Computers in Human Behavior Reports 20 (2025) 100799

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Robot-assisted handwriting training: An intervention for children with neurodevelopmental disorders

Jianling Zou [a], Soizic Gauthier [b,d], Dominique Archambault [a], Mohamed Chetouani [c], David Cohen [b,c,*], Salvatore M. Anzalone [a,**], the iReCheck Study Group [1]

[a] Laboratoire de Cognitions Humaine et Artificielle (CHArt), Université Paris 8-Vincennes-Saint-Denis, Saint-Denis, France
[b] Service de Psychiatrie de l'Enfant et de l'Adolescent (SPEA-PSL), Hòpital de la Pitié-Salpêtrière, AP-HP, Paris, France
[c] Institut des Systèmes Intelligents et de Robotique (ISIR), Sorbonne Université, CNRS, Paris, France
[d] Forward College, Paris, France

A R T I C L E I N F O

Keywords:
Social robotics
Neurodevelopmental disorders
Long-term interaction
Learning-by-teaching
Handwriting and dysgraphia

A B S T R A C T

Social robots offer promising possibilities for special education, particularly in supporting children with complex neurodevelopmental disorders (NDDs) who often experience learning difficulties, such as challenges with handwriting skills. However, most studies to date have been short-term and conducted in controlled lab settings. This study presents a robot-assisted intervention using a Wizard of Oz interface, the R2C3 system (Rehabilitation Robotic Companion for Children and Caregivers), to support handwriting reeducation in a real, long-term educational setting. The intervention adopted a learning-by-teaching scenario involving 18 children with NDD and dysgraphia, alongside 8 caregivers (teachers or occupational therapists), who participated in 9 handwriting-focused sessions facilitated through a serious game. Results showed strong engagement with low attrition (2 dropouts) and substantial improvement: 73.3 % of children improved their handwriting scores by at least one standard deviation on the BHK test. Improvement was significantly associated with engagement in the game and the nature of verbal exchanges, emphasizing the social robot's role in the design of our sessions. In a subsample of 11 children, sessions with the R2C3 system were compared with a session where the robot operated autonomously with minimal behaviors, further underscoring the impact of the robot's proactive behaviors on engagement and outcomes.

1. Introduction

Learning foundational academic skills such as reading, writing, and math is a challenge faced by school-aged children. For those with neurodevelopmental disorders (NDDs), these challenges are often amplified, impacting not only their educational progress, but also their overall cognitive and social development (Xavier & Cohen, 2020). This group of disorders, comprising conditions that typically manifest early in life, such as autism spectrum disorder, attention deficit hyperactivity disorder, or developmental coordination disorder, significantly affects the development of children and has significant repercussions not only on their personal and social sphere but also on their academic functioning (American Psychiatric Association, 2013). Moreover, as comorbidities between NDD are very common (Xavier & Cohen, 2020), the diverse combinations of disorders and their varying degrees of severity result in highly distinct profiles that require individually tailored support. In this context, technology-based solutions for NDD are considered promising approaches to provide engaging, personalized, and adaptive learning (Scassellati, 2007).

Socially assistive robotics - i.e., the use of robots with a focus on engaging individuals in interpersonal manners to assist them (Breazeal et al., 2016) - combined or not with the use of serious games, has the potential to efficiently support children with NDD and their caregivers in facing a number of difficulties, such as improving their social skills (E. Y. Chung et al., 2024; Scassellati et al., 2018), reeducating their handwriting (Gargot et al., 2021; Palsbo & Hood-Szivek, 2012), eliciting their body awareness (Costa et al., 2015). However, most studies have been exploratory and limited to short-term interactions (Kouroupa et al.,

* Corresponding author. Service de Psychiatrie de l'Enfant et de l'Adolescent (SPEA-PSL), Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France.
** Corresponding author.
E-mail addresses: david.cohen@aphp.fr (D. Cohen), sanzalone@univ-paris8.fr (S.M. Anzalone).

[1] The members of the iReCheck Study Group are listed in Appendix A.

https://doi.org/10.1016/j.chbr.2025.100799
Received 24 March 2025; Received in revised form 5 September 2025; Accepted 6 September 2025
Available online 12 September 2025

2022). Although some elicited great excitement on social media, they offer little clinical or educational value (Grossard et al., 2018). In the largest study so far, 12 autistic children engaged in a home-based triadic interaction with a caregiver and a robot: for 30 min every day, for one month, children completed social skills activities provided by a serious game (Scassellati et al., 2018). Despite this, recent advances in the reliability of robot hardware, in the compliance of control systems to human presence, and in the ergonomics of user-friendly interfaces pave the way for real-world, long-term scenarios (Jung & Hinds, 2018) that right now are rare (van Straten, Peter, & Kühne, 2020). These developments can support the design of extended experimental settings, moving beyond punctual, cross-sectional investigations or short-term interactions limited to only a few sessions, towards real-world contexts in which robots adapt to unstructured environments and tailor their behaviours to the specific needs of both children and caregivers (Anzalone et al., 2019). Ultimately, this opens the way to longer interactions conducted in real environments, with real users and addressing real needs.

Difficulty with written production is a prevalent and significant challenge among children with NDDs, affecting a large portion of this population (Kushki, Chau, & Anagnostou, 2011; Yoshimasu et al., 2011). This difficulty with written production, which falls below expectations relative to an individual's chronological age, cognitive abilities, and educational level, is known as dysgraphia (American Psychiatric Association, 2013; Hamstra-Bletz & Blöte, 1993). Dysgraphia has been associated with lower self-perception, reduced self-esteem, and weaker social functioning (P. J. Chung et al., 2020), hence accurately portraying the social difficulties faced by children with NDDs. Children with dysgraphia may fall into a vicious circle by practicing writing less and less, consequently increasing the gap with typically developing children of the same age (Gargot et al., 2020). They may avoid writing because it is difficult for them and it provokes anxiety. As a result, they may further get discouraged and avoid writing practice; their self-esteem may be further reduced as well as their motivation. In extreme cases, children reject and stop writing altogether.

The use of social robots (Gargot et al., 2021) in a learning-by-teaching scenario can help break this vicious circle by enhancing motivation. When children teach a robot how to write, they take on the role of evaluator and advisor, shifting from being judged or critiqued to actively guiding their partner. In a single-case longitudinal study, Gargot et al. implemented the learning-by-teaching paradigm with the humanoid robot Nao, alongside the serious game Dynamilis[2] for 20 consecutive training sessions. The child could engage with the robot, reduce writing avoidance, increase commitment over time, and enhance handwriting skills (Gargot et al., 2021).

Building upon this single case and Scassellati's proposal of a triadic interaction between a child, a therapist, and a robotic platform combined with a serious game (Scassellati et al., 2018), we developed a new scenario to support children with complex NDD and dysgraphia by transitioning to the use of QTRobot[3] and by refining both the robot's social interactions and control accessibility (Zou et al., 2022, 2024). To ensure accessibility for caregivers (e.g., teachers, occupational therapists) without a background in computer science, we engaged in a collaborative design with caregivers and dysgraphic children to develop a Wizard of Oz interface (WoZ). Together with a library of robot behaviors, it allows tailored interactions with the serious games application Dynamilis, to meet the specific needs of children and the requirements of caregivers. The system, named R2C3 (Rehabilitation Robotic Companion for Children and Caregivers), is designed to support handwriting training sessions, conducted by specialized teachers or therapists, for children with NDDs and dysgraphia.

Here, we report a study involving 18 children with complex NDD and dysgraphia who received robot-assisted intervention using the R2C3 system to support the reeducation of handwriting skills in the specific context of a daycare center. This study seeks to determine: (i) whether the system improves children's writing skills and self-esteem; (ii) whether specific interaction traces between children, caregivers, and the robot during the sessions significantly influence children's writing evolution; and (iii) how the degree of personalization required in robot behaviors affects the writing training. To explore this last question, we randomly introduced a session with the robot only performing minimal stereotyped behaviors to support essential functions for maintaining the writing training process. To assess children's writing skills and self-esteem, we used the Concise Assessment Scale for Children's Handwriting (known by the acronym BHK) (Hamstra-Bletz et al., 1987) and the Self-Perception Profile for Children (SPPC) (Harter, 2012). Accordingly, we hypothesise that:

- Research question 1: Children's scores on the BHK and the SPPC will significantly increase between the pre- and post-test assessments.
- Research question 2: Differences in social interaction (as measured by verbal and visual interaction traces) will correlate with children's writing evolution (as measured by comparison of pre- and post-test BHK scores).
- Research question 3: Children's engagement during sessions conducted with a personalized social robot controlled by caregivers will be higher compared to engagement during sessions conducted with an autonomous robot proposing stereotyped behaviors.

2. Materials and methods

The presented robot-assisted intervention for handwriting training comprised 9 sessions for each child, each one lasting approximately 20 min. These training sessions involved triadic interactions between the social robot, the child, and the caregiver within a learning-by-teaching scenario (Fig. 1): the robot asked children to teach him how to successfully complete the serious game (Zou et al., 2022). In a first condition, caregivers were given instructions to have the robot behave as they prefer whenever they deemed it necessary or desirable, freely selecting games on the Dynamilis app, to engage the children in handwriting activities. In a second condition, a proactive autonomous robot performed stereotyped interactive behaviors. We set up data collection equipment which included two cameras: one positioned in front of the child to capture their facial expressions and another placed on a tripod next to the desk to record the desk area. To collect audio data related to the discussions between children, caregivers, and the robot, we use the ZOOM H4N Pro multi-track recorder. To synchronize all these different data collected, we exploited the ROS[4] data recording capabilities, creating a rosbag[5] file for each session synchronizing and saving the videos from two cameras, the audio from the recorder, and the system logs of the WoZ interface.

