AI-Driven Scenario-Based Perspective-Taking Intervention (AISPT).

A 12-session, individually delivered, voice-activated chatbot programme targeting cognitive and emotional perspective-taking in preschool-aged children through structured narrative dialogues and guided prosocial problem-solving.

AI-Driven Scenario-Based Perspective-Taking Intervention (AISPT).

A 12-session, individually delivered, voice-activated chatbot programme targeting cognitive and emotional perspective-taking in preschool-aged children through structured narrative dialogues and guided prosocial problem-solving.

Why:

The intervention is grounded in socio-constructivist and social–emotional development theories. Drawing on Vygotsky's (1978) concept of the zone of proximal development, the scenario-based chatbot is designed to function as a scaffolded social partner that guides children through age-appropriate narrative dilemmas.
The underlying rationale is that structured story prompts, emotion labelling, and role-taking exercises provide opportunities for children to practise both cognitive and emotional perspective-taking in a supportive and safe context. Cognitive perspective-taking tasks were based on false-belief paradigms; emotional perspective-taking tasks engaged children in emotion inference, causal reasoning, and prosocial problem-solving. By engaging children in reflective dialogue and prosocial problem solving, the intervention aims to strengthen early social reasoning, enhance empathic understanding, and promote the foundational components of social–emotional learning (SEL) in early childhood. The framework integrates three theoretical axes: (1) social-cognitive development (Vygotsky, 1978; Piaget, 1932), (2) SEL (CASEL, 2023), and (3) human–technology interaction (Reeves & Nass, 1996).

What (material):

Core digital tool: A commercially available, web-based generative AI chatbot operating in Turkish, delivered via a tablet or laptop. The chatbot communicated with the child **directly through voice output**; the child responded by speaking aloud. The facilitator did not mediate this communication.
Scenario scripts: Twelve structured scenario scripts covering social–emotional dilemmas sequenced by complexity (sessions 1–12). Scenarios spanned six thematic domains: (1) Inclusion & Loneliness, (2) Misunderstanding & Apology, (3) Fairness & Generosity, (4) Intentions & Consequences, (5) Emotional Regulation & Forgiveness, (6) Inclusion & Social Advocacy, (7) Miscommunication & Clarification, (8) Self-Regulation & Social Awareness, (9) Hurt Feelings & Intentionality, (10) Perspective Differences & Compromise, (11) Trust & Repair, and (12) Loss, Change & Transition. Scenario 1 served as the worked model for all sessions. Each script specified the chatbot's voice prompts, the facilitator's observational role, and tiered scaffold prompts for use when a child did not respond.
Facilitator materials: (a) S1: Scenario Scripts — 12 session scripts with turn-by-turn chatbot prompts, facilitator notes, and scaffold prompts. (b) S2: Facilitator Flowcharts — 12-session implementation timeline, universal session flow diagram, situational decision guide, and five-level scaffolding ladder. (c) S3: Fidelity Checklist — per-session checklist (Sections A–I), programme-level summary tracking sheet, and fidelity calculation guide.
Training materials: A structured facilitator orientation (60–90 minutes) comprising: theoretical background on perspective-taking and SEL; the standardised session flow; the voice-based chatbot interaction model; scaffolding procedures; and at least one

supervised practice session. Facilitators were required to reach ≥85% fidelity before delivering independently.
All scenario scripts, facilitator tools, and fidelity checklists are available as supplementary materials (S1, S2, S3). If required by the journal, these materials can also be hosted in a public repository (e.g., OSF).

What (procedures):

Each intervention session followed a structured, developmentally aligned sequence:
Step 1 — Welcome & Orientation (1–2 min): Facilitator greeted the child and directed them to the device. Chatbot opened the session with a warm-up prompt ("Are you ready?").
Step 2 — Scenario Delivery (15–18 min): The chatbot presented the scenario narrative by voice and guided the child through the following phases: (i) narrative presentation, (ii) emotion inference ("How do you think [character] feels?"), (iii) causal reasoning ("Why do you think that?"), (iv) self-connection (personal experience link), (v) prosocial repair ("What would you do?"), (vi) outcome prediction, and (vii) closing affirmation.
Step 3 — Session Closing (2 min): Chatbot delivered a closing statement by voice. Facilitator affirmed the child's effort and returned them to the classroom routine.
Sessions were conducted individually, lasted 20–25 minutes, and followed identical sequencing across all 12 sessions. Post-session, the facilitator completed the fidelity checklist and recorded observation notes.

Who provided:

The intervention was delivered by facilitators with backgrounds in early childhood education or developmental psychology.
All providers completed a structured training session (60–90 minutes) covering: (1) the theoretical basis of perspective-taking and SEL, (2) the standardised session flow and scenario content, (3) the voice-based chatbot interaction model and child–chatbot dynamics, (4) the scaffolding ladder and tiered support procedures, and (5) the fidelity monitoring protocol.
Each facilitator conducted at least one supervised practice session before independent delivery and was required to meet the fidelity threshold (≥85%) to proceed.

