Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-026-07519-6

ORIGINAL ARTICLE

Parent's AI Coach (PaiCoach): A Single-Case Experimental Study of an AI-Supported Parent Coaching System for Autistic Children

Yusuf Akemoğlu[1,3] · **Zhaohui Li**[1] · **Qingxiao Zheng**[1,4] · **Emily Brooke Roberts**[1] · **Samar Pratap Singh**[1] · **Jinjun Xiong**[1,2]

Received: 18 May 2026 / Accepted: 25 August 2026
© The Author(s) 2026

Abstract

Purpose Artificial intelligence (AI) may enhance the scalability and immediacy of parent coaching by providing automated, performance-based feedback within naturalistic routines. In this study, we examined the effects of Parent's AI Coach (PaiCoach), an AI-supported parent coaching system, on parent implementation of evidence-based practices in improving autistic children's communicative responsiveness during shared book reading.

Method Participants included four mother-child dyads with minimally verbal autistic children between 45 and 56 months of age. A concurrent multiple-baseline across participants single-case experimental design was used. Primary dependent variables included parent implementation fidelity and child communicative responsiveness. Visual analysis and Tau-U effect sizes were used to examine functional relations between the introduction of PaiCoach and changes in parent and child outcomes.

Results Visual analysis showed functional relations between the introduction of PaiCoach and increases in parent implementation fidelity with corresponding improvements in children's responsiveness. Through human validation, we show that PaiCoach adopts 68.6% of AI-generated analyses without modification, whereas 31.4% required some expert editing. Social validity and usability findings indicate high parent satisfaction with PaiCoach and its good usability. Overall, findings provide preliminary evidence for the feasibility of AI-assisted parent coaching and the continued importance of human oversight in current AI-supported systems.

Keywords Artificial intelligence · Autism · Parent-mediated intervention · Single-case experimental design · Human-in-the-loop · Multimodal AI

✉ Yusuf Akemoğlu
yusufakemoglu@gmail.com; yusufakemoglu@duzce.edu.tr

Zhaohui Li
zli253@buffalo.edu

Qingxiao Zheng
zqxjuly@gmail.com

Emily Brooke Roberts
ebr2@buffalo.edu

Samar Pratap Singh
samarpra@buffalo.edu

Jinjun Xiong
jinjun@buffalo.edu; jinjun@utsa.edu

[1] National AI Institute for Exceptional Education, University at Buffalo, Buffalo, NY 14260, USA

[2] College of AI, Cyber and Computing, University of Texas at San Antonio, San Antonio, TX 78204, USA

[3] Department of Special Education, Duzce University, 81620 Konuralp, Düzce, Turkey

[4] Department of Human-Centered Computing and Social Informatics, The Pennsylvania State University, State College, PA 16802, USA

Introduction

Autism spectrum disorder (ASD) is characterized by persistent differences in social communication and interaction that emerge early in development (American Psychiatric Association, 2013). Autistic children often experience challenges in joint attention, reciprocal engagement, and functional language use, which are foundational for later academic achievement, peer relationships, and adaptive functioning. Early intervention targeting these skills is therefore critical, particularly when embedded within meaningful, everyday interactions (Schreibman et al., 2015). However, despite a strong evidence base for early intervention, many families still face barriers in accessing consistent and quality services (Vivanti et al., 2018).

Parent-mediated interventions have emerged as a central and effective approach for addressing early social-communication challenges. These interventions position parents as primary implementers of evidence-based strategies within natural routines, thereby increasing intervention intensity and promoting generalization across contexts (Ousley et al., 2025). Within naturalistic developmental behavioral interventions, parents are taught to use strategies such as modeling, mand-model procedures, and time delay to create and respond to communication opportunities based on the child's interests (Meadan et al., 2016). A robust body of literature indicates that when parents are trained and supported, they can implement these strategies with fidelity which in turn result in improvements in autistic children's communicative responding and initiating (Akemoglu et al., 2020). However, despite the promise of parent-mediated interventions, their effectiveness depends heavily on implementation fidelity, which refers to the extent to which intervention strategies are delivered as intended (Ousley et al., 2025).

Maintaining high levels of implementation fidelity over time can be challenging for many families. Parents often receive coaching during scheduled sessions but may have limited access to ongoing support, feedback, and performance monitoring during everyday routines (Hinton et al., 2024). In addition, the demands of daily family life, competing responsibilities, and the complexity of implementing multiple communication strategies consistently across interactions may reduce sustained strategy use over time (Vivanti et al., 2018). Delays between parent performance and professional feedback may further limit opportunities for immediate correction and refinement (Akemoglu et al., 2020). As a result, there remains a need for more accessible and continuous coaching systems that can support parents between traditional intervention sessions.

Artificial intelligence (AI) offers a promising pathway to address these limitations by enabling more immediate, personalized, and scalable support (Marino et al., 2023; Perry et al., 2026). Within human-centered AI frameworks, AI systems could be designed to augment human learning by providing adaptive, context-sensitive feedback while maintaining alignment with user needs and values (Amershi et al., 2019; Meadan et al., 2026). In educational contexts, AI has been used to develop intelligent tutoring systems that adapt to learner performance and provide targeted feedback to support skill acquisition (Feng et al., 2025; Hopcan et al., 2023). Advances in large language models (LLM) further extend these capabilities by enabling systems to generate contextually relevant feedback and support interactive, language-based learning at scale (Rakap & Balikci, 2024). In parallel, developments in multimodal AI enable the analysis of complex behavioral data, including video-recorded interactions to detect patterns in communication and generate feedback (Holyfield et al., 2024; Li et al., 2025). Multimodal LLMs can integrate visual, auditory, and linguistic information to produce structured descriptions of human behavior and potentially bridge the gap between raw interaction data and expert interpretation (Bommasani et al., 2021). These systems shift the role of AI from delivering static metrics toward generating interpretable, language-based feedback aligned with human reasoning processes (Waterfield et al., 2024).

In this regard, some technologies incorporated automated analysis of speech and interaction patterns to provide feedback on features such as turn-taking and language use. For example, systems such as Captivate (Kwon et al., 2022) have incorporated multimodal sensing including gaze tracking and audio analysis to capture attention and engagement patterns during interaction. Similarly, the work of AACessTalk (Choi et al., 2025) has supported minimally verbal autistic children by combining vocabulary recommendations and guidance for parents to improve turn-taking. Together, these systems reflect a progression toward increasingly automated and data-driven approaches for supporting parent-child interaction.

In light of this, emerging research in autism intervention further supports the adoption of AI to enhance social-communication outcomes and intervention delivery (Feng et al., 2025; Kotsi et al., 2025). For example, Koegel et al. (2025) demonstrated that an LLM based conversational program could provide structured practice opportunities and immediate feedback on verbal empathetic responses. Their LLM based app Noora resulted in significant improvements in empathetic responding that generalized to natural social conversations among autistic adolescents and adults. Complementing this, other studies have shown that AI-supported tools can improve implementation fidelity by guiding users through evidence-based procedures and providing performance-based feedback (Griffen et al., 2024). AI-assisted

systems have also been shown to increase opportunities for practice and reduce delays between performance and feedback, addressing a central limitation of traditional coaching models (Benway & Preston, 2024). Taken together, this emerging literature suggests that AI can support continuous, data-informed learning processes that are particularly relevant for coaching complex interactional skills. Yet, despite these advances, the application of AI to parent-mediated autism intervention remains limited. Many technology-supported approaches do not provide individualized, time-anchored feedback that allows parents to evaluate and adapt their interactions within everyday routines (Akemoglu et al., 2026).

In response to these limitations, the Parent's AI Coach (PaiCoach) was developed as a human-in-the-loop, AI-supported parent coaching system designed to augment parent implementation of evidence-based communication strategies within naturalistic routines (Li et al., 2025; Zheng et al., under review). PaiCoach is grounded in the PiCSS (Parent-implemented Communication Strategies-Storybook) program (Akemoglu & Tomeny, 2021; Akemoglu et al., 2022a, 2022b), which itself is rooted in the PiCS framework (Meadan et al., 2016). The PiCSS program integrates naturalistic communication strategies, including modeling, mand-model, and time delay, within shared reading routines to support the social-communication development of young autistic children. Modeling involves demonstrating target words or communicative responses for the child, mand-model combines prompts (e.g., "say, 'toy'") or questions with language models to encourage child responding, and time delay creates opportunities for children to initiate communication by intentionally pausing before providing prompts. These strategies are designed to increase children's communicative responses and initiations during book reading.

The original PiCSS study demonstrated that parents of young children with developmental disabilities could learn to implement these strategies with fidelity following in-person training and coaching and could result in improvements in child communication outcomes (Akamoglu & Meadan, 2019). A subsequent replication study extended these findings to autistic preschoolers and further demonstrated improvements in parent implementation fidelity and child communicative responses and initiations during shared reading (Akemoglu & Tomeny, 2021). Building on this foundation, the intervention was adapted for telepractice delivery through the internet-based Parent-Implemented Communication Strategies-Storybook (i-PiCSS) for remote training and coaching (Akemoglu et al., 2022a). The model was subsequently expanded into a module-based telepractice format that combined self-directed online learning with individualized coaching to improve accessibility and scalability (Akemoglu et al., 2022b). Collectively, these studies provide empirical support for the instructional content, coaching procedures, and technology-assisted delivery methods underlying PaiCoach.

PaiCoach was developed through multiple user-study iterations within a human-in-the-loop framework in which parents record brief videos of their reading interactions with their children and upload the videos to the platform (Zheng et al., under review). Specifically, PaiCoach integrates computer vision, automatic speech recognition, and LLM-supported processing to analyze verbal and nonverbal interaction patterns within shared reading sessions. The system extracts information related to parent strategy use (e.g., modeling, mand-model, and time delay), child communicative responses, response timing, and conversational exchanges. Next, it generates preliminary time-stamped, strategy-specific feedback intended to support parent reflection and implementation (Li et al., 2025). Importantly, all AI-generated outputs are reviewed and validated by trained human reviewers prior to delivery to parents.

By combining AI-assisted feedback with human validation in a timely and accessible manner, PaiCoach aims to address limitations of traditional coaching models and augment human professionals such as special educators or speech-language pathologists that work with families (Zheng et al., under review). That is, PaiCoach maintains a human-centered approach by allowing human-expert reviews of the AI-generated analyses and feedback before such AI-generated results are made available to the parents (Amershi et al., 2019; Kotsi et al., 2025).

Despite the growing interest in AI-supported autism intervention, much of the existing literature remains preliminary and consists primarily of feasibility studies, prototype development, and descriptive evaluations (Akemoglu et al., 2026; Kotsi et al., 2025). This means that there is a limited number of rigorous experimental studies examining the effectiveness of AI tools. To address this gap in the literature, we aimed to evaluate the emerging potential of PaiCoach as a human-in-the-loop parent coaching system by examining its capacity to support parent implementation of communication strategies during shared reading. Thus, our study was guided by the following research questions:

(1) Is there a functional relation between the introduction of PaiCoach and changes in parents' fidelity of implementing naturalistic communication strategies?
(2) Is there a functional relation between parents' use of PaiCoach and changes in children's communicative responses?
(3) To what extent is PaiCoach perceived as socially valid by participating families?

Method

Participant Recruitment

After receiving approval from the Institutional Review Board, we distributed recruitment materials via email and community outreach through local autism centers, early childhood programs, school districts, and parent support networks in Western New York. Inclusion criteria for participation were as follows: (a) children were between 3 and 5 years of age, (b) children had a documented diagnosis of ASD and confirmed by the research team, and (c) children had approximately 20–30 functional spoken words as reported by parents. A total of eight families expressed interest in the study. Six families met the eligibility criteria following screening procedures and four families agreed to participate. As a token of appreciation for participation, each family received a $100 digital gift card.

Participants

Participants included four parent-child dyads. All participating parents were mothers and served as the primary adult interacting with the child during shared book reading routines. Parents ranged in age from 18 to 45 years and represented diverse educational and socioeconomic backgrounds with education levels ranging from GED to master's degrees and household incomes ranging from US$25,000 to over US$100,000.

Children ranged in age from 45 to 56 months and all had a documented ASD received between 24 and 36 months of age. All children were receiving intervention services, including early childhood special education, speech-language therapy, and occupational therapy, with some also receiving applied behavior analysis and physical therapy. Children demonstrated limited expressive communication characterized by gestures, vocalizations, single words, signs, and emerging two-word utterances. Assessment data demonstrate autism-related characteristics across participants. All study procedures were conducted remotely. See Table 1 for detailed parent and child demographic information and Table 2 for child assessment data.

Materials and Technology

Storybooks

A set of 10 developmentally appropriate children's storybooks was selected by the research team for use during shared reading sessions and delivered to the families (See S1 in Supplementary Materials for the book list). These books were chosen based on their suitability for young children, including features such as simple language, repetitive structures, and engaging illustrations to support interaction and communication. Four books were designated for use during the baseline and maintenance phases and six books were designated for use during the intervention phase. Parents were instructed to shuffle the books prior to each session.

Experimental Design

We used a concurrent multiple-baseline across participants single-case experimental design (Ledford & Gast, 2018) consisting of baseline, intervention, and maintenance phases. Dyads were staggered into intervention following stable baseline patterns in parent implementation fidelity across a minimum of three consecutive sessions. Stability was evaluated through visual analysis of level and trend.

Table 1 Demographic information on participating parents and children

Dyad	Age	Parents			Children			Primary mode of communication
		Race/Ethnicity	Education/ Marital status	Income	Age/Gender	Age of ASD diagnosis	Services	
1	36–45	White	Master's/Married	>$100,000	48 months/ Male	30 months	ECSE, SLT, PT, OT	Combination of gestures, sounds, single words and some two-word utterances.
2	36–45	White	Master's/Married	>$100,000	50 months/ Male	36 months	ECSE, ABA, SLT, OT	Gestures (pointing, reaching, leading), single words and signs.
3	18–24	African American	GED/Single	$25,000-$49,999	45 months/ Male	24 months	ECSE, SLT, OT	Combination of single words and some two-word utterances.
4	36–45	White	Associate/Divorced	$25,000-$49,999	56 months/ Male	36 months	ECSE, ABA, SLT, PT, OT	Gestures (reaching, leading), single words and signs.

