Task-Related Electrodermal Activity Dynamics in Robot-mediated Interaction: A Six-Case Study of Children with Neurodevelopmental Disorders

Jaeryoung Lee, *Member, IEEE,* and Dimitar Stefanov, *Member, IEEE*

***Abstract*—Socially assistive robots (SARs) have been explored as platforms to support structured interaction with children with neurodevelopmental disorders (NDDs), however, most systems are designed for individuals with autism spectrum disorder (ASD) and may not be generalizable to other conditions. Emotional engagement and physiological arousal often differ across diagnostic categories and symptom severities, highlighting the challenge of capturing heterogeneous arousal patterns observed in small, diverse participant samples. This study presents a six case exploratory analysis of electrodermal activity (EDA) responses during four task based emotional interaction sessions in robot-assisted training involving children diagnosed with ASD, pervasive developmental disorder (PDD), developmental coordination disorder (DCD), and intellectual disability (ID). A multimodal platform integrating a NAO robot, an E4 wristband, facial expression tracking, and therapist control was used to synchronize behavioral and physiological data streams using unified timestamps; physiological signals were recorded for offline analysis. EDA signals were interpreted as task related arousal dynamics rather than discrete emotional states. Descriptive time series inspection and vector autoregressive impulse response visualization were applied to characterize individual variability. The findings revealed heterogeneous arousal trajectories shaped by task type and individual characteristics, reflecting pronounced individual differences in task related arousal dynamics. Overall, this study presents a timestamp synchronized multimodal data logging platform and an exploratory six-case analysis of task related EDA arousal dynamics during robot-mediated interaction, emphasizing pronounced individual variability.**

***Index Terms*—Human Robot Interaction, Physiological Arousal, Socially Assistive Robots, Children with NDDs, Robot-mediated Interaction**

I. Introduction

SARs have been widely explored as research platforms to support engagement, social interaction, and tasks related to emotion among children with NDDs in controlled settings [1], [2], [3]. Among various NDDs, autism spectrum disorder (ASD) has been a primary focus of therapeutic interventions using SARs. ASD is a highly prevalent developmental condition, that affects approximately one in 32 children in the United States [4]. Although ASD has been the primary focus of most SAR research, systems designed specifically for ASD may not be generalizable effectively to other developmental conditions such as pervasive developmental disorder (PDD), developmental coordination disorder (DCD), or intellectual disability (ID). Although these conditions share overlapping features, they also exhibit substantial heterogeneity in their cognitive, behavioral, and emotional reactivity profiles [5], [6]. As a result, ASD-focused robot-mediated interaction systems may not be generalizable effectively across diverse diagnostic populations. However, broader generalization remains an open challenge, given that educational environments for NDDs typically encompass children with multiple and overlapping diagnoses. Supporting this perspective, Márquez-Caraveo et al. [7] evaluated 74 children and adolescents aged 6–15 years who were diagnosed with NDDs, including ID, ASD, and attention-deficit hyperactivity disorder (ADHD). Using standardized cognitive and functional assessment tools, the authors comprehensively measured the participants' cognitive, behavioral, and social functioning. Their analysis revealed heterogeneous cognitive profiles, particularly in working memory and perceptual reasoning indices, emphasizing the need for individualized diagnostic and intervention approaches that reflect each child's cognitive characteristics. Furthermore, Liu et al. [8] conducted a comprehensive review of differentiated educational approaches for students with special educational needs. Their findings highlighted that children with ID, ASD, and learning disabilities exhibit distinct emotional and sensory profiles, thus requiring specialized pedagogical strategies and adopted learning environments to support their educational development effectively. In this context, there is growing interest in developing SAR research systems that can characterize behavioral and physiological variability across heterogeneous developmental profiles during interaction.