2.1. Study design

This study focuses on exploring the reeducation of handwriting in children with complex NDD and dysgraphia through a robot-assisted intervention combined with the serious game Dynamilis using the R2C3 system (Zou et al., 2022, 2024). Specifically, the study investigates three operational research questions. (RQ1) Does children's participation in 9 writing training sessions affect their handwriting skills and self-esteem? (RQ2) Which modalities of the triadic (child--robot-caregiver) interaction during the writing training sessions correlate with children's handwriting improvement? (RQ3) Does the degree

[2] Dynamilis, developed by School Rebound: https://dynamilis.com/fr/.
[3] QTRobot from LuxAI: https://luxai.com.
[4] ROS: https://www.ros.org/.
[5] rosbag: https://wiki.ros.org/rosbag.

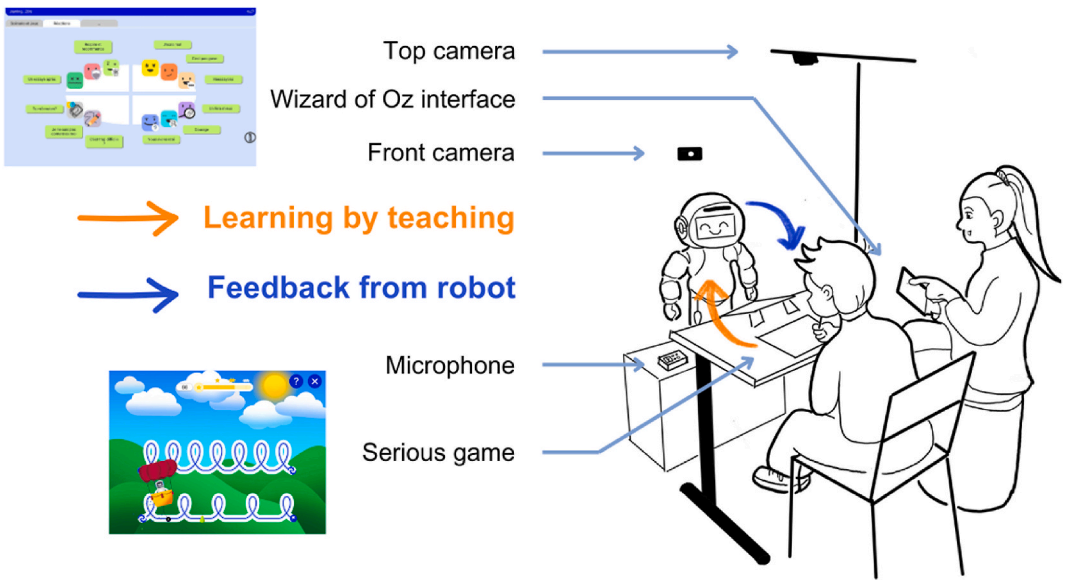

Fig. 1. Experimental Setup. The R2C3 system includes a social robot, an iPad with an Apple Pencil, a tactile tablet, two RGB cameras, and a microphone, designed to facilitate triadic interactions between the robot, child, and caregiver. In this study, children with dysgraphia engaged in handwriting-focused serious games on the iPad, while caregivers used the tablet's Wizard of Oz interface to control the robot and deliver feedback and guidance.

of robot personalization affect the triadic interaction during the writing training sessions? The design is open, non-blind, prospective, and longitudinal. For each participant, we continuously recorded a total of 9, 20-min writing training sessions, conducted once per week, that constituted the intervention. The study had two phases. Seven children participated in a pilot study (phase 1), and all nine experiment sessions were conducted with a caregiver-controlled robot displaying tailored behaviors (Zou et al., 2024). In phase 2, 11 new children participated in 9 sessions, including one session randomly scheduled with a QT robot having stereotyped behaviors to understand the impact of different degrees of robot personalization. The first and last sessions were excluded from randomisation, as the interaction during these sessions can be altered by the children discovering the robot or having to say goodbye. We first used randomisation without replacement to assign one of the seven remaining sessions in the Stereotyped Robot condition to each child, so that each session was assigned once to the first seven children. For the remaining four children, we repeated the randomisation procedure, again without replacement. An overview of the experimental process is depicted in Fig. 2.

During the proposed intervention for handwriting reeducation, we observed that children's performances during the first and last sessions were affected due to the excitement of encountering the robot for the first time or the emotions associated with parting and saying goodbyes. To mitigate the impact of this phenomenon on our results, we excluded the data from the first and last sessions of each participant in the analysis conducted.

2.2. Participants

Altogether, 18 children participated in this study (see Table S1 for a description of each participant). Participants were recruited in the Department of Child and Adolescent Psychiatry of the Pitié-Salpêtrière Hospital. Inclusion criteria were: (i) presenting a diagnosis of dysgraphia, associated with a diagnosis of complex NDDs (one or more comorbid NDD and the need for specialized outpatient day care center); (ii) being between 7 and 15 years old, and (iii) last but not least, being willing to participate in the study. Exclusion criteria were: (i) major non-psychiatric medical health issues; (ii) presenting persistent delusions or hallucinations; (iii) being frightened of or particularly aggressive towards robots. For each child, the diagnosis was based on all available medical information (including direct interviews and tests, family history data, and treatment records). The study was reviewed and approved by the ethics committee of Sorbonne University [ID: CER-2020-103] and included written parental consent and oral children consent.

Fig. 2 shows the flow diagram of the experimental process and the participant inclusion. Out of 18 participants, two children were excluded from the study: one child was excluded because he did not complete all the necessary sessions, as he was discharged from the hospital before he could do so, and another child was excluded because his participation was interrupted several times for long periods due to his condition.

An additional 5 children were excluded from the multimodal interactive feature analysis (Child1 to Child5 in Table S1): as part of the study was conducted during the COVID-19 period, these 5 children were wearing facial masks, which caused difficulties in face recognition and gaze analysis. However, the evolution of their writing and self-perception was still investigated.

Children were supervised during the writing training sessions by 8 caregivers. Each child participated in their sessions with one specific caregiver. This group of caregivers comprised two speech therapists, 3 psychomotor therapists, and 3 specialized teachers, and all of them but one were women. Their mean age was 36.6 years old (max 54, min 23), and their mean years of experience was 10.3 years (max 25, min 1).

Before their first session with a child, caregivers received training to become familiar with using the Wizard of Oz interface and R2C3. During this training session, caregivers first watch a video tutorial explaining how to use the Wizard of Oz interface.[6] Then, we asked them to complete the following tasks: make QTRobot welcome a child; show specific buttons on the Wizard of OZ interface; make QTRobot ask whether what it just did was correct; make QTRobot conclude a session and say goodbye; make QTRobot welcome another child; make QTRobot explain the goal/scenario of the sessions; make QTRobot ask for a specific game to be played; make QTRobot react to someone telling it "you really suck!" aggressively.

[6] A tutorial for the R2C3 system and the WoZ interface: https://www.youtube.com/watch?v=h7eIE5TxgVU.

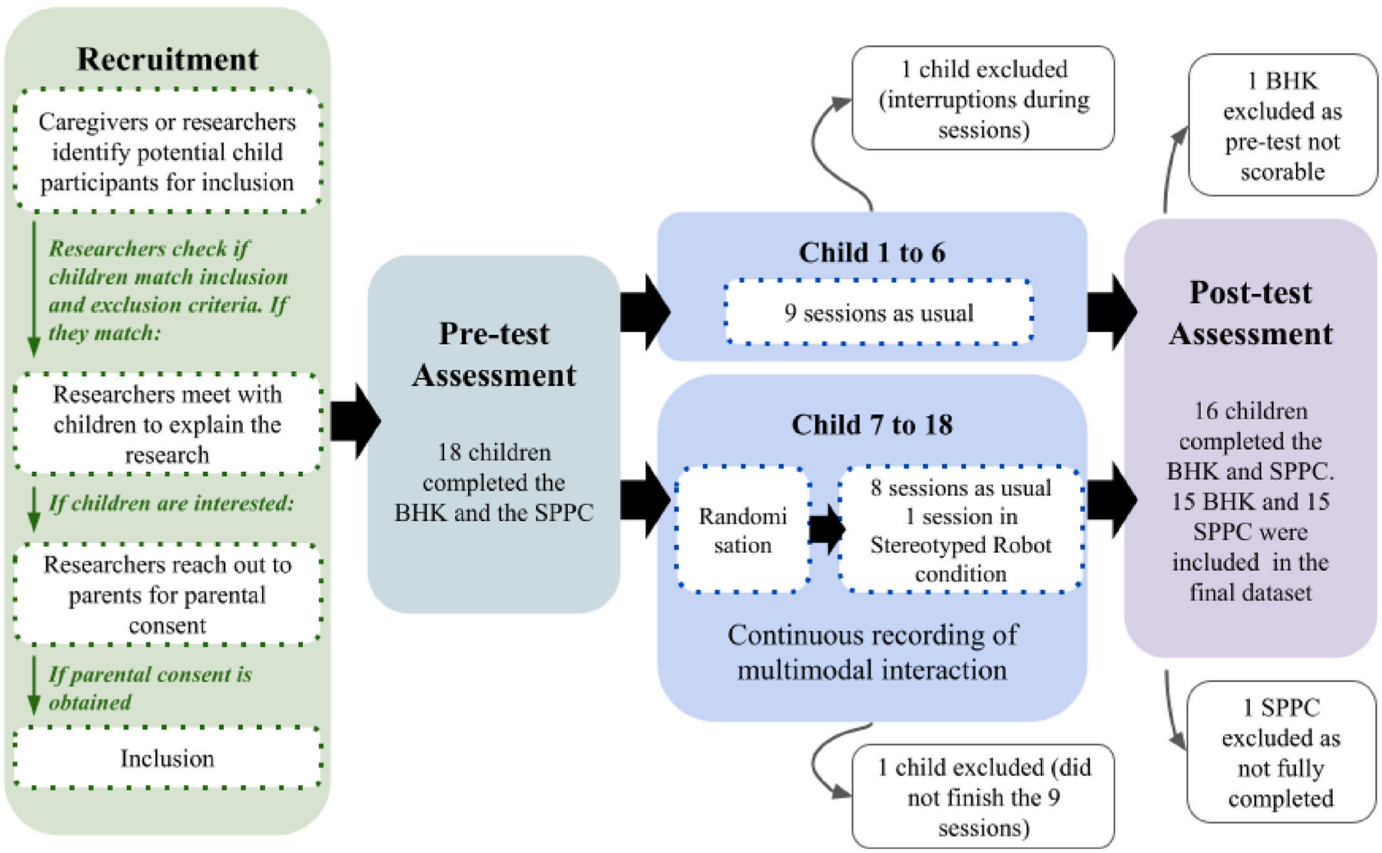

Fig. 2. Flow diagram of experimental process. A total of 18 child participants were initially enrolled, but two of them had to be excluded due to interrupted sessions. Additionally, five child participants were only included in the exploration of the first research question, since the data collected from these participants did not allow the exploration of the other research questions.