ABA Applied behavior analysis, *ASD* Autism spectrum disorder, *ECSE* Early childhood special education, *SLT* Speech and language therapy, *OT* Occupational therapy, *PT* Physical therapy

Table 2 Child assessment data

Measure	Child-1	Child-2	Child-3	Child-4
MCDI				
Words Understood(out of 396)	159	88	190	67
Words Produced(out of 371)	65	23	45	17
ASQ-SE-2	130 (48-month cutoff: 85)	200 (48-month cutoff: 85)	145 (48-month cutoff: 85)	150 (60-month cutoff: 95)
GARS-2 Autism Disorder Index (6 subscales)	60 (Probable ASD; Level 1)	89 (Very likely ASD; Level 2)	71 (Very likely ASD; Level 2)	94 (Very likely ASD; Level 2)

ASD Autism spectrum disorder, *ASQ SE-2* Ages and stages questionnaire social-emotional-2, *GARS-2* Gilliam autism rating scale-2, *MCDI* MacArthur-bates communication development Inventory; Words Produced = Number of reported expressive words

Data were collected across all phases to examine changes in level, trend, variability, overlap, and immediacy of effect following introduction of the intervention.

Procedures

The study consisted of three primary phases: baseline, intervention, and maintenance. Across dyads, participants completed between 3 and 7 baseline sessions, 6–7 intervention sessions, and 1–2 maintenance probes across approximately 10–12 weeks.

Baseline Phase

During baseline, parents engaged in shared book reading as they typically would without receiving instruction or feedback or having access to PaiCoach. Parents recorded 3–5-minute videos and uploaded them to a secure institutional Box folder. Baseline continued until stable patterns of parent implementation and child communication were observed.

Intervention Phase

The intervention consisted of an individually delivered Zoom-based i-PiCSS training and AI-supported (PaiCoach) coaching.

Training. Following baseline, parents participated in a one-time training session adapted from the i-PiCSS program. The session lasted approximately one hour and was delivered via Zoom. It introduced reading prompts and three naturalistic communication strategies (modeling, mand-model, time delay). Training included explicit instruction and video demonstrations. In addition to the live training session, parents had ongoing access to embedded training materials within the PaiCoach platform, including brief instructional videos, step-by-step strategy descriptions, and a visual flowchart outlining key interaction sequences (e.g., establishing joint attention, presenting a strategy, waiting for a response, and responding contingently). See S2 in Supplemental Materials for screenshot examples.

AI-Supported (PaiCoach) Coaching. After training, parents began using PaiCoach. Upon video upload, the system generated a notification to the research team. The uploaded video was processed using the PaiCoach AI model to analyze parent strategy use and child communication behaviors. After AI finished its processing of the videos with the results, the first author and trained research team members reviewed and validated the AI-generated outputs on strategy detection and feedback accuracy before releasing the results to parents. Next, parents were notified that feedback about their implementation of intervention strategies was available. PaiCoach's feedback was time-anchored to specific moments within the video and included both strengths and suggestions for improvement. See Fig. 1 for the overall PaiCoach system flow.

Maintenance Phase

The maintenance phase was conducted approximately three weeks after the intervention phase ended. Except for Parent 4, who uploaded only one maintenance video due to family circumstances, all other parents recorded and uploaded two maintenance videos to PaiCoach. In this phase, the feedback features were disabled and parents did not receive any feedback.

Dependent Variables (DVs)

Primary DVs included parent implementation fidelity and children's communicative responsiveness derived from video-recorded reading sessions.

Parent DVs

Parent implementation fidelity was calculated using a 4-point fidelity scoring system applied to each instance of strategy use (i.e., modeling, mand-model, and time delay).

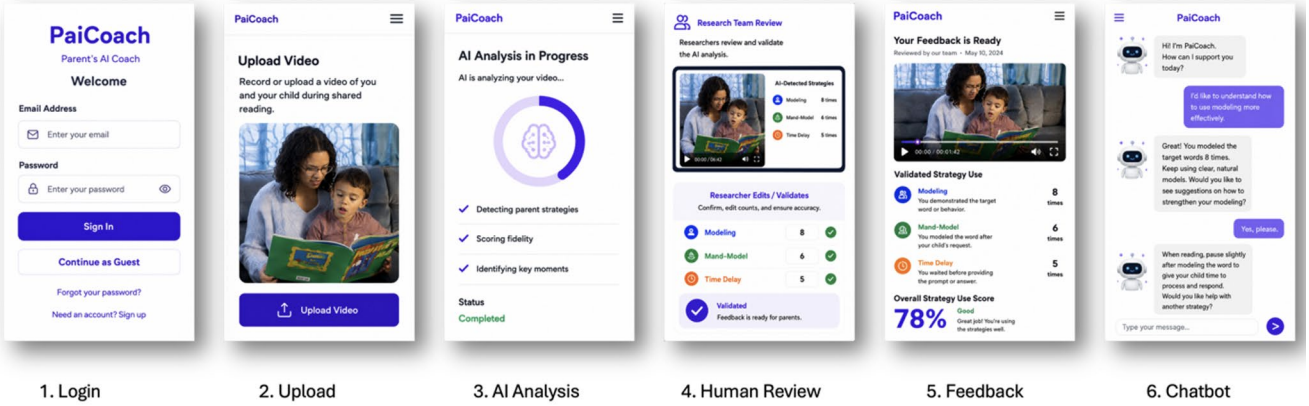


Fig. 1 PaiCoach overall interaction workflow. The figure illustrates the process from parents uploading interaction videos to the system's AI-generated analyses, including human-in-the-loop process and opportunities for parent reflection and interaction with the conversational agent

Overall fidelity DV was operationalized as the percentage of highest-quality implementations (score of 4) relative to the total number of scored strategy instances within a session (i.e., total number of 4 s divided by total number of all scores [1, 2, 3, 4] x 100). This approach was selected to be consistent with prior i-PiCSS fidelity procedures (Akemoglu et al., 2022a, 2022b). The fidelity scores reflect the extent to which parents implemented the strategies with high quality (see S3 in Supplementary Materials for fidelity scoring).

Child DV

Child DV included communicative responsiveness during reading. Child response rate was calculated as the percentage of child responses relative to parent-provided opportunities, operationalized as the total number of responses divided by the number of parent-provided opportunities (i.e., instances scored as fidelity 2, 3 or 4) x 100. Child response rates were calculated overall across all strategies (modeling, mand-model, and time delay). Child responses included verbal and nonverbal behaviors (e.g., vocalizations, words, gestures) directed toward the parent's use of a strategy.

PaiCoach Development and Evaluation

PaiCoach was developed based on the ASD-HI benchmark dataset and AI pipeline (Li et al., 2025), which were derived from naturalistic parent-child shared-reading videos collected through the PiCSS and i-PiCSS studies. The ASD-HI dataset contains 478 expert-labeled parent strategy-use instances extracted from 48 real-world shared-reading sessions across three families, with annotations for modeling, mand-model, time delay, and corresponding fidelity scores. The dataset was partitioned into training ($n=239$), validation ($n=120$), and test ($n=119$) sets for model development and evaluation.