For this purpose, recent advances in multimodal sensing have enabled approaches for collecting behavioral and physiological signals during interaction. Such technology has enabled robots to capture behavioral and physiological cues during therapy, including facial expressions, gaze, and physiological signals. Among these, EDA offers a non-invasive physiological index of arousal regulated by the sympathetic nervous system [9]. Although widely used as an indicator of emotional engagement, EDA reflects general physiological activation rather than discrete emotional categories. Therefore, interpreting EDA within task-related or interactional contexts may provide additional context for understanding task-related arousal dynamics during interaction [10], [11]. In line with these developments, a recent systematic review on technology-based interventions for NDDs [12] also highlighted this imbalance. Similar to what has been observed in educational

This study was supported by JSPS KAKENHI Grant Number JP20K19327.
This study was approved by the Ethics Committee of Chubu University.

J. Lee is with the Department of AI Robotics, Chubu University, Kasugai, Aichi, 487-8501 Japan e-mail: jaeryounglee@fsc.chubu.ac.jp

D. Stefanov is with the Department of Design Engineering & Maths, Middlesex University, The Burroughs, Hendon, London NW4 4BT, UK

and clinical contexts, studies have predominantly focused on ASD, followed by ADHD, with markedly fewer investigations targeting other developmental conditions. Based on these technological advancements, the present study applies EDA-based assessment to the context of robot-mediated interaction, aiming to describe task-related variations in physiological arousal among children with NDDs.

From this perspective, the study explores how task-related EDA dynamics manifest during robot-mediated interaction for children with different developmental conditions. To address the large inter-individual variability and small sample size typical of such studies, we adopt an exploratory multiple single-case design involving six children diagnosed with ASD, PDD, DCD, or ID. A multimodal platform integrating a NAO robot, an E4 wristband, facial expression recognition for automated robot control, and optional manual control by the therapist was used to synchronize behavioral and physiological data streams using unified timestamps; physiological signals were recorded for offline analysis. Furthermore, the synchronization architecture of the system ensures coherence across heterogeneous modalities through unified temporal alignment, enhancing usability in interaction settings and enabling precise cross-modal analysis of emotional responses during interaction.

This study makes the following contributions:

- **EDA characterization across heterogeneous NDD profiles**: We analyze task-related EDA patterns observed during robot-mediated interaction in six children representing heterogeneous neurodevelopmental disorders, extending prior SAR research beyond ASD-focused populations.
- **Exploration of cross-phase arousal dynamics using VAR/IRF**: Given the sequential presentation of multiple robot-generated expressive phases within a single emotion expression training session, we employ vector autoregression (VAR) and impulse response functions (IRF) as descriptive tools to examine temporal dependencies in physiological arousal across successive task phases.
- **Exploratory characterization for individual variability**: We report descriptive, single-case differences in individual arousal dynamics during robot-mediated interaction, demonstrating that physiological responses vary markedly across children even under identical robot-defined task conditions.

Within this context, the central aim of this work is to introduce a timestamp synchronized multimodal data logging platform and to explore individual arousal dynamics at the single case level during robot-mediated interaction, with a focus on individual heterogeneity. These exploratory observations motivate future research on SARs that consider individual arousal patterns and cognitive profiles in robot-mediated interaction settings. In line with this scope, the present study does not infer or classify participants' discrete emotional states. All analyses treat EDA as an arousal indicator, and the labels "Anger," "Sadness," "Joy," and "Surprise" refer to robot-generated expressive phases rather than children's subjective emotional experiences.

II. Related Work

A. Socially Assistive Robots for Neurodevelopmental Disorders

In the early stages of SAR development, much attention was given to the physical appearance and basic functionalities of robots to attract the attention of children, particularly those with ASD. Many new robot designs were created with the goal of stimulating focus and engagement through appealing shapes and predictable actions. Representative examples include KASPAR [13], Probo [14], Keepon [15], NAO [16], and Zeno [17]. These robots are appreciated for their predictable movements, simplified facial expressions, and ability to support turn-taking, imitation, and emotion-recognition tasks. Later, as robotic platform efficiency improved and deep learning research gained momentum, the focus shifted from the robot's appearance to the effectiveness of therapeutic content, specifically the robot's ability to perceive human behