <i>Multimodal sensory engagement</i>	Speech + screen-based content; <i>Visual-symbolic representations</i>
Primary Outcomes	Engagement (90%), Affect (70%); <i>Affective emphasis</i>	Learning gains (82%), Cognitive (71%); <i>Cognitive emphasis</i>
Study Settings	Laboratory (60%), Clinic (30%); <i>Controlled environments</i>	Classroom (71%), After-school (24%); <i>Ecological validity priority</i>

5. Discussion

This review’s findings reveal consistent patterns and gaps in social robot use in early childhood and elementary education. This section interprets results across the four research questions,

examines methodological limitations, and discusses implications for researchers, practitioners, and designers, with particular attention to the Brazilian educational context.

5.1. Robot Platforms: Dominance, Accessibility, and Morphology

NAO's dominance reflects its maturity in research infrastructure, yet raises concerns about generalizability and accessibility. High costs (e.g., NAO, Pepper) create significant barriers for widespread adoption, especially in low-resource public schools [Belpaeme et al. 2018]. Dependence on commercial products also entails risks, as Jibo's discontinuation illustrated. The 17 single-occurrence platforms suggest exploration of alternatives, but no affordable option has achieved critical mass for systematic replication.

Limited diversity in humanoid robot morphology constrains understanding of how embodiment affects learning. Non-humanoid designs, potentially reducing social anxiety and anthropomorphization, remain underexplored (2020–2025). Recent integration of Large Language Models [Elgarf et al. 2024] extends robot capabilities beyond preprogrammed interactions, offering adaptive dialogue. However, this also introduces risks regarding content reliability, output unpredictability, and ethical implications for young children

5.2. Affective Engagement vs Cognitive Evidence: A Critical Tension

High rates of affective engagement are educationally meaningful, but engagement does not automatically translate into sustained cognitive gains. Many interventions relied on observational measures, self-reports, or short-term assessments, providing limited evidence of long-term learning.

Although a substantial proportion of elementary education studies reported positive learning outcomes (Table 3), these findings should be interpreted cautiously. Most reported gains were measured immediately after relatively short interventions and were based on small or medium-sized samples, with only six studies including follow-up assessments. Therefore, while short-term improvements were frequently observed, the current evidence remains insufficient to conclude that social robots produce sustained cognitive gains over time.

Only 6 of 27 studies (22.2%) included follow-up assessments, making it difficult to determine if effects are sustained or novelty-driven [Belpaeme et al. 2018]. Even with comparable literacy outcomes for robot and human tutors [Neumann et al. 2023], underlying mechanisms remain unclear (e.g., robot affordances, practice time, Hawthorne effect).

Moreover, a strong design focus on positive emotions may overlook productive struggle in deep learning [Kapur e Bielaczyc 2012]. Current robot designs, optimized for engagement, might inadvertently reduce the cognitive demand necessary for durable learning. Future designs should support learners through difficulty, not merely keep them engaged.

5.3. Learning Domains: Concentration, Gaps, and Untapped Potential

Concentration on Mathematics/STEM and Literacy reflects broader educational technology trends, but exposes missed opportunities. Social-Emotional Learning (SEL) applications, including creativity and empathy, show robots can support interpersonal development beyond academics [Belpaeme et al. 2018], yet explicit SEL integration with curricular goals is rare. This is particularly relevant in Brazil, where BNCC emphasizes socioemotional competencies.

Second language learning is similarly underrepresented, despite robots serving as patient practice partners [Konijn et al. 2022]. Domains like physical education, arts, music, and inquiry-based science are virtually absent, reflecting assumptions about robot roles rather than systematic exploration. These gaps suggest the field is shaped more by technical affordances and computer-science preferences than by pedagogical need.

5.4. Age-Specific Considerations and Developmental Appropriateness

Systematic age-related differences in robot deployment (platform, role, session length, outcome focus) suggest intuitive developmental appropriateness. Preschool studies favored smaller platforms in tutor roles for foundational skills (10–15 min sessions, affective focus). Elementary interventions used taller, peer-like robots in peer roles for abstract concepts (20–30 min sessions, cognitive focus).

However, few studies explicitly justified these design choices developmentally. Age-appropriateness was often assumed, despite younger children's theory of mind, executive function, and literacy levels fundamentally shaping robot interaction [Belpaeme et al. 2018]. The absence of research with children under 3 and few studies with adolescents leaves significant gaps across the developmental spectrum.

5.5. Methodological Limitations and Research Gaps

Consistent methodological patterns limit current evidence's robustness and generalizability. Most experimental and quasi-experimental studies relied on relatively small samples (median $n = 24$) and short intervention periods, which may reduce statistical power for detecting long-term learning effects. However, these characteristics should be interpreted in light of each study's methodological approach, since qualitative and exploratory designs often employ smaller samples while still providing valuable evidence regarding interaction processes and educational experiences.

In addition to the limitations observed in the primary studies, this review also presents methodological constraints that should be acknowledged. The search was restricted to studies published between 2020 and 2025 and to publications in English or Portuguese indexed in the selected databases. Consequently, relevant studies published in other languages, available exclusively as grey literature, or indexed in different repositories may not have been captured. Furthermore, although study selection was conducted independently by two reviewers with disagreements resolved by a third reviewer, no inter-rater agreement statistic (e.g., Cohen's kappa) was calculated. These aspects should be considered when interpreting the comprehensiveness of the present synthesis.

Geographic concentration in high-resource contexts [Kim et al. 2021, Konijn et al. 2022] raises equity and generalizability concerns, especially for Brazil. Public school infrastructure, class sizes, linguistic diversity, and socioeconomic variability differ substantially, reinforcing the need for nationally situated research adapting international protocols to local frameworks like the BNCC.

Longitudinal studies are rare, limiting insight into novelty effects. Few studies formulate theory-driven hypotheses on how robot affordances support learning. Costbenefit analyses are absent, despite substantial robot costs. Ethical issues (data privacy, consent, manipulation, delegation of instruction) were only occasionally discussed.

5.6. Implications for Research and Practice

Effective robot-mediated interventions require careful developmental alignment. Researchers need adequately powered, longitudinal, theory-driven studies in underexplored domains and populations, particularly in Latin American and other underrepresented contexts. Practitioners should view robots as complements, not replacements, considering technical demands, costs, and limited long-term evidence. Designers must ground choices in pedagogical theory and developmental psychology, exploring lower-cost, diverse robot forms, and engaging educators and children as co-designers.

6. Conclusion

This scoping review synthesized 27 empirical studies (2020–2025) on social and educational robots in early childhood and elementary education, addressing four research questions.

RQ1 revealed NAO's dominance, highlighting cost and fragmentation as barriers. RQ2 showed consistent high affective engagement, but limited evidence for sustained cognitive gains. RQ3 identified Mathematics/STEM and Literacy as primary foci, with underrepresentation in SEL and L2. RQ4 demonstrated systematic age-related differences in robot deployment, often lacking explicit developmental justification.

The corpus reveals structural limitations: small samples, short interventions, geographic concentration in high-resource contexts, and limited ethical reflection. These patterns constrain generalizability, particularly for contexts like Brazilian public schools, reinforcing the need for nationally situated empirical work.

Advancing the field requires accessible platforms, longitudinal and adequately powered studies, stronger theoretical grounding, and explicit engagement with equity, ethics, and cultural context. As LLM-based interaction capabilities expand [Elgarf et al. 2024], educational robot use must prioritize pedagogical goals over technological novelty, aiming to augment rather than automate children's learning experiences.

Finally, this review should be interpreted as a synthesis of the recent international literature within a well-defined scope, rather than as an exhaustive representation of all educational robotics research, particularly considering the diversity of publication venues and indexing practices in the Brazilian context.

Declaration on the Use of Artificial Intelligence

Generative Artificial Intelligence tools were used exclusively to support grammatical revision, textual refinement, structural organisation of the manuscript, and the generation of the conceptual synthesis figure presented in this study. All conceptual decisions, literature selection, methodological procedures, interpretation of results, and final scientific content were critically reviewed and validated by the authors, who assume full responsibility for the accuracy, originality, and integrity of the manuscript.

Acknowledgment

The authors acknowledge the financial support provided by the Financiadora de Estudos e Projetos (Finep), the Ministry of Science, Technology and Innovation (MCTI), and the National Fund for Scientific and Technological Development (FNDCT) under the MCTI/Finep/FNDCT ProInfra Regional Development – North, Northeast and Central-West (NNECO 2024) Public Call, through the subproject “Assistive and Collaborative Technological Solutions to Support Educational Inclusion” (STecACAIE). This support covered the expenses associated with participation in and presentation of this work.

The second author acknowledges CAPES for support under the Program for Reducing Asymmetries in Graduate Education (PRAPG), Call No. 14/2023, linked to the Graduate Program in Computer Science (UERN–UFERSA).

References

Almousa, O. e Alghowinem, S. (2022). Conceptualization and development of an autonomous and personalized early literacy content and robot tutor behavior for preschool children. *User Modeling and User-Adapted Interaction*, 33:1–31.

- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., e Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, 3(21):eaat5954.
- Cagiltay, B., Mutlu, B., e Michaelis, J. E. (2023). “my unconditional homework buddy:” exploring children’s preferences for a homework companion robot. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference, IDC ’23*, page 375–387, New York, NY, USA. Association for Computing Machinery.
- Ekström, S., Pareto, L., e Ljungblad, S. (2024). Teaching in a collaborative mathematic learning activity with and without a social robot. *Education and Information Technologies*, 30:1301–1328.
- Elgarf, M., Salam, H., e Peters, C. (2024). Fostering children’s creativity through llm-driven storytelling with a social robot. *Frontiers in Robotics and AI*, 11:1457429.
- Ferreira, F. e Lima, D. (2023). The use of educational robotics in the development of students’ computational thinking and cognitive skills: a systematic literature review. *RENOTE*, 21:363–372.
- Kapur, M. e Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, 21(1):45–83.
- Kewalramani, S., Kidman, G., e Palaiologou, I. (2024). A scoping review of the use of robotics technologies for promoting social-emotional learning in children with autism spectrum disorder. *Education Sciences*, 14(1):24.
- Kim, Y., Marx, S., Pham, H., e Nguyen, T. (2021). Designing for robot-mediated interaction among culturally and linguistically diverse children. *Educational Technology Research and Development*, 69:1–22.
- Konijn, E., Jansen, B., Bustos, V., Hobbelink, V., e Vanegas, D. (2022). Social robots for (second) language learning in (migrant) primary school children. *International Journal of Social Robotics*, 14:1–17.
- Konijn, E. A. e Hoorn, J. F. (2020). Robot tutor and pupils’ educational ability: Teaching the times tables. *Computers & Education*, 157:103970.
- Kouroupa, A., Laws, K. R., Irvine, K., Mengoni, S. E., Baird, A., e Sharma, S. (2022). The use of social robots with children and young people on the autism spectrum: A systematic review and meta-analysis. *PLOS ONE*, 17(6):e0269800.
- Neumann, M. M., Koch, L.-C., Zagami, J., Reilly, D., e Neumann, D. L. (2023). Preschool children’s engagement with a social robot compared to a human instructor. *Early Childhood Research Quarterly*, 65:332–341.
- Peters, M., Marnie, C., Tricco, A., Pollock, D., Alexander, L., Mcinerney, P., Godfrey, C., e Khalil, H. (2021). Updated methodological guidance for the conduct of scoping reviews. *JBI evidence implementation*, 19:3–10.
- Sena, M. K. B. e Paradedá, R. B. (2025). Supplementary table: Studies included in the review — robots in early childhood education. https://osf.io/txz7d/overview?view_only=a575f26f4c264a11b1a88de561cc2034. Comprehensive, year-by-year list of studies included in the scope review.

- Sokolonski, A., de Sá, A., e Macêdo, R. (2020). Robótica educacional como facilitadora do aprendizado do raciocínio computacional: Revisão sistemática da literatura. In *Anais do XXXI Simpósio Brasileiro de Informática na Educação (SBIE 2020)*, pages 1503–1512.
- Souza, A., Sousa, C., Torres, D., e Paradedda, R. (2024). Roboldo: An affective robot for empathetic social interactions in education. In *Anais do XVI Simpósio Brasileiro de Robótica e XV Workshop de Robótica na Educação*, pages 307–312, Porto Alegre, RS, Brasil. SBC.
- Tolksdorf, N., Viertel, F., e Rohlfing, K. (2021). Do shy preschoolers interact differently when learning language with a social robot? an analysis of interactional behavior and word learning. *Frontiers in Robotics and AI*, 8:676123.
- Tricco, A., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Levac, D., Moher, D., Peters, M., Weeks, L., Hempel, S., Akl, E., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M., Garritty, C., e Straus, S. (2018). Prisma extension for scoping reviews (prisma-scr): Checklist and explanation. *Annals of Internal Medicine*, 169.
- van Otterdijk, M. T. H., de Korte, M. W. P., van den Berk-Smeekens, I., Hendrix, J., van Dongen-Boomsma, M., den Boer, J. C., Buitelaar, J. K., Lourens, T., Glennon, J. C., Staal, W. G., e Barakova, E. I. (2020). The effects of long-term child–robot interaction on the attention and the engagement of children with autism. *Robotics*, 9(4).
- Vitale, A. e Dello Iacono, U. (2024). Using social robots as inclusive educational technology for mathematics learning through storytelling. *European Public & Social Innovation Review*, 9:1–17.
- Woo, H., LeTendre, G. K., Pham-Shouse, T., e Xiong, Y. (2021). The use of social robots in classrooms: A review of field-based studies. *Educational Research Review*, 33:100388.
- Zhang, X., Breazeal, C., e Park, H. W. (2023). A social robot reading partner for explorative guidance. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction, HRI '23*, page 341–349, New York, NY, USA. Association for Computing Machinery.