Strategy classification was based on the observed interaction sequence rather than an autonomous determination of which strategy a parent should use. For example, modeling was identified when a parent provided a verbal language model without first requiring a child response, whereas time delay was identified when the parent created a communication opportunity and intentionally paused to allow the child an opportunity to initiate or respond before providing additional prompting or a model. Thus, the model learned to differentiate strategies from expert-coded examples of their observable interactional features and temporal sequences rather than independently determining that a time delay was clinically preferable to modeling at a particular moment.

The PaiCoach backend was built upon the AI pipeline described by Li et al. (2025), which integrates automatic speech recognition (ASR), multimodal video analysis, and LLM-based prompting to identify parent strategy use, estimate fidelity scores, and generate preliminary feedback. For the version evaluated in the present study, the backend was iteratively refined by replacing the original Whisper ASR component with Qwen3 and updating the classification and feedback model from GPT-4o to GPT-5.1 to improve transcription quality and behavioral classification performance. During internal evaluation of this updated backend, the strategy detection model achieved approximately 82% recall, and the GPT-based multimodal fidelity assessment model achieved 68.6% accuracy.

PaiCoach DV Detection and Metrics

The PaiCoach system identified strategy instances and generated preliminary fidelity scores using multimodal analysis of parent-child interaction data. These outputs were then reviewed and edited by the research team for accuracy prior

to final metric calculation. While the parent-facing interface provided parents with access to their own fidelity scores as well as timestamped and summary feedback, the researcher-facing interface presented session-level metrics derived from coded interaction data for monitoring purposes. These included (a) overall parent implementation fidelity, (b) strategy-specific fidelity (time delay, modeling, and mand-model combined), and (c) child response rate. The system then generated a longitudinal visualization (single-case design graphs) for human-edited data in which each parent-specific visualization distinguished baseline and intervention phases and supported interpretation of participants' performance changes over time.

Human Validation Procedures for AI-Generated Outputs

The purpose of the human validation procedures was to verify and refine AI-generated outputs prior to delivery so that parents received expert-reviewed feedback. Therefore, the present study evaluated PaiCoach as a human-in-the-loop coaching system rather than a fully autonomous AI system. Therefore, all PaiCoach-generated annotations and scores were reviewed by trained human coders across all phases, and feedback delivered during the intervention phase was validated by human reviewers prior to parent access. Human reviewers examined AI-generated outputs for accuracy in (a) temporal segmentation of interaction events, (b) identification of communication strategies, (c) fidelity scoring, and (d) generated feedback statements. When discrepancies were identified, reviewers corrected strategy labels, adjusted timing markers, modified fidelity scores, and refined generated feedback prior to parent access. Edits were tracked and aggregated across sessions. For each participant, the proportion of codes requiring modification (editing) and those requiring no modification (unedited) were calculated to characterize alignment between PaiCoach outputs and human judgment within a human-in-the-loop framework (see Table 3).

Reliability of Human Validation

Agreement between AI-generated outputs and expert review was evaluated by determining whether AI-generated outputs required modification following independent expert review. The first author conducted the primary review of all AI-generated outputs. Prior to independent review, a graduate assistant (GA; second author) received training on the coding procedures using previously coded parent-child interaction videos and practice AI-generated outputs until agreement exceeded 80% across two consecutive practice sessions. Next, the GA independently reviewed randomly assigned 20% of sessions and corresponding AI-generated outputs across study phases.

Agreement was defined as both reviewers making the same decision regarding AI-generated codes (i.e., whether a code required modification and the resulting corrected score). Disagreements were resolved through discussion and consensus review between the first author and the GA prior to finalization of the validated outputs. Agreement was calculated using point-by-point agreement by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100. Mean agreement across sessions and outputs was 92.8% (range = 88%–98%).

Table 3 Summary of human edits to AI-generated coding

Participant	Unedited	Edited
Dyad-1	68.1%	31.9%
Dyad-2	72%	28%
Dyad-3	71.2%	28.8%
Dyad-4	58.6%	41.4%
Mean	68.6%	31.4%

Traditional Interobserver Agreement (IOA)

IOA was calculated separately from the AI-validation procedures to assess reliability of the observational coding system used for parent fidelity and child communication variables. Across all three phases, IOA was calculated for approximately 33% of sessions and was conducted between the first author and the GA (second author). Agreement was calculated using point-by-point agreement, defined as the number of agreements divided by the total number of agreements plus disagreements multiplied by 100. Mean IOA for parent fidelity was 91% (range = 85%–97%), and mean IOA for child communication variables was 89% (range = 83%–95%).

Data Analysis

Visual analysis was the primary data analysis method as it is consistent with single-case experimental design (SCED). Visual analysis focused on changes in level, trend, variability, immediacy of effect, overlap, and consistency of data patterns across similar phases (Ledford & Gast, 2018). In addition, Tau-U values were used to support visual analysis findings and effect sizes were calculated to quantify non-overlapping data between baseline and intervention phases (Parker et al., 2011).

Social Validity and Usability

To assess social validity, parents participated in pre- and post-intervention semi-structured interviews conducted via video conferencing. All interviews were conducted by

a GA using Zoom and were audio- and video-recorded with participant consent. Interviews lasted approximately 20–30 min and explored parent perceptions of PaiCoach, including usability, acceptability, and perceived effectiveness of the intervention. Zoom's automatic transcription feature was used to generate initial transcripts, which were reviewed by the research team for accuracy. Following transcription, the research team developed structured summaries of each interview to synthesize parent responses across key domains (e.g., goals, procedures, outcomes). In addition to interviews, parents completed the System Usability Scale (SUS) and a researcher-developed post-study social validity questionnaire.

Results

A multiple baseline across participants design was used with staggered introduction of PaiCoach across the four dyads. We used visual analysis to examine changes in parent implementation fidelity and child communicative responsiveness across baseline and intervention phases for each dyad. Figure 2 depicts human-validated/edited graphs.

Dyad 1

During baseline, parent implementation fidelity for Parent 1 (P1) was characterized by a low-to-moderate level, variable responding, and no consistent upward trend, with scores ranging from 0% to 44.4%. Although one baseline session demonstrated moderate fidelity, performance decreased during later baseline sessions. Similarly, child responsiveness during baseline was variable as the data ranged from 25% to 44.4% without stable increasing trend across the three baseline sessions.

A clear immediacy of effect was observed for P1's implementation fidelity upon the introduction of PaiCoach-supported coaching. P1's fidelity increased substantially during the first intervention session to levels above baseline performance and remained elevated throughout intervention. P1's fidelity during intervention ranged from 57.1% to 88.9% and demonstrated a generally increasing trend with reduced variability compared to baseline phase. Her intervention data demonstrated minimal overlap with baseline and replicated intervention effects.