2.3. Children evaluation scales

To investigate whether children's participation in writing training sessions using our R2C3 system leads to improvements in children's writing skills and self-esteem, we used the Concise Assessment Scale for Children's Handwriting (BHK) (Hamstra-Bletz et al., 1987) and the Self-Perception Profile for Children (SPPC) (Harter, 2012) before the first session (pre-test), and after the last one (post-test). The BHK test is a widely used assessment of writing skills in European countries (Biotteau et al., 2019). In the context of BHK, participants are instructed to copy a text for 5 min on a white paper (Hamstra-Bletz et al., 1987). The therapist evaluates two scores: the first score is based on 13 criteria that assess the quality/legibility of the written content, while the second score is determined by the speed of the writing process (the number of letters written within 5 min). The diagnostic threshold for dysgraphia is established at two standard deviations below the standardized average performance for each grade level: a negative score represents difficulties in writing. The SPPC, a widely used questionnaire (Harter, 2012), comprises 36 items aiming at gauging self-esteem in children aged 8 years and older. This instrument yields scores across five distinct domains of self-esteem: scholastic, social, athletic, physical appearance, and behavioral, as well as global self-esteem scale: a high score translates to positive self-esteem. In this research, we focus on three specific domains of the SPPC: scholastic competence, social competence, and global self-esteem. The BHK test and SPPC questionnaire were administered as pre-tests before the initial session and as post-tests after the final session. All the BHK scores were assessed by psychomotor therapist IZ who participated in our study, after the texts produced were anonymized. The SPPC questionnaire's scores were calculated by researcher SG.

2.4. Serious game

We used Dynamilis, an application specifically designed for handwriting training and reeducation, installed on an iPad Pro 12.9 2020 and operated with an Apple Pencil. Dynamilis has been developed with specific algorithms to train pressure, tilt direction, kinematics, and smoothness when using a pencil (Asselborn et al., 2018) that allows personalization of training (Asselborn et al., 2020; Gargot et al., 2020). Dynamilis includes diverse serious games, all aiming at supporting the improvement of writing skills, in terms of the static, kinematic, pressure, or tilt features of writing. Static features directly refer to the letters and words' shape. Kinematic features refer to the dynamics of writing (writing speed, pause between words, between letters …). Pressure features capture the characteristics of the pressure exerted between the pen tip and the tablet surface, and tilt features encompass the attributes of pen tilt (Asselborn et al., 2018). The co-writer game, one of the serious games included in the Dynamilis app, is the activity that aligns the most closely with our learning-by-teaching scenario (Zou et al., 2022). In this game, the child assumes the role of a robot's teacher, instructing it on how to write. However, we made sure that the role of the robot and the learning-by-teaching situation were replicated in the other games as well thanks to specifically drafted instructions given during the writing training sessions by our robot (Zou et al., 2024).

2.5. Robot

For our study, we selected the social robot platform QTRobot designed by the Luxai company. Our decision was guided by its reliability and stability, making it suitable for longer sessions, and by the wide range of expressive capabilities it can display, in part thanks to the screen it is equipped with as its face. This socially expressive robot is specifically engineered to enhance the efficacy of rehabilitation and reeducation for children with autism (Grossard et al., 2018).

In our prior work, we collaborated with caregivers and children with dysgraphia to conceive and co-develop a system known as the Rehabilitation Robotic Companion for Children and Caregivers (R2C3) (Zou et al., 2022, 2024). This system incorporates a Wizard of Oz (WoZ) interface for robot teleoperation. The behavior repertoire encompasses 120 distinct robot behaviors, comprising facial expressions, speeches, and gestures. The WoZ methodology involves substituting any automated decision algorithm in the robot's controller with a manual selection of the robot's behavior. During the interaction, a human "wizard" (here, the caregiver), serving as the robot's operator, selects the most appropriate behavior for the robot based on its current state, its environment, and its ongoing interaction with human partners through an interface (Riek, 2012; Steinfeld et al., 2009, pp. 101–108). In this study (Fig. 3, top), we used the R2C3 system, controlled by caregivers. A typical WoZ session begins with a greeting phase, during which the caregiver selects a welcome behaviour. The caregiver is then free to initiate any game with the child, using the robot to deliver instructions, feedback, and encouragement, before, after, and during the game. The choice of the game, the transition to a new one, and the provision of feedback are left to the caregiver's judgment, allowing for continuous personalization and adaptation to the child's needs, engagement states, and responses. The session concludes at the caregiver's discretion.[7]

To examine the influence of the robot's social capabilities on children's engagement, we developed an alternate version of its behaviors, restricting them to the minimum stereotyped behaviors to support the functions essential for maintaining the writing training process. Fig. S1, in the supplementary material, displays the algorithm diagram followed in this Stereotyped Robot Condition. The session commences upon recognition of a child's face with the robot greeting the child. The robot is then limited to providing game instructions and feedback such as "Congratulations" or "Take courage and try again" based on game performance. Game selection and transitions are not personalized, with a change in the game occurring only after three consecutive successes or failures. The session concludes after five different games have been played or when the session duration reaches 20 min.[8] In this study (Fig. 3, bottom), the Stereotyped Robot Condition was exclusively employed during a singular session during phase 2 of the study (and thus limited to 11 child participants).

2.6. Multimodal features extraction

As aforementioned, during the experiments, different multimodal information such as video data, audio data, and log features, were recorded and synchronized using Rosbag. Video, audio, and contextual features (see Table 1 for more details), were then extracted and stored in a multimodal database.

Utilizing the front-facing video data, acquired from the USB camera fixed on the wall facing the child participant, we conducted an examination of gaze behavior. The extraction of gaze relies on the open-source library OpenFace[9] (Baltrusaitis et al., 2018, pp. 59–66), which generates head directions, thus enabling the estimation of an individual's gaze direction in radians in world coordinates. This enables us to determine whether a person is looking straight ahead, left or right, and up or down, and thus, whether a child is looking or not at the tablet, the caregiver, or the robot. Gaze values outside the regions of interest (e.g., when the child was looking at the ceiling), as well as frames with missing gaze estimates, were first labeled as Unknown Values and then discarded. Additionally, we applied voice activity detection (VAD) to the audio stream, assigning each frame to silence, speech, or overlapped speech; consequently, no unassigned values were observed. These computations were performed using the Pyannote Python open-source package[10] (Bredin et al., 2020, pp. 7124–7128). Based on the outputs of each modality, the time spent in gazing and vocalisation activities was calculated. Table 1 reports these features normalised with respect to the total duration of each session. Lastly, the density of robot behaviours for each session was calculated using the logs generated by the WoZ interface, which recorded the selected behaviour and its timestamp each time a button was clicked.

2.7. Statistical analysis

We performed all statistical analysis using R. To compare pre- and post-variables, we apple non-parametric Wilcoxon rank tests. To explore how the specific interaction traces correlated with children's writing evolution (Table 1), we conducted linear mixed-effects regression models. Each feature and outcome combination was subjected to an individual model. Specifically, we implemented models through the 'lme4' package in R. The model formula used was "variable(feature) ∼ scale(outcome) + (1|subject)". Given the observed deviations from the assumptions of classical linear regression, we opted for a robust approach. Using the ' boot ' package, we estimated 95 % confidence intervals and computed p-values through bootstrapping with 10,000 replications. The regression coefficient of linear mixed-effects regression can be interpreted as the partial correlation coefficient. Finally, to explore how robot session condition correlates with interaction traces, we used linear mixed-effects regression models, specifically implemented through the 'lme4' package in R with one model per feature using the formula: "scale(feature) ∼ robot condition + (1|subject)". All the models were applied upon prior verification of their respective assumptions.

3. Results

3.1. Evolution of handwriting skills and self-esteem

The BHK test and the SPPC scale were administered before the first session (pre-test), and after the last one (post-test) to verify eventual improvements in children's writing skills and self-esteem (Research question 1). As reported in Fig. 2, data from the two children who interrupted their sessions were fully excluded from the dataset. For the remaining 16 participants, both SPPC and BHK measures were included, even in cases of partial completion: SPPC scores were retained for the one child whose BHK was unscorable, and BHK scores were retained for the one child who did not complete the final SPPC. This resulted in a final sample of 15 SPPC and 15 BHK measures. Fig. 4 and Table 2 summarizes the statistics and distribution of BHK variables (speed and quality) and SPPC scores (scholastic scale, social scale, and global scale scores) in pre- and post-session assessments. Qualitatively, among the 15 participants, 9 (60 %) children improved by at least one standard deviation in BHK quality or speed scores without experiencing a decline in the other, while 2 (13.3 %) other children improved one BHK score but decreased in the other (see Fig. 5 for examples). This can still be considered as a general improvement in writing abilities given the relatively limited duration of our study, as the development of writing speed and quality is not necessarily parallel, and improvements in quality with limited improvement in speed, or vice-versa, is typical during children's writing acquisition (Feder & Majnemer, 2007). Additionally, it has been reported that improvements in writing speed or writing quality can induce a momentary delay in the other

[7] A demonstration of the R2C3 system and the WoZ interface: https://www.youtube.com/watch?v=iZzBAZbiSVA.