How (mode of delivery; individual or group):

The intervention was delivered face to face in the preschool setting on an individual basis.
Each child interacted directly and individually with the voice-activated chatbot on a tablet or laptop. The chatbot spoke aloud to the child; the child responded by speaking directly to the chatbot. Sessions were conducted in real time through voice-based dialogue.
A trained facilitator was present throughout each session to: (a) observe the child's engagement and welfare, (b) monitor for distress or adverse reactions, and (c) deliver scaffolding support when a child did not respond after approximately 15 seconds, following the tiered scaffolding protocol. The facilitator did not mediate the communication between child and chatbot and did not speak on the child's behalf.

Where:

The intervention took place in a quiet, low-distraction area within the preschool environment, typically a small guidance room or a designated classroom corner equipped with a table and chairs appropriate for individual child–device interaction.
Sessions required only basic infrastructure: a tablet or laptop with a stable internet connection, functioning microphone, and appropriate audio output. No additional room equipment was required.

When and how much:

he intervention was delivered over a 6-week period (June–July 2025), with children participating in two sessions per week, for a total of 12 sessions. Each session lasted approximately 20–25 minutes, yielding a total individual exposure time of 240–300 minutes per child.
One scenario was delivered per session. Scenario content was sequenced from lower to higher complexity across the 12 sessions. The structure and duration of each session were kept consistent across the programme to maintain standardisation.

AI-Driven Scenario-Based Perspective-Taking Intervention (AISPT)

A 12-session, individually delivered, voice-activated chatbot programme targeting cognitive and emotional perspective-taking in preschool-aged children through virtual dialogue and developmentally appropriate scaffolding.

Tailoring:

Yes. The intervention included planned, developmentally appropriate adaptations.
The first assessment week (Week 1) was used for baseline data collection. The 12 intervention sessions were delivered across Weeks 2–7. Post-intervention assessments were administered during Week 8.
Although the core structure and scenario content were identical for all participants, the chatbot's generative responses were contextually adapted to each child's spoken input, allowing clarification when misunderstandings occurred and providing simplified or supportive language when needed. This responsiveness was an inherent property of the generative AI model rather than a manual adjustment.
The facilitator could also offer brief verbal scaffolding or emotional support in cases of confusion or heightened affect, following the standardised scaffolding ladder (S2). Five escalating support levels were defined, from extended wait time (Level 1) to session suspension for child welfare (Level 5). These adjustments ensured that each child received an appropriate level of support while maintaining fidelity to the standardised session framework.

Modification:

Yes. Minor, unavoidable modifications occurred during the intervention due to technical constraints.
Specifically, occasional latency in the chatbot's voice responses and brief internet interruptions required facilitators to pause the session and resume when connectivity was restored. On a small number of occasions, the chatbot's voice recognition did not detect the child's speech, requiring the child to repeat their response or the facilitator to offer a brief cue to repeat.
These adjustments did not alter the core structure, content, or objectives of the intervention. All modifications were documented in session fidelity logs (S3, Section D), including the nature of the disruption, its duration, and the corrective action taken.

How well (planned):

Fidelity monitoring procedures were established prior to data collection and are described in the Session Fidelity Checklist (Supplementary Material S3).
Session-level monitoring: Facilitators completed a fidelity checklist (S3) after each session, documenting adherence to all mandatory procedural steps across nine sections: (A) Administrative information, (B) Pre-session setup, (C) Scenario delivery fidelity, (D) Technical disruptions, (E) Child engagement observations, (F) Scaffold use log, (G) Safety and welfare, (H) Overall fidelity rating, and (I) Qualitative field notes.
Independent audit: Approximately 10% of sessions (randomly selected) were reviewed by an independent observer using the same checklist to verify adherence and ensure inter-rater consistency.
Fidelity threshold: Sessions were classified as protocol-adherent when ≥85% of mandatory steps were completed as specified. Programme-level fidelity was calculated as the mean of all 12 session-level fidelity scores.
Fidelity criteria: Completing all core scenario phases; maintaining session duration within 20–25 minutes; providing only standardised, level-appropriate scaffolding support; and not mediating child–chatbot communication.

How well (actual):

Fidelity was monitored throughout the intervention using the procedures described under Item 11.
Facilitators completed session-level fidelity checklists after each of the 12 sessions. Review of fidelity logs indicated that sessions were implemented as planned, with no substantive procedural deviations. Minor adjustments (e.g., brief pacing modifications in response to technical latency) were documented but did not alter core intervention components.
All 15 participants attended all 12 sessions; no post-baseline attrition occurred. Session duration was maintained within the 20–25 minute window across all sessions.
The planned fidelity threshold (≥85% adherence) was met. Detailed fidelity data are available in the session-level fidelity logs submitted as Supplementary Material S3.