Child 1 (C1) demonstrated a corresponding increase in his responsiveness following the introduction of intervention (PaiCoach). His responsiveness increased from baseline levels to intervention levels as they ranged from 58.3% to 92.3% with an upward trend across sessions and relatively stable performance during later intervention sessions. Although some variability was observed during early intervention, C1's responsiveness consistently exceeded baseline ranges and approached high levels during later sessions. Dyad-1's data indicated preliminary evidence of maintenance for both variables with parent fidelity ranging

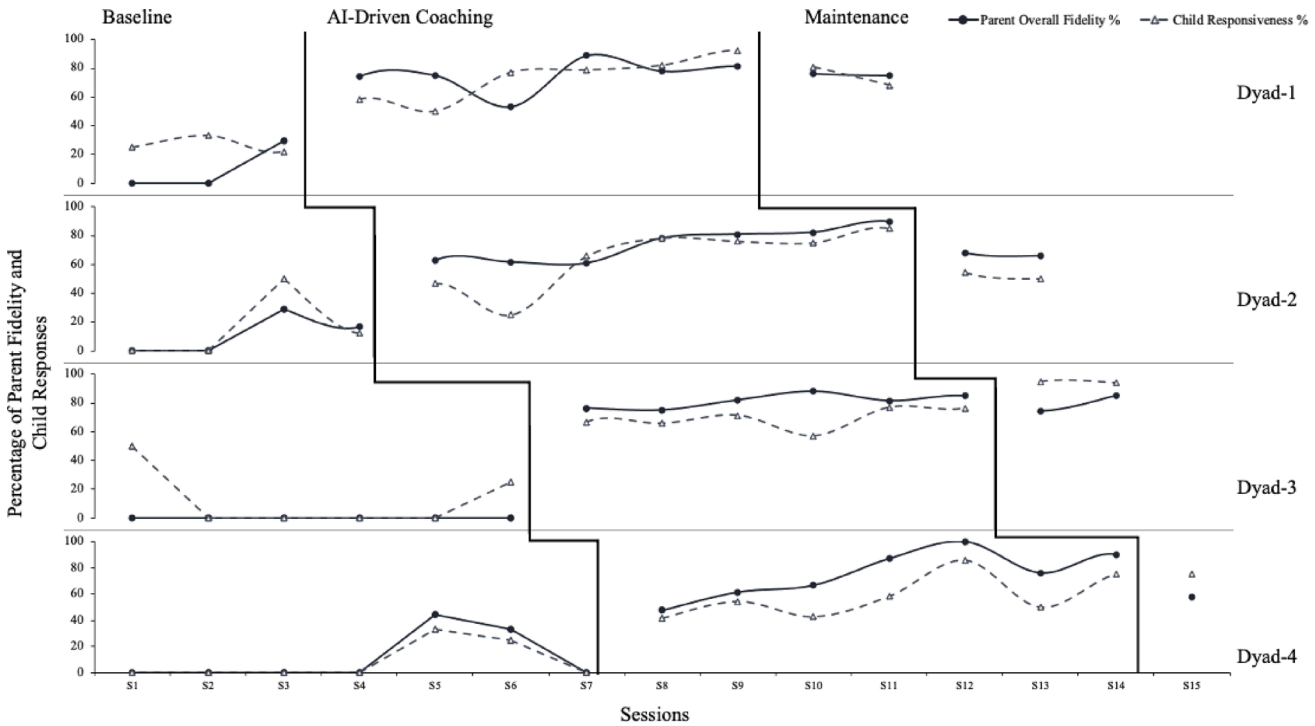


Fig. 2 Percentage of parent implementation fidelity and child responses. This graph displays the human-edited data. Session labels on the x-axis correspond to the session sequence for the bottom dyad only and are included as visual reference points

from 75% to 81.6% and child responsiveness ranging from 68.2% to 75%.

Dyad 2

During baseline, parent implementation fidelity for Parent 2 (P2) remained low overall with scores ranging from 0% to 28.6%. Her baseline data demonstrated a low level, minimal upward trend, and moderate variability. Child 2 (C2)'s responsiveness during baseline was variable as it was between 0% and 50%. Although one session demonstrated elevated responsiveness, subsequent sessions decreased substantially which resulted in an overall unstable and inconsistent baseline pattern.

Following the introduction of PaiCoach, P2's implementation fidelity demonstrated an immediate increase in level. Her fidelity scores during intervention ranged from 60.8% to 89.5% and were characterized by a positive trend and reduced variability as opposed to baseline. Following an initial increase, P2's fidelity remained consistently elevated throughout intervention with minimal overlap between baseline and intervention phases. C2's responsiveness also increased in the intervention phase. His responsiveness during intervention ranged from 25% to 85% with higher and more stable performance observed during later sessions. Dyad 2's maintenance data demonstrated sustained gains as parent fidelity remained between 65.8% and 67.9% and child responsiveness between 50% and 54.5%.

Dyad 3

During baseline, parent implementation fidelity for Parent 3 (P3) showed a low and stable pattern with no evidence of targeted strategy implementation prior to intervention and remained at 0% across all baseline sessions. However, Child 3 (C3)'s responsiveness during baseline showed variability. His baseline data initially occurred at approximately 50% before decreasing to 0% across subsequent sessions. P3's intervention data demonstrated a strong immediacy of effect following the introduction of PaiCoach. Her fidelity increased from 0% during baseline to 76.4% during the first intervention session and fidelity scores ranged from 74.3% to 88.2%. Overall, P3 demonstrated a stable, high-level pattern with minimal variability throughout intervention. Similarly, C3's responsiveness increased following the introduction of PaiCoach and ranged from 57.1% to 94.7% with an overall increasing trend and reduced variability relative to baseline. Dyad 3's maintenance data remained elevated for both variables with parent fidelity at 85.2% and child responsiveness at 93.8%.

Dyad 4

During baseline, Parent 4 (P4)'s parent implementation fidelity remained at or near 0% across most sessions with two sessions demonstrating moderate increases to 44.4% and 33.3%, respectively. Her baseline data resulted in an overall low level, variable pattern, and lack of consistent upward trend. Child 4 (C4)'s responsiveness during baseline similarly remained low overall and showed considerable variability without stable improvement.

Following the introduction of PaiCoach, P4's fidelity demonstrated a clear increase in level and a positive upward trend as her scores ranged from 47.6% to 100% with later sessions demonstrating consistently high performance. Also, P4's variability decreased during intervention relative to baseline and intervention data mostly exceeded baseline ranges. In parallel, C4's responsiveness also demonstrated increases in level and trend following intervention onset as it was between 41.7% and 85.7% during intervention. Later intervention sessions demonstrated higher and more stable levels of responsiveness corresponding with increased parent fidelity. A single maintenance probe was collected due to family sickness. In that data point, parent fidelity was 57.8% and child responsiveness was 75%.

Human Validation and AI Alignment

The human validation findings indicated moderate-to-strong alignment between PaiCoach-generated outputs and expert human review (see Table 3). Across 53 reviewed sessions and 879 coded instances, approximately two-thirds (69%) of AI-generated outputs did not require modification following human review, 31.4% of the outputs were edited. Domain-specific edit rates varied. For example, the system had lower human validation for time detection (28.1%) and strategy detection (25.1%) and higher rates for fidelity scoring (37.7%) and feedback outputs (34.5%). Also, there were participant-level differences in editing rates. For example, P4 demonstrated the highest overall edit rates across sessions, whereas P3 demonstrated lowest editing rates. Several sessions across participants demonstrated very low edit rates and indicated stronger alignment between PaiCoach outputs and human judgment (see S4 for editing data in Supplementary Materials).