[8] A demonstration of the Stereotyped Robot Condition: https://www.youtube.com/watch?v=9i9lJmiJnAU.

[9] OpenFace: https://github.com/TadasBaltrusaitis/OpenFace.

[10] PyAnnote Audio: https://github.com/pyannote/pyannote-audio.

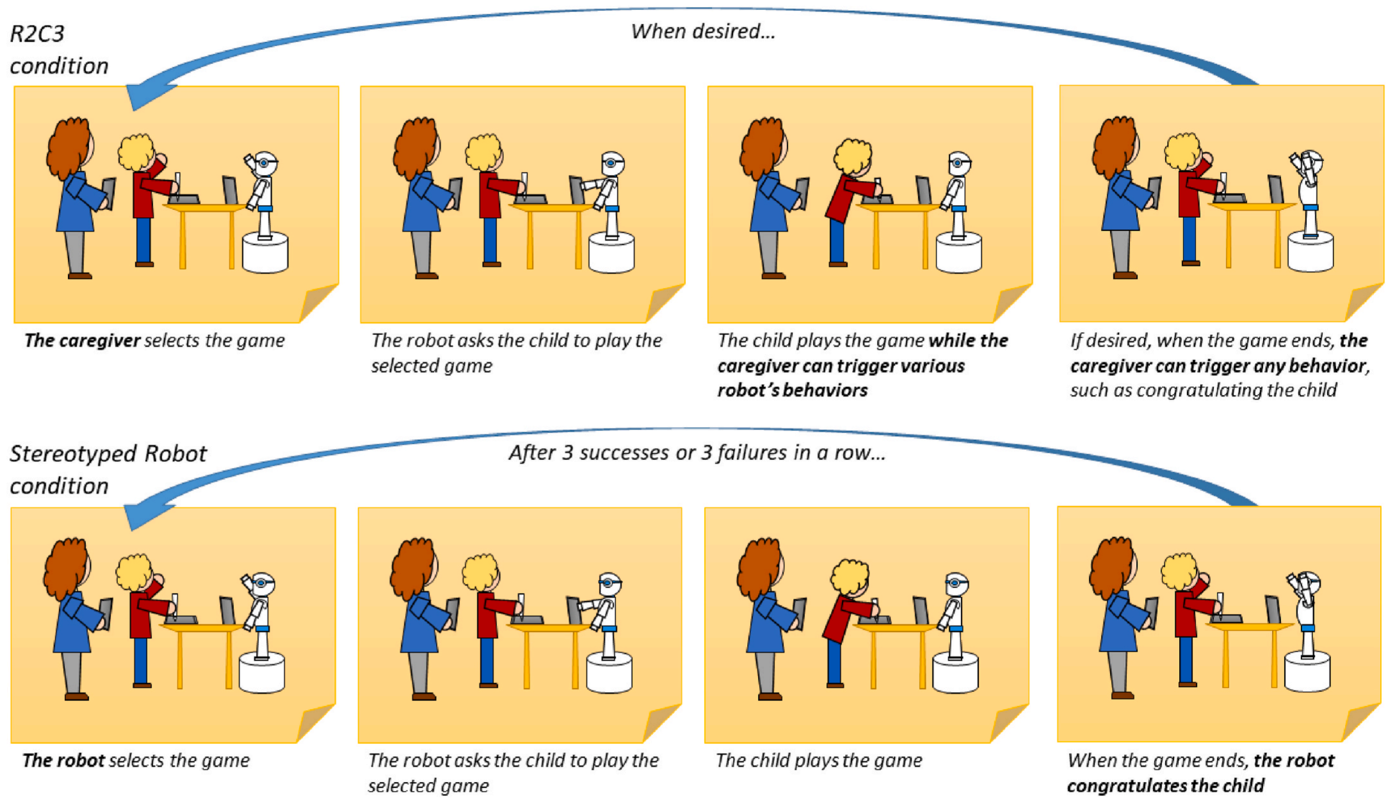

Fig. 3. The robot behavior in the two conditions. On the top, the teleoperated R2C3 condition, where the caregiver continuously controls and personalizes the behavior of the robot according to the child's needs; on the bottom, the stereotyped robot condition, in which behaviors are selected autonomously.

characteristics (Gargot et al., 2021) as part of the typical writing acquisition trajectory. However, we found no statistically significant difference between the pretest and post-test scores for the two measures and a considerable variability among children (Fig. 4 A-E).

Fig. S2 shows the correlation between pretest scores and the writing or self-esteem evolution. We found a discernible decreasing trend ($\rho = -0.5$, $p = 0.06$) between the pretest score on the BHK Speed Scale and the writing speed evolution. Child participants with the most severe writing difficulties in speed showed higher levels of improvement post-test. We also observed a statistically significant decreasing correlation ($\rho = -0.71$, $p = 0.004$) between the initial scores on the SPPC Global Scale and the global self-esteem evolution. Child participants with lower global self-esteem manifested higher levels of improvement post-test.

3.2. *Correlation between interactive traces with the robot and writing evolution*

To investigate how children, caregivers, and the robot interact during the handwriting training sessions and how the specific traces of the interaction correlate with children's writing evolution (Research question 2), we collected a variety of multimodal features - as listed and detailed in Table 1. To identify the main interaction traces within the triad, we performed a Principal Component Analysis (PCA). The top 2 principal components account for a substantial portion of the data's variance, with a variance-explained ratio of 86.25 %. Fig. 6 presents the loadings of the top 2 principal components. We observed that in the first principal component (PC1), the percentage of time per session during which the child gazes at the tablet (T_pad) carries a higher weight (65.91 %), indicating that it primarily captures variations related to the percentage of time the child is playing with Dynamilis. The second principal component (PC2) is dominated by the percentage of time per session during which a member of the triad (the child/the caregiver/the robot) is speaking (T_speechActivity) and the percentage of time per session when all the members remain silent (T_silence), which have higher weights (29.35 % and 40.72 % respectively). This suggests that PC2 mainly encompasses the dynamics of speech turn-taking. Taken together, PC1 and PC2 underline that the triadic interaction during the session mainly differs from one triad to another in terms of time spent by the child looking at the tablet and in terms of discussion activity. This confirms the importance of social interactions during the educative process which translates into the importance of collecting a multimodal dataset comprising visual and audio information.

To explore how the specific interaction traces correlate with children's writing evolution (Table 1), after confirming that model assumptions were met, we conducted linear mixed-effects regression models. Each feature and outcome combination was subjected to an individual model. Models are displayed in Table 3. The writing evolution is significantly correlated to the duration of speech activity ($\beta = -0.5$, $p = 0.018$). When the duration of verbal exchange within a triad decreases by approximately 0.5 standard deviations, the child's average writing evolution increases by one standard deviation. Although the effect on T_pad and T_silence is not statistically significant ($p = 0.07$ and $p = 0.051$, respectively), there exist trends: for a decrease of 0.46 standard deviations in the time a child spends engaged in the serious games, the child's average writing evolution increases by one standard deviation. Additionally, when the duration of silence tends to rise by 0.44 standard deviations, the child's average writing evolution increases by one standard deviation.

3.3. *Impact of robot's social capabilities*

To investigate the impact of the robot's social capabilities on engagement with the robot (Research question 3), we created two different conditions highlighting the type of behaviors displayed by the robot. Among the 9 sessions, one was conducted with a Stereotyped Robot and was randomly distributed for each child among 7 out of 9

Table 1
Multimodal features for the analysis.

Modality	Feature	Description	Source
Interactive traces			
Visual	Gaze at Robot (T_robot)	The percentage of time per session during which the child gazes at the robot.	Autom. extracted
	Gaze at Caregiver (T_caregiver)	The percentage of time per session during which the child gazes at the caregiver.	Autom. extracted
	Gaze at Tablet (T_pad)	The percentage of time per session during which the child gazes at the tablet.	Autom. extracted
Audio	Speech Activity (T_speechActivity)	The percentage of time per session during which a member (the child/the caregiver/the robot) is speaking.	Autom. extracted
	Speech Overlap (T_overlapSpeech)	The percentage of time per session during which the speech of at least 2 members overlaps.	Autom. extracted
	Long Pause (T_longPause)	The percentage of time per session when there are brief pauses (lasting 1–2 s) in the speech.	Autom. extracted
	Short Pause (T_shortPause)	The percentage of time per session when there are brief pauses (lasting 0.5–1 s) in the speech.	Autom. extracted
	Silence (T_silence)	The percentage of time per session when all the members remain silent (more than 2 s).	Autom. extracted
Contextual	Density of behaviors (density_behavior)	The ratio of the quantity of behaviors to the total duration (behaviors per second).	Logs
	Total Duration	The total duration of the session (in seconds).	Logs
	Session Type	The type of robot control (teleoperated or stereotyped) during the session.	Logs
Clinical Outcomes			
	Writing Evolution	The average difference between pretest and posttest scores on both BHK scales (quality and speed).	Scored by experts
	SPPC-scholastic Scale Evolution	The variance in scores between the pretest and posttest on the scholastic scale of SPPC.	Scored by experts
	SPPC-social Scale Evolution	The variance in scores between the pretest and posttest on the social scale of SPPC	Scored by experts
	SPPC-global Scale Evolution	The variance in scores between the pretest and posttest on the global scale of SPPC	Scored by experts

sessions, excluding the first and last sessions. The Stereotyped Robot contrasted with the R2C3 system because it lacked the high level of personalization provided by caregivers' control through the WoZ interface. To explore how robot session condition correlates with interaction traces, we used linear mixed-effects regression models. The models reveal significant differences between the two robot conditions for several interaction traces including T_pad, T_shortPause, T_longPause, T_speechActivity, T_silence, and density_behavior. Fig. 7 displays the differences between the two conditions for each feature. Details can be found in Table 4.

The Stereotyped Robot Condition creates a significantly higher density of behaviors compared to the R2C3 Condition. Taken together, the percentage of speech activity time with short and long pause times can be considered as a description of the vocal interaction dynamics among the child participant, caregiver, and robot. Interestingly, these metrics are significantly reduced in the Stereotyped Robot Condition, presenting a paradox with the increased density of robot behaviors: in this condition, the robot displays a higher number of behaviors, but the amount of vocal interaction is lower. Furthermore, the percentage of time that child participants gaze at the tablet is also significantly lower in the Stereotyped Robot Condition, indicating reduced engagement in the serious game.

The results underline the unique behavioral patterns recorded during the two different conditions.

4. Discussion

Socially assistive robotics aims to provide personalized, on-demand, and structured reeducation to augment the efforts of teachers, parents and clinicians (Belpaeme et al., 2018). Here, we set up a triadic scenario with a child, QT robot, and Dynamilis under the control of a practitioner using R2C3 to offer a writing intervention to children with complex NDD and dysgraphia. We address three fundamental research questions, each bearing significant relevance. First, we sought to ascertain the impact of this writing intervention on children's writing skills and self-esteem. Second, we explored the intricate dynamics of interactions between children and the robotic system during the writing training sessions, examining how interaction traces correlate with writing evolution. Lastly, we explored the level of personalization required in the robot's behaviors to facilitate effective handwriting reeducation. Our findings shed light on the potential of robot-assisted interventions for improving the handwriting skills of children with dysgraphia. The low attrition rate observed throughout the study underlines both the feasibility and acceptability of the proposed intervention, and points to the robust engagement of participants and caregivers. Beyond these outcomes, our study offers several noteworthy contributions: (i) it was conducted in a long-term, real-world child–robot–caregiver setting, supported by an interdisciplinary collaboration at the intersection of robotics, education, and clinical sciences; (ii) it involved the collection and exploitation of real-world multimodal data in a sustained interactive scenario; (iii) it provided an analysis of how the social capabilities of the companion robot influenced children's engagement and performance. Together, these contributions demonstrate the value of moving beyond short-term laboratory studies towards more ecological, interdisciplinary, and impactful approaches.

4.1. Social robot motivates Children's participation

From our experience with this study, it appears that the robot, with its playful aspect and thanks to the learning-by-teaching paradigm, can act as a motivator for children to practice writing, even when this task is particularly difficult for them. This is highlighted by the descriptions made by the children - several of our participants stated to the researchers during their participation in the study that working with a robot motivated them to engage in writing training sessions (Zou et al., 2024). This is of particular importance when working with children who tend to avoid writing - and who may consequently see their writing skills deteriorate even further.

Only 2 out of 18 children (11 %) either failed to complete the 9 sessions of the intervention or took extended breaks from the intervention. It is hard to state whether this dropout rate is better or worse compared to other interventions, as dropout rates are rarely reported in studies describing interventions designed to support the writing skills of children with dysgraphia. Working with a similar population, Yanjana and Kumar reported a drop-out rate of 22 % for a 3-month behavioral intervention targeting writing skills in children with dysgraphia (Yanjana et al., 2020), whereas Mehta and Nandgaonkar reported a drop-out rate of 9 % for a visual-perceptual training (Mehta & Nandgaonkar, 2019). Additionally, factors external to the intervention *per se*

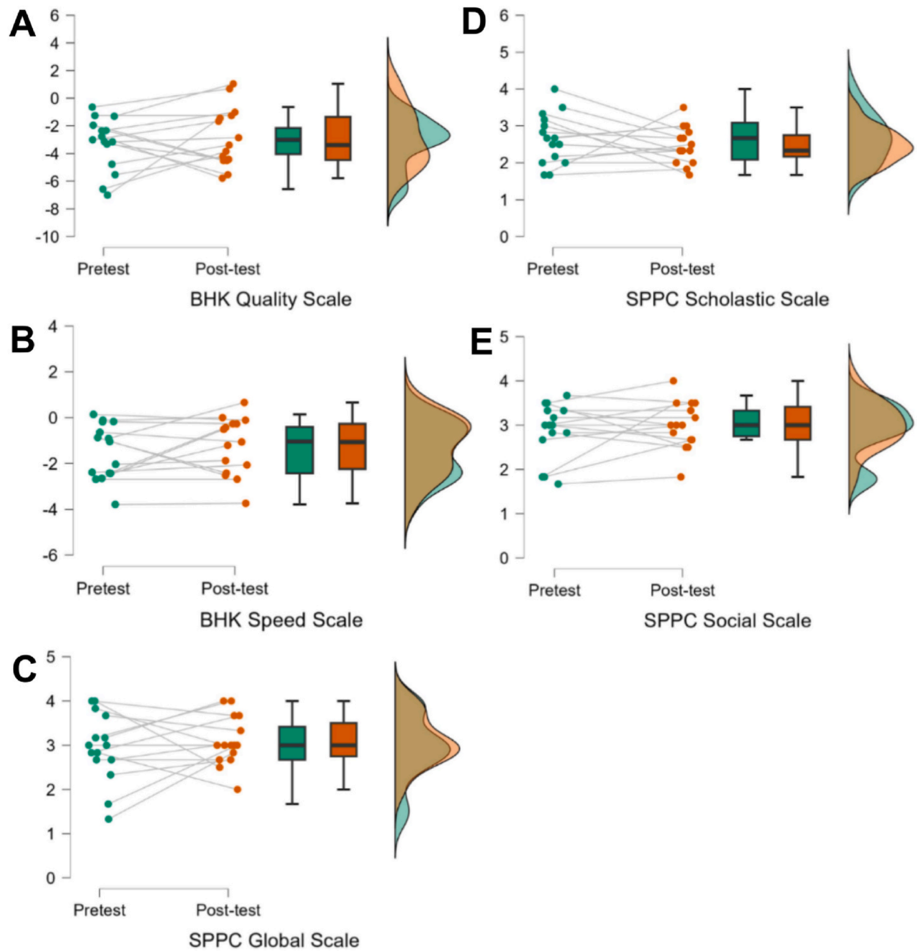

Fig. 4. Pretest v.s. post-test result. Results from the pretest and post-test comparisons, including: (A) the distribution of BHK quality scale scores, (B) the distribution of BHK speed scale scores, (C) the distribution of SPPC global scale scores, (D) the distribution of SPPC scholastic scale scores, and (E) the distribution of SPPC social scale scores. For all scales, higher scores indicate better outcomes. However, no significant differences were found between pretest and post-test results across all scales.

Table 2
Statistical analysis results comparing BHK and SPPC variables in pre-test and post-test conditions.

Pre/Post variable	W	z	p	Rank-Biserial Correlation	SE Rank-Biserial Correlation	95 % CI Lower	95 % CI Upper	r
BHK Quality	44	−0.534	0.616	−0.162	0.294	−0.64	0.406	−0.137
BHK Speed	41	−0.722	0.49	−0.219	0.294	−0.674	0.356	−0.186
SPPC Scholastic	18.5	−1.608	0.116	−0.526	0.316	−0.841	0.056	−0.415
SPPC Social	30.5	1.75	0.092	0.564	0.377	−0.079	0.927	0.4518
SPPC Global	33	0	1	0	0.328	−0.584	0.584	0

could influence the engagement of children and bias this so-called “drop-out” rate. Notably, children can be motivated to participate because they want to help researchers, because they want to please the caregiver they are working with or their parents, or because they appreciate the individualized attention they gain by participating. However, although the observed effects are encouraging, we cannot fully rule out alternative explanations, such as novelty effects or caregiver enthusiasm, which may also have contributed to children’s motivation and engagement.

4.2. *Effectiveness on writing skills and self-esteem*

Despite the absence of statistically significant pre-test to post-test evolution of BHK scores, it would be premature to dismiss the utility of our system regarding writing improvement, for at least three reasons. First, child participants do not only have dysgraphia but also a range of diverse comorbidities (see Supplementary Table S1 for more detail), which means that statistical generalization may be difficult for these children as they represent a very heterogeneous group. Second, due to constraints imposed by the context (such as children’s participation in other activities) or related to the children’s disorders, we planned our intervention over a relatively limited duration (1 session/week during 9 weeks), whereas improvement of writing skills after writing intervention usually requires 2 sessions per week and no less than 20 sessions (Hoy et al., 2011). Despite our lower frequency and total duration, it is noteworthy that 73.3 % of participants demonstrated an improvement of at least one standard deviation or more in at least one of the BHK scores. Finally, prolonged writing may lead to changes in the kinetics and kinematics of handwriting, which could potentially be a factor influencing the post-test BHK results. According to Kushki et al. (Kushki, Schwellnus, et al., 2011), horizontal stroke speed, grip force, and pressure force increase as children are writing for a prolonged time, while

Fig. 5. Examples of writing samples produced after the intervention with the R2C3 system. Even though the pretest and post-test comparisons yielded no significant differences, the progress made by some children can be qualitatively observed.

vertical stroke speed decreases. Since we requested child participants to complete the BHK test after the final session, they had already undergone 20 min of handwriting practice at this point. Thus, conducting the BHK test at this specific moment might have hindered their performance.

We also found no statistical differences between the pre-test and the post-test for the different SPPC scores, contrary to the expected results. As previously mentioned, our child participants are presenting complex NDD with many comorbidities, in addition to dysgraphia. NDD has been repeatedly associated with lower self-esteem in children (Akyurek & Murattoğlu, 2021), but self-esteem in children with NDD has also been found to be associated with other factors, such as parenting style, socio-economic factors (Arim et al., 2015), or the type of treatment they received for their disorder (Harpin et al., 2016). It might be that our intervention, targeting self-esteem impairment related to writing, was not impactful enough to modify the self-esteem of our participants, influenced by a multitude of other factors. Additionally, a nine-session intervention might have been too short to significantly influence children's self-esteem. Finally, some participants rated their self-esteem extremely high (at the maximum possible score) even pre-intervention. This is likely due to limited self-awareness that translates into inflated self-esteem scores. This tendency is labeled as positive illusory bias (Schuck et al., 2018).

4.3. Interaction traces during the sessions

Based on the results of the PCA, we identified two dimensions that encapsulate variations in terms of interaction amongst the triad during the session: verbal exchange within the triad and the non-verbal behavior of the child, predominantly characterized by the duration of gaze towards the tablet. We interpret the verbal exchange within the triad as a social component, reflecting the dynamics of interpersonal communication (Solomon et al., 2021). On the other hand, the non-verbal behavior of the child is considered a task component, indicating the child's engagement with the task - playing the serious game. These two components, social and task engagement, provide a comprehensive view of the interaction dynamics during this scenario.

We did expect children's writing evolution to be associated with task engagement, as the time children spent looking at the tablet is a proxy for engaging with the serious game. However, with the children with complex NDD and dysgraphia who participated in our learning-by-teaching scenario, we also found that social engagement is associated with writing improvement, as verbal exchanges during the training sessions predict writing evolution. Taken together, these results seem to

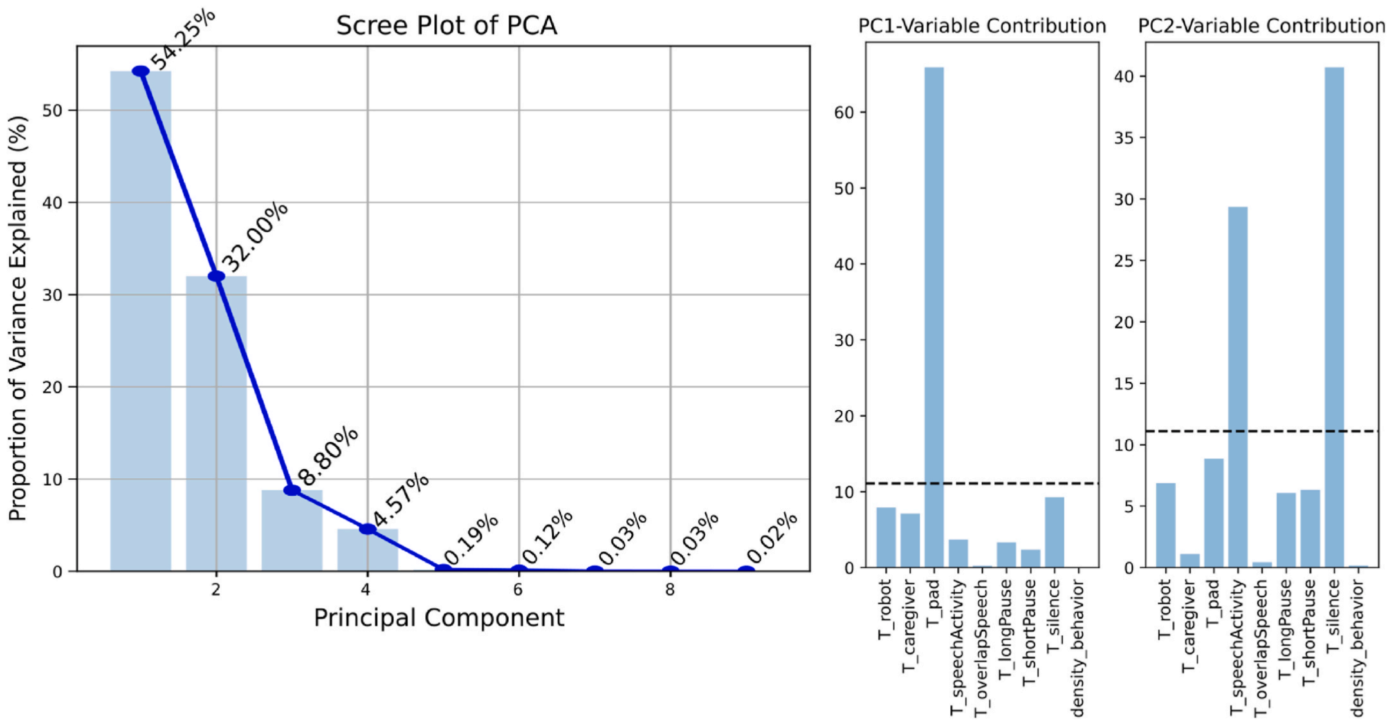

Fig. 6. PCA outcome related to interactive traces. The top two components represent 86.25 % of the variance, with the first component primarily influenced by the percentage of time child participants gaze at the tablet, and the second component largely determined by the percentages of speech activity time and silent time.

Table 3
Statistical analysis results about the correlation between the variation of BHK scores and interaction traces.

Dependent variable	std_estimate	95 % CI_low	95 % CI_up	p-value
T_robot	−0.04	−0.55	0.47	0.883
T_caregiver	−0.03	−0.53	0.47	0.928
T_pad (PC1)	−0.46	−0.95	0.04	0.07 (.)
T_speechActivity (PC2)	−0.5	−0.91	−0.09	0.018 (*)
T_overlapSpeech	−0.02	−0.41	0.37	0.932
T_longPause	−0.21	−0.72	0.31	0.427
T_shortPause	−0.37	−0.86	0.12	0.143
T_silence (PC2)	0.44	0	0.89	0.051 (.)
density_behavior	0.06	−0.24	0.37	0.685

underline the importance and meaningfulness of socially assistive robotics, and the social interaction it allows during the training sessions, to support handwriting reeducation in children with complex NDD.

4.4. Repetitive robot behaviors decrease quality of triadic interaction

For an intervention to be effective, it must maintain the engagement of its participants. High-frequency but repetitive robot behaviors do not appear to guide participants to engage in interactions effectively, confirming the importance of personalization of social robots (Leyzberg et al., 2014) and the importance of involving end-users in the designing processes (Neerincx et al., 2023). When the control of robot behaviors is in the hands of caregivers, with our developed R2C3 system comprising 120 different robot behaviors, caregivers make real-time selections of appropriate robot behaviors based on the child's state as well as the game state (not only game final score but also ongoing game state). In contrast, the robot in the Stereotyped Robot Condition provides only the most basic feedback based on the game's final score, completely disregarding the child's state. In the sessions involving the stereotyped robot condition, we noticed on several occasions that child participants expressed frustration or distress as the robot persistently requested them to play the same game in which they consistently struggled to achieve a favorable score. This might suggest that the interaction approach of the stereotyped robot is not as effective as that of the tailored robot. Therefore, the next steps will be to further optimize the interaction design of the stereotyped robot to make it more closely resemble that of the tailored robot.

4.5. Challenges and limitations

Regarding our results, a first limitation concerning the generalisability of the proposed findings lies in our sample size. To validate the efficiency of our system to ease writing reeducation with these complex cases, we need longer duration and more frequent sessions to reach statistical power to conclude. In addition, the single-site nature of the study further limits the generalisability of our findings. Recent works include both cross-sectional (Panceri et al., 2021) and longitudinal studies (Freitas et al., 2024; Mutawa et al., 2023; Scassellati et al., 2018), involving children with NDD, in which a robot operating in coordination with a tablet interface is used to enhance children's motivation (Freitas et al., 2024), engagement or basic socio-cognitive skills such as eye contact (Mutawa et al., 2023) and joint attention (Panceri et al., 2021). However, while this is not the first work focusing on a robot for handwriting training (Hood et al., 2015), none of these previous studies involved children with complex NDD and dysgraphia. Thus, to our knowledge, the present study is the largest conducted so far in this specific context.

Nevertheless, some methodological limitations should be noted. The recruitment strategy may have introduced a degree of selection bias, as the applied inclusion and exclusion criteria may not fully capture the diversity of children with NDD. Formal blinding was not feasible due to the overt presence of the robot and the involvement of caregivers; as a consequence, both caregivers and children may have formed expectations about the intervention that could have affected engagement and behaviour. We attempted to mitigate these expectation effects by randomising robot conditions and by keeping instructions to caregivers as neutral as possible. However, novelty effects or caregiver enthusiasm could still have influenced children's motivation and engagement.