Importantly, although PaiCoach relied on human validation, the system substantially reduced the manual effort associated with observational coding and feedback generation. Based on our implementation experience across previous PiCSS and i-PiCSS intervention studies, manually coding a 5-min parent-child shared-reading session typically requires repeated viewing, pausing, coding, and note taking, taking approximately 20–30 min for an experienced coder and up

Table 4 Comparison of parent coaching delivery models: traditional in-person, video-based, and PaiCoach AI-supported coaching

Feature	Traditional in-person coaching	Video-based coaching	PaiCoach AI-supported coaching
Coaching model	Parent and child participate in scheduled coaching with a professional, typically in person	Parent records parent-child interactions and subsequently reviews the video with a professional	Parent uploads a brief interaction video; PaiCoach generates automated, time-anchored analysis and feedback
Interval between opportunities for feedback	Typically constrained by scheduled appointments; may range from days to weeks depending on provider and family availability	Typically constrained by scheduling a subsequent video-review/coaching session; may range from days to weeks	Potentially minutes with fully automated processing; with human-in-the-loop validation, feedback can be provided asynchronously without requiring simultaneous parent-professional availability
Parent and professional scheduling	Requires parent, child, and professional availability at the same time	Requires parent and professional availability at the same time for video-review	Does not require simultaneous availability. Parents can upload videos and review feedback asynchronously at times convenient for them
Approximate time associated with a coaching/feedback episode	Potentially 1–2 h when transportation, preparation, observation, and feedback are considered*	Approximately 30–60 min for professional-parent video review and feedback*	AI processing occurs automatically. In the present study, human verification/editing of AI-generated coding and feedback typically required approximately 5–10 min per 5-minute video; parents could subsequently review feedback at their own pace*

*The present study also did not prospectively record processing or reviewer time. The reported 5–10-minute PaiCoach human-review estimate is based on the research team's implementation experience

to 60 min for a newly trained coder. In contrast, PaiCoach automatically analyzed a video, generated transcripts, preliminary behavioral coding, fidelity scores, and draft parent feedback within a few minutes. This allowed us (reviewers) to focus on verification and editing of AI-generated outputs, which typically required approximately 5–10 min per session. Table 4 compares PaiCoach with traditional in-person and video-based parent coaching models.

Tau-U Results

Human-Validated

Tau-U analyses supported the visual analysis findings. Tau-U values for parent implementation fidelity were 1.00 across all four dyads. This indicates complete nonoverlap between baseline and AI-supported coaching phases. Child responsiveness Tau-U values ranged from 0.72 to 1.00, with a mean of 0.84, indicating large to very large intervention effects.

Social Validity

Goals

During preintervention interviews, all parents indicated that they wanted their children to increase functional communication, including using more words, responding during interactions, and engaging more consistently during shared reading. In addition, parents expressed a desire for accessible and practical support that could help them implement communication strategies more effectively within everyday routines. For example, P1 stated, "I want him to communicate more, even if it's just using more words or trying to repeat things during reading." P2 emphasized engagement, noting, "I would love to see him be more involved with books instead of just flipping pages or losing interest quickly." Similarly, P4 described a broader goal of interaction, explaining, "I want him to interact more with me when we read, not just listen but actually respond or participate." Although parents did not initially describe the use of AI-supported tools, their goals reflected a need for ongoing and nonjudgmental feedback. Consistent with this need, P2 noted, "I know it [AI] won't judge me and I will have access to it all the time." P1 similarly highlighted the value of AI-generated feedback and said, "I want [PaiCoach] to notice areas where I can improve and also to tell me things that I did well."

Procedures

All four parents reported generally positive perceptions of the procedures of PaiCoach while also identifying areas for improvement. Parents highlighted the usefulness of PaiCoach feedback in supporting their implementation of communication strategies. For example, P2 noted, "It notices… areas where you can improve, and it also tells you things that you did well, so it is instructive criticism." Similarly, P4 described how feedback supported her practice, "Going through and just seeing… little areas that I missed that I could have used to increase participation… it was definitely useful."

P3 also emphasized the accessibility and flexibility of PaiCoach and said, "I had access whenever I needed it. I uploaded the videos at night right before bedtime. That was convenient for us." She further explained that, after managing daily caregiving responsibilities, attending additional appointments or traveling to receive coaching support was

often difficult due to fatigue and scheduling demands. In this regard, the ability to access PaiCoach asynchronously within home routines allowed her to engage with coaching support at times that were realistic and manageable for her family. These comments suggest that PaiCoach may help reduce practical barriers associated with traditional service delivery models for busy families. In addition, parents expressed a preference for concise, time-anchored feedback. For example, P2 noted, "I definitely used the tips and points more… I just didn't really have time to watch the whole video again," which indicated that targeted summary feedback was more feasible within everyday routines. At the same time, parents reported challenges related to the usability of the platform. P2 explained, "I wouldn't say it was super easy when it came to actually uploading the video… you kind of have to go through the tabs and remember where it is," and further noted confusion with navigation, as she stated, "I'm like, okay, am I ending? Am I restoring? Where do I upload the video?"

Outcomes

All four parents reported positive perceptions of the outcomes of the PaiCoach intervention and noted increases in their children's engagement and participation. For example, P2 explained that "staying on one topic for a little longer… definitely supported him in participating." This highlighted how the use of intervention strategies influenced her sustained interaction. P1 explicitly linked these gains to strategy use and said, "Because of that, he definitely repeats things more often," which demonstrated that changes in parent interaction supported by PaiCoach contributed to increased child responding. Similarly, P3 described how PaiCoach-generated prompts supported her child's language development, she said, "the suggestions it [PaiCoach] tailors… helped… being able to ask him different things that I might not have thought to ask," and that "it [PaiCoach] definitely got him talking more."

Quantitative Social Validity and Usability Findings

In addition to interviews, parents also completed a post-study social validity survey and SUS. The PaiCoach survey assessed perceptions of intervention goals, procedures, and outcomes using a 5-point Likert scale. Overall, parents reported high levels of satisfaction across survey domains. Across four parents, mean ratings were high for training clarity and accessibility (M = 4.5–5.0), feasibility of procedures (M = 4.3–4.7), and usefulness of PaiCoach strategies for supporting child communication goals (M = 4.5–5.0). Parents also reported strong satisfaction with their child's progress (M = 4.3–4.7) and their own confidence and competence in using communication strategies (M = 4.5–5.0). Ratings for overall usefulness and likelihood of recommending PaiCoach were high (M = 4.7–5.0). SUS findings further supported these results. PaiCoach yielded a mean SUS score of 77.5 (range = 72.5–82.5), indicating good usability and acceptability. These findings demonstrate that PaiCoach was found to be socially acceptable.

Discussion

We examined the effects of PaiCoach, an AI-supported parent coaching system, on parent implementation fidelity and autistic children's communicative responsiveness within shared reading. Our findings provide preliminary evidence of a functional relation between the introduction of PaiCoach and improvements in parent fidelity across dyads with corresponding gains in child communication. These findings extend the literature on parent-mediated intervention by demonstrating that AI can be leveraged to enhance the delivery of coaching and feedback within everyday routines (Akemoglu et al., 2020; Perry et al., 2026).

PaiCoach as a Human-in-the-Loop AI Coaching System

A central contribution of this study is the positioning of PaiCoach as a human-in-the-loop, AI-supported coaching system that augments parent implementation rather than replacing human involvement in intervention delivery (Meadan et al., 2026). Consistent with emerging perspectives in educational AI, effective systems are those that operate within human-in-the-loop frameworks, where AI supports decision-making, monitoring, and feedback while humans remain the primary agents of interpretation and action (Amershi et al., 2019; Rakap, 2024).

In the current study, PaiCoach used multimodal AI to generate preliminary analyses of parent strategy use and child communication behaviors, which were subsequently reviewed and validated by trained human reviewers prior to feedback delivery. This distinction is important because parent-child interactions involve substantial contextual and behavioral complexity that remains difficult for fully autonomous systems to interpret reliably (Koegel et al., 2025). Thus, the present findings are best interpreted as preliminary evidence supporting the feasibility of AI-augmented parent coaching rather than autonomous AI-delivered intervention.

Parent Fidelity and AI-Supported Feedback

The observed improvements in parent implementation fidelity are consistent with extensive research demonstrating that

coaching with performance-based feedback is necessary to achieve and maintain high levels of fidelity (Akemoglu et al., 2022a; Meadan et al., 2016). Traditional parent coaching models, however, are constrained by limited opportunities for observation and feedback, often resulting in insufficient support between sessions.

As an emerging AI tool, PaiCoach provides preliminary findings to address this gap by using AI to generate structured, time-anchored feedback based on recorded interactions, enabling iterative cycles of practice and refinement. These findings are consistent with prior work demonstrating that AI-supported systems can improve implementer fidelity by providing continuous, context-sensitive guidance and supporting procedural accuracy following the introduction of AI-based coaching systems (Griffen et al., 2024).

Child Communication and AI-Supported Interaction

PaiCoach did not directly deliver intervention to children, instead, it influenced child outcomes indirectly by enhancing the quality and consistency of parent implementation. This indirect pathway is particularly significant in the context of AI-supported intervention (Perry et al., 2026). Much of the existing literature has focused on technologies designed to teach children directly, such as applications targeting emotion recognition or social skills (Feng et al., 2025). In contrast, PaiCoach represents a shift toward supporting the mechanisms of interaction, where AI enhances the learning environment rather than replacing human partners. This approach is supported by emerging research on AI-based interventions targeting social communication. For example, Koegel et al. (2025) demonstrated that an AI system using a LLM improved empathetic responding in autistic individuals with gains generalizing to natural conversations. Together, these findings suggest that AI can support communication outcomes both by direct instruction and by enhancing the interactional context.

Human Validation and AI-Supported Coding

Human validation findings provide important insight into the current capabilities and limitations of AI-supported coding. Across sessions, the majority of PaiCoach-generated codes did not require modification which means that the system can approximate human judgment in detecting parent strategy use and child communication behaviors. At the same time, the proportion of edited codes demonstrates the continued importance of human oversight since automated systems may struggle with ambiguity and nuanced interpretation (Bommasani et al., 2021).

Closer examination of edits demonstrated that many discrepancies were associated with limitations in underlying multimodal components, including automatic speech recognition and computer vision. For example, gesture-based child responses were occasionally missed when visual information was partially obscured (e.g., book positioning), and overlapping speech or brief response latencies affected the detection of parent strategies and fidelity scoring. In addition, variability in natural language use (e.g., differences in acknowledgment or praise statements such as “good job” vs. “okay”) contributed to misclassification of strategy components despite correct implementation. These instances resulted in fidelity scores being reduced from “Fidelity 4” to “Fidelity 3.” These findings support the use of a human-in-the-loop framework, in which AI assists with initial analysis while human reviewers ensure accuracy and contextual interpretation (Meadan et al., 2026). Human-centered AI frameworks emphasize that effective systems augment, rather than replace, human decision-making by integrating human expertise into the evaluation process (Amershi et al., 2019).

Emerging Evidence on AI in Autism Intervention

While much of the existing work on AI in this field has relied on feasibility studies (Akemoglu et al., 2026; Kotsi et al., 2025), the use of an experimental design in the current study allows for demonstration of functional relations at the individual level. The present study contributes to the evolving evidence base on AI in autism intervention by providing experimental data within a SCED method that remains underrepresented in the AI literature (Akemoglu et al., 2026). However, the findings highlight important areas for continued development in AI-supported intervention systems. These include improving the accuracy and autonomy of AI models, expanding datasets to better capture variability in real-world interactions, and refining feedback mechanisms (Li et al., 2025).

Limitations

An important limitation of the current study relates to the accuracy and autonomy of the PaiCoach system. Although multimodal AI was used to analyze parent-child interaction videos and generate preliminary feedback, all outputs were reviewed and validated by human researchers prior to parent delivery. While this human-in-the-loop approach increased the accuracy and contextual appropriateness of feedback, it also limits the extent to which PaiCoach currently functions as a fully autonomous system. Across reviewed sessions, approximately 69% of AI-generated outputs were retained without modification. These findings reflect both the emerging capabilities and limitations of multimodal AI for analyzing parent-child interactions.

Also, child communication behaviors appeared substantially more difficult for the system to classify reliably than parent strategy use. These discrepancies suggest that current multimodal AI systems may struggle to consistently interpret complex communication behaviors. Therefore, the findings should be interpreted as preliminary evidence supporting the feasibility of a human-in-the-loop AI-supported coaching system.

Future Directions

Future research should focus on improving the accuracy, robustness, and scalability of PaiCoach through advances in machine learning and data development. Increasing system accuracy will require larger and more diverse datasets of parent-child interactions to support model training and validation. Expanding the dataset could potentially allow the system to better capture variability in communication styles and contexts. In addition, the development of multimodal AI models that integrate visual, auditory, and linguistic data may enhance the system's ability to detect strategies and interpret interaction dynamics. Advances in LLMs and computer vision provide promising avenues for improving both feedback generation and contextual understanding.

Future work should also explore approaches to reducing reliance on human validation, such as implementing confidence-based thresholds, adaptive feedback systems, or hybrid review models, where human oversight is applied selectively. These developments would support more scalable and near real-time feedback delivery, which is critical for maximizing the practical impact of AI-supported coaching systems. Future studies should examine the effectiveness of AI-supported parent training and coaching systems across larger and more diverse populations, as well as its application across different routines and intervention targets. Comparative studies evaluating AI tools like PaiCoach relative to traditional coaching models would further clarify its role within service delivery systems.

Conclusion

This study provides preliminary evidence that PaiCoach, an AI-supported parent coaching system can improve parent implementation fidelity and support child communicative responsiveness within naturalistic home-based interactions. By integrating AI within a parent-mediated intervention framework, PaiCoach demonstrates the potential to enhance the precision, consistency, and accessibility of coaching while maintaining a human-in-the-loop approach.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10803-026-07519-6.

Acknowledgements The first author gratefully acknowledges support from the Scientific and Technological Research Council of Turkey (TÜBİTAK) through the 2219 International Postdoctoral Research Fellowship Program. The views expressed in this article are solely those of the author and do not necessarily reflect the views of TÜBİTAK.

Author Contribution Yusuf Akemoğlu led the study, including study design, implementation, data collection and analysis, and manuscript preparation. Zhaohui Li contributed to the development and implementation of the AI models, and Qingxiao Zheng contributed to the development of the PaiCoach user interface. Emily Brooke Roberts assisted with recruitment, data organization, and reliability coding. Samar Pratap Singh provided technical and backend support for PaiCoach. Jinjun Xiong supervised the overall study and served as principal investigator of the supporting grants. All authors reviewed and approved the final manuscript.