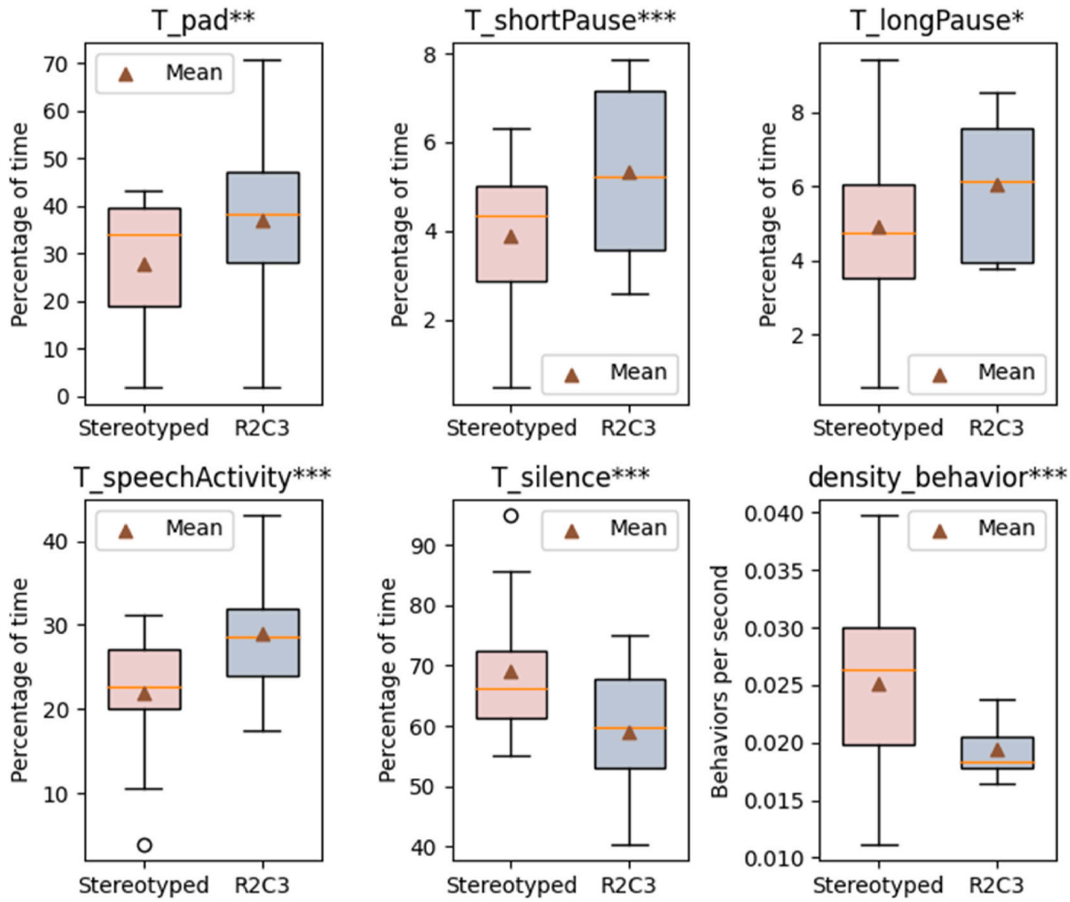

Fig. 7. Comparison of interaction traces between two robot conditions (*: $p < 0.05$; **: $p < 0.005$; ***: $p < 0.001$). Each feature is represented by a box plot, with the R2C3 Condition denoted by grey boxes and the Stereotyped Robot Condition by pink boxes. Despite the robot behaving more frequently in the Stereotyped Robot Condition (density of behavior is significantly higher in Stereotyped Robot Condition), it results in less interaction among participants (less time in speech activity with less pause) and less engagement in the serious game from child participants (less time of gaze to the tablet). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Statistical analysis results comparing the Stereotyped Robot Condition and the R2C3 condition.

Variable	std_estimate	95 %CI_low	95 %CI_up	p value
T_robot	−0.07	−0.52	0.37	0.727
T_caregiver	0.16	−0.3	0.62	0.509
T_pad	−0.49	−0.8	−0.17	0.002 (**)
T_speechActivity	−0.8	−1.17	−0.42	<0.001 (***)
T_overlapSpeech	0.2	−0.36	0.77	0.495
T_longPause	−0.48	−0.88	−0.08	0.018 (*)
T_shortPause	−0.67	−1.02	−0.31	<0.001 (***)
T_silence	0.78	0.42	1.15	<0.001 (***)
density_behavior	0.99	0.42	1.55	<0.001 (***)

Future experiments could further disentangle the specific contribution of the robot by comparing more proactive, autonomous systems, WoZ controlled robots and non-robotic tools. A final remark concerns the duration of the study (1 session/week during 9 weeks per child) that did not allow a long-term retention assessment of the observed handwriting improvements. As a consequence, conclusions about maintenance over time should be drawn with caution.

Technical limitations should also be highlighted. To understand the dynamics of triadic interactions involving a child, a caregiver, and a social robot, a comprehensive visual and audio analysis is imperative. Due to our experimental setup aiming to closely simulate real-world conditions, we did not impose restrictions on children's writing postures and attire. Consequently, during gaze tracking, we often encountered challenges to retrieving data, created for example by children entirely facing down, wearing glasses or a hat, and consequently limited eye detection. This was particularly evident as the experiments took place during the COVID-19 pandemic, while five children wore facial masks; their data were excluded from the multimodal interaction feature analysis, although their SPPC and BHK scores were retained. While mask wearing did not influence the standardised test scores, this clearly hindered the reliability of video-based gaze measures. At the same time, we described speech dynamics globally without detailed speech turns analysis. Speaker diarization and speech recognition were our initial choices for audio analysis. However, implementing this analysis in real-world settings presents significant challenges. Specifically, the employment of the QT robot's French text-to-speech system, featuring the voice of a young boy, poses difficulties in distinguishing this artificial voice from the actual child participants. Furthermore, through post hoc analysis of recorded audio, it was observed that caregivers tend to modify their vocal habits, making their voices softer and more soothing when addressing children during the training sessions. This alteration in their voice characteristics also results in variations within the same audio segment, making it difficult to discern that the voice belongs to the same caregiver, even though it does. Consequently, the shifts in caregivers' vocal characteristics diminish the accuracy of speaker diarization.

Despite these limitations, our findings suggest that socially assistive robots can play a meaningful role in supporting handwriting reeducation for children with complex NDD, highlighting their potential for integration into educational and clinical practice. In such contexts, robots could complement existing methods by sustaining children's motivation and engagement in tasks that are often perceived as difficult or discouraging. As the system was co-developed with caregivers and children, its progressive adoption by clinicians in autonomous use was facilitated, reinforcing not only the system's ecological validity and practical relevance, but also the value of the participatory design approach adopted. Nevertheless, its broader adoption still raises important challenges. Future work should further investigate the scalability of the system through larger, multi-site studies and collaborations with schools and clinical institutions, as well as its embedding into everyday routines. This will call for new practices, pedagogies, and didactic strategies that explicitly take into account the presence of the robot, addressing both its strengths in fostering motivation and engagement and its potential weaknesses, such as acting as a possible distractor, while also considering its capacity to adapt to the evolving needs of children and caregivers.

CRediT authorship contribution statement

Jianling Zou: Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Soizic Gauthier:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Dominique Archambault:** Validation, Supervision, Resources, Project administration, Funding acquisition. **Mohamed Chetouani:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization. **David Cohen:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Salvatore M. Anzalone:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Data and materials availability

The dataset generated and/or analysed during the current study are available upon request to the authors.

Funding

Agence Nationale de la Recherche (ANR), ANR-FNFS project iRe-Check - ANR-19-CE19-0029 - FNS 200021E 189475/1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors thank all families that participated in the study. Authors also thank Hugues Pellerin for supervising statistical analysis. Finally, authors thank the therapists that participated in the study, as well as the teaching team of the Heuyer School Centre.

Appendix A. The iRecheck Study Group

Barbara Bruno (Karlsruhe Institut für Technologie, Germany), Daniel C. Tozadore, Chenyang Wang, Thibault Asselborn, Pierre Dillenbourg (École Polytechnique Fédérale de Lausanne, Switzerland), Soraya El Farhane, Ingrid Zammouri, Atika Lemerle (Assistance publique - Hôpitaux de Paris, France), Thomas Gargot (Université de Tours, France).

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbr.2025.100799.

Data availability

Data will be made available on request.

References

Akyurek, G., & Murattoğlu, H. (2021). Study of self concept and social support perception: Neurodevelopmental disorders. *Journal of the Neurological Sciences, 429*. https://doi.org/10.1016/j.jns.2021.119222

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Publishing Incorporated.

Arim, R. G., Kohen, D. E., Garner, R. E., Lach, L. M., Brehaut, J. C., MacKenzie, M. J., & Rosenbaum, P. L. (2015). Psychosocial functioning in children with neurodevelopmental disorders and externalizing behavior problems. *Disability & Rehabilitation, 37*(4), 345–354. https://doi.org/10.3109/09638288.2014.919361

Asselborn, T., Chapatte, M., & Dillenbourg, P. (2020). Extending the spectrum of dysgraphia: A data driven strategy to estimate handwriting quality. *Scientific Reports, 10*(1), 3140. https://doi.org/10.1038/s41598-020-60011-8

Asselborn, T., Gargot, T., Kidziński, Ł., Johal, W., Cohen, D., Jolly, C., & Dillenbourg, P. (2018). Automated human-level diagnosis of dysgraphia using a consumer tablet. *npj Digital Medicine, 1*(1), 1–9. https://doi.org/10.1038/s41746-018-0049-x

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L.-P. (2018). OpenFace 2.0: Facial behavior analysis toolkit. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (pp. 59–66). https://doi.org/10.1109/FG.2018.00019

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3*(21), Article eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Biotteau, M., Danna, J., Baudou, É., Puyjarinet, F., Velay, J.-L., Albaret, J.-M., & Chaix, Y. (2019). Developmental coordination disorder and dysgraphia: Signs and symptoms, diagnosis, and rehabilitation. *Neuropsychiatric Disease and Treatment, 15*, 1873–1885. https://doi.org/10.2147/NDT.S120514

Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social robotics. In B. Siciliano, & O. Khatib (Eds.), *Springer handbook of robotics* (pp. 1935–1972). Springer International Publishing. https://doi.org/10.1007/978-3-319-32552-1_72.

Bredin, H., Yin, R., Coria, J. M., Gelly, G., Korshunov, P., Lavechin, M., Fustes, D., Titeux, H., Bouaziz, W., & Gill, M.-P. (2020). Pyannote.audio: Neural building blocks for speaker diarization. *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing* (pp. 7124–7128). https://doi.org/10.48550/arXiv.1911.01255

Chung, E. Y., Kuen-fung Sin, K., & Chow, D. H. (2024). Effectiveness of robotic intervention on improving social development and participation of children with autism spectrum disorder – A randomised controlled trial. *Journal of Autism and Developmental Disorders*, 1–8. https://doi.org/10.1007/s10803-024-06236-2

Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and dysgraphia: Definition, diagnosis, and management. *Translational Pediatrics, 9*(Suppl 1), S46–S54. https://doi.org/10.21037/tp.2019.11.01

Costa, S., Lehmann, H., Dautenhahn, K., Robins, B., & Soares, F. (2015). Using a humanoid robot to elicit body awareness and appropriate physical interaction in children with autism. *International Journal of Social Robotics, 7*(2), 265–278. https://doi.org/10.1007/s12369-014-0250-2

Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine and Child Neurology, 49*(4), 312–317. https://doi.org/10.1111/j.1469-8749.2007.00312.x

Freitas, É. V. da S., Panceri, J. A. C., Schreider, S. da L., Caldeira, E. M. de O., & Filho, T. F. B. (2024). Cognitive serious games dynamically modulated as a therapeutic tool for applied behavior analysis therapy in children with autism spectrum disorder. *International Journal of Emerging Technologies in Learning (iJET), 19*(5), Article 05. https://doi.org/10.3991/ijet.v19i05.48677

Gargot, T., Asselborn, T., Pellerin, H., Zammouri, I., Anzalone, S. M., Casteran, L., Johal, W., Dillenbourg, P., Cohen, D., & Jolly, C. (2020). Acquisition of handwriting in children with and without dysgraphia: A computational approach. *PLoS One, 15*(9), Article e0237575. https://doi.org/10.1371/journal.pone.0237575

Gargot, T., Asselborn, T., Zammouri, I., Brunelle, J., Johal, W., Dillenbourg, P., Archambault, D., Chetouani, M., Cohen, D., & Anzalone, S. M. (2021). "It is not the robot who learns, it is me"; treating severe dysgraphia using Child-Robot interaction. *Frontiers in Psychiatry, 12*. https://doi.org/10.3389/fpsyt.2021.596055

Grossard, C., Palestra, G., Xavier, J., Chetouani, M., Grynszpan, O., & Cohen, D. (2018). ICT and autism care: State of the art. *Current Opinion in Psychiatry, 31*(6), 474–483. https://doi.org/10.1097/YCO.0000000000000455

Hamstra-Bletz, L., & Blöte, A. W. (1993). A longitudinal study on dysgraphic handwriting in primary school. *Journal of Learning Disabilities, 26*(10), 689–699. https://doi.org/10.1177/002221949302601007

Hamstra-Bletz, L., DeBie, J., & Ben Brinker, B. (1987). *Concise evaluation scale for children's handwriting, 1*. Lisse: Swets.

Harpin, V., Mazzone, L., Raynaud, J. P., Kahle, J., & Hodgkins, P. (2016). Long-term outcomes of ADHD: A systematic review of self-esteem and social function. *Journal of Attention Disorders, 20*(4), 295–305. https://doi.org/10.1177/1087054713486516

Harter, S. (2012). *Self-perception profile for children: Manual and questionnaires*. Univeristy of Denver, Department of Psychology.

Hood, D., Lemaignan, S., & Dillenbourg, P. (2015). When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 83–90. https://doi.org/10.1145/2696454.2696479

Hoy, M. M. P., Egan, M. Y., & Feder, K. P. (2011). A systematic review of interventions to improve handwriting. *Canadian Journal of Occupational Therapy, 78*(1), 13–25. https://doi.org/10.2182/cjot.2011.78.1.3

Kouroupa, A., Laws, K. R., Irvine, K., Mengoni, S. E., Baird, A., & 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), Article e0269800. https://doi.org/10.1371/journal.pone.0269800

Kushki, A., Chau, T., & Anagnostou, E. (2011). Handwriting difficulties in children with autism spectrum disorders: A scoping review. *Journal of Autism and Developmental Disorders, 41*(12), 1706–1716. https://doi.org/10.1007/s10803-011-1206-0

Kushki, A., Schwellnus, H., Ilyas, F., & Chau, T. (2011). Changes in kinetics and kinematics of handwriting during a prolonged writing task in children with and without dysgraphia. *Research in Developmental Disabilities, 32*(3), 1058–1064. https://doi.org/10.1016/j.ridd.2011.01.026

Leyzberg, D., Spaulding, S., & Scassellati, B. (2014). Personalizing robot tutors to individuals' learning differences. *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, 423–430. https://doi.org/10.1145/2559636.2559671

Mehta, P. P., & Nandgaonkar, H. P. (2019). Visual-perceptual training for handwriting legibility and speed in children with handwriting difficulties: A single-arm interventional study. *Indian Journal of Occupational Therapy, 51*(1), 14. https://doi.org/10.4103/ijoth.IJOTH_11_18

Mutawa, A. M., Al Mudhahkah, H. M., Al-Huwais, A., Al-Khaldi, N., Al-Otaibi, R., & Al-Ansari, A. (2023). Augmenting Mobile app with NAO robot for autism education. *Machines, 11*(8), Article 8. https://doi.org/10.3390/machines11080833

Neerincx, A., Veldhuis, D., Masthoff, J. M. F., & de Graaf, M. M. A. (2023). Co-designing a social robot for child health care. *International Journal of Child-Computer Interaction, 38*, Article 100615. https://doi.org/10.1016/j.ijcci.2023.100615

Palsbo, S. E., & Hood-Szivek, P. (2012). Effect of robotic-assisted three-dimensional repetitive motion to improve hand motor function and control in children with handwriting deficits: A nonrandomized phase 2 device trial. *American Journal of Occupational Therapy, 66*(6), 682–690. https://doi.org/10.5014/ajot.2012.004556

Panceri, J. A. C., Freitas, É., de Souza, J. C., da Luz Schreider, S., Caldeira, E., & Bastos, T. F. (2021). A new socially assistive robot with integrated serious games for therapies with children with autism spectrum disorder and Down syndrome: A pilot study. *Sensors, 21*(24), Article 24. https://doi.org/10.3390/s21248414

Riek, L. (2012). Wizard of Oz studies in HRI: A systematic review and New Reporting Guidelines. *Journal of Human-Robot Interaction*, 119–136. https://doi.org/10.5898/JHRI.1.1.Riek

Scassellati, B. (2007). How social robots will help Us to diagnose, treat, and understand autism. In S. Thrun, R. Brooks, & H. Durrant-Whyte (Eds.), *Robotics research: Result of the 12th international symposium ISRR, 28* pp. 552–563). Springer. https://doi.org/10.1007/978-3-540-48113-3_47.

Scassellati, B., Boccanfuso, L., Huang, C.-M., Mademtzi, M., Qin, M., Salomons, N., Ventola, P., & Shic, F. (2018). Improving social skills in children with ASD using a long-term, in-home social robot. *Science Robotics, 3*(21), Article eaat7544. https://doi.org/10.1126/scirobotics.aat7544

Schuck, S. E. B., Johnson, H. L., Abdullah, M. M., Stehli, A., Fine, A. H., & Lakes, K. D. (2018). The role of animal assisted intervention on improving self-esteem in children with attention Deficit/Hyperactivity disorder. *Frontiers in Pediatrics, 6*(300). https://doi.org/10.3389/fped.2018.00300

Solomon, D. H., Brinberg, M., Bodie, G. D., Jones, S., & Ram, N. (2021). A dynamic dyadic systems approach to interpersonal communication. *Journal of Communication, 71*(6), 1001–1026. https://doi.org/10.1093/joc/jqab035

Steinfeld, A., Jenkins, O. C., & Scassellati, B. (2009). The oz of wizard: Simulating the human for interaction research. *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (pp. 101–108). https://doi.org/10.1145/1514095.1514115

Xavier, J., & Cohen, D. (2020). Multidimensional impairments. In A. Gallagher, C. Bulteau, D. Cohen, & J. L. Michaud (Eds.), *Handbook of clinical neurology, 174* pp. 159–169). Elsevier. https://doi.org/10.1016/B978-0-444-64148-9.00012-0.

Yanjana, Singh, P., & Kumar, M. (2020). Behavioral intervention with fine motor training for Dysgraphia in school going children. *International Journal of Current Research and Review, 12*(18), 180–187. https://doi.org/10.31782/IJCRR.2020.121827

Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Killian, J. M., Voigt, R. G., Weaver, A. L., & Katusic, S. K. (2011). Written-Language disorder among children with and without ADHD in a population-based birth cohort. *Pediatrics, 128*(3), e605–e612. https://doi.org/10.1542/peds.2010-2581

Zou, J., Gauthier, S., Anzalone, S. M., Cohen, D., & Archambault, D. (2022). A wizard of Oz interface with Qtrobot for facilitating the handwriting learning in children with Dysgraphia and its usability evaluation. In K. Miesenberger, G. Kouroupetroglou, K. Mavrou, R. Manduchi, M. Covarrubias Rodriguez, & P. Penáz (Eds.), *Computers helping people with special needs, 13342* pp. 219–225). Springer International Publishing. https://doi.org/10.1007/978-3-031-08645-8_26.

Zou, J., Gauthier, S., Pellerin, H., Gargot, T., Archambault, D., Chetouani, M., Cohen, D., & Anzalone, S. M. (2024). R2C3, A rehabilitation robotic companion for children and caregivers: The collaborative design of a social robot for children with neurodevelopmental disorders. *International Journal of Social Robotics, 16*(3), 599–617. https://doi.org/10.1007/s12369-024-01104-6