Funding This material is based upon work supported under the AI Research Institutes program by National Science Foundation and the Institute of Education Sciences, U.S. Department of Education, through Award # 2229873- National AI Institute for Exceptional Education. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the U.S. Department of Education. Similarly, this work also was supported in part by the Organization for Autism Research under its Applied Research Grant. The views expressed in this article are those of the authors and do not necessarily reflect the views of the sponsor.

Declarations

Conflict of interest The authors have no conflict of interest. Dr. Yusuf Akemoğlu is an Associate Editor for JADD but was not involved in the editorial process.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akamoglu, Y., & Meadan, H. (2019). Parent-implemented communication strategies during storybook reading. *Journal of Early Intervention, 41*(4), 300–320.

Akemoglu, Y., & Tomeny, K. R. (2021). A parent-implemented shared-reading intervention to promote communication skills of preschoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 51*(8), 2974–2987.

Akemoglu, Y., Muharib, R., & Meadan, H. (2020). A systematic and quality review of parent-implemented language and

communication interventions conducted via telepractice. *Journal of Behavioral Education*, *29*, 282–316.

Akemoglu, Y., Hinton, V., Laroue, D., & Jefferson, V. (2022a). Parent-implemented shared reading intervention via telepractice. *Journal of Early Intervention*, *44*(2), 190–210.

Akemoğlu, Y., Laroue, D., Kudesey, C., & Stahlman, M. (2022b). A module-based telepractice intervention for parents of children with developmental disabilities. *Journal of Autism and Developmental Disorders*, *52*, 5177–5190.

Akemoglu, Y., Roberts, E., & Xiong, J. (2026). A systematic review and quality appraisal of AI-augmented intervention studies for autistic individuals. *Autism, 30*(9), 2191–2207.

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., & Horvitz, E. (2019). Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–13).

Benway, N. R., & Preston, J. L. (2024). Artificial intelligence–assisted speech therapy for/ɹ: A single-case experimental study. *American journal of speech-language pathology, 33*(5), 2461–2486.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Choi, D., Park, S., Lee, K., Hong, H., & Kim, Y. H. (2025). Aaccesstalk: Fostering communication between minimally verbal autistic children and parents with contextual guidance and card recommendation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–25).

Feng, X., Guan, W., & Xu, E. (2025). The impact of generative artificial intelligence on the social-emotional skills of children with autism spectrum disorder: A meta-analysis of 38 experimental and quasi-experimental studies. *International Journal of Developmental Disabilities*. https://doi.org/10.1080/20473869.2025.2511661

Griffen, B., Lorah, E. R., Caldwell, N., Hantula, D. A., Nosek, J., Tincani, M., & Lemley, S. (2024). The effects of artificial intelligence on implementors' fidelity of instructional strategies during handwashing acquisition in children with autism. *Journal of Developmental and Physical Disabilities, 36*(5), 793–819.

Hinton, V., Akemoğlu, Y., Tomeny, K., & McWilliam, R. A. (2024). Supporting families from a distance: Implementing routines-based home visits via telepractice. *Early Childhood Education Journal, 52*, 629–636.

Holyfield, C., MacNeil, S., Caldwell, N., O'Neill Zimmerman, T., Lorah, E., Dragut, E., & Vucetic, S. (2024). Leveraging communication partner speech to automate augmented input for children on the autism spectrum who are minimally verbal: Prototype development and preliminary efficacy investigation. *American Journal of Speech-Language Pathology*, *33*, 1174–1192.

Hopcan, S., Polat, E., Ozturk, M. E., & Ozturk, L. (2023). Artificial intelligence in special education: A systematic review. *Interactive Learning Environments*, *31*(10), 7335–7353.

Koegel, L. K., Ponder, E., Bruzzese, T., Wang, M., Semnani, S. J., Chi, N., & Lam, M. S. (2025). Using artificial intelligence to improve empathetic statements in autistic adolescents and adults: A randomized clinical trial. *Journal of Autism and Developmental Disorders*, 1–17.

Kotsi, S., Handrinou, S., Iatraki, G., & Soulis, S. G. (2025). A review of artificial intelligence interventions for students with autism spectrum disorder. *Disabilities*, *5*(1), 7.

Kwon, T., Jeong, M., Ko, E. S., & Lee, Y. (2022). Captivate! contextual language guidance for parent-child interaction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–17).

Ledford, J. R., & Gast, D. L. (2018). *Single case research methodology: Applications in special education and behavioral sciences*. Routledge.

Li, Z., Akemoglu, Y., Lyu, J., Zheng, Q., & Xiong, J. (2025). ASD-HI: A parent-child interaction dataset for automated assessment of home intervention. In International Conference on Artificial Intelligence in Education (pp. 48–62). Springer Nature Switzerland.

Marino, M. T., Vasquez, E., Dieker, L., Basham, J., & Blackorby, J. (2023). The future of artificial intelligence in special education technology. *Journal of Special Education Technology, 38*(3), 404–416.

Meadan, H., Snodgrass, M. R., Meyer, L. E., Fisher, K. W., Chung, M. Y., & Halle, J. W. (2016). Internet-based parent-implemented intervention for young children with autism: A pilot study. *Journal of Early Intervention*, *38*(1), 3–23.

Meadan, H., Alqunaysi, R., Jennings, S., Douglas, S. N., Reilly, A. M., & Xiong, J. (2026). "Nothing can replace authentic human interaction": EI/ECSE professionals' perception of AI. *Topics in Early Childhood Special Education*. https://doi.org/10.1177/02711214261435935

Ousley, C. L., Raulston, T. J., & Gilhuber, C. S. (2025). Telecoaching for parents of young autistic children using strength-based video feedback. *Journal of Autism and Developmental Disorders*, *55*(1), 14–29.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*(2), 284–299.

Perry, N., Boulton, K. A., Hankin, L., Roisenberg, B. B., & Guastella, A. J. (2026). A qualitative study investigating caregiver perspectives of an artificial intelligence assistive device to support daily activities in families with children with autism spectrum disorder. *Digital Health, 12*, Article 20552076251411228.

Rakap, S. (2024). Navigating the role of artificial intelligence in special education: Advantages, disadvantages, and ethical considerations. *Practice*, *6*(2–3), 99–104.

Rakap, S., & Balikci, S. (2024). Enhancing IEP goal development for preschoolers with autism: a preliminary study on ChatGPT integration. *Journal of Autism and Developmental Disorders*, 1–6.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A., & Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(8), 2411–2428.

Vivanti, G., Kasari, C., Green, J., Mandell, D., Maye, M., & Hudry, K. (2018). Implementing and evaluating early intervention for children with autism: Where are the gaps and what should we do?. *Autism research, 11*(1), 16–23.

Waterfield, D. A., Watson, L., & Day, J. (2024). Applying artificial intelligence in special education: Exploring availability and functionality of AI platforms for special educators. *Journal of Special Education Technology, 39*(3), 448–454.

Zheng, Q., Akemoglu, Y., Li, Z., Singh, S., Nassereldine, A., & Xiong, J. (Under review). PaiCoach: A family-centered AI coach for parents' home interventions with children of developmental differences. *The ACM CHI Conference on Human Factors in Computing Systems).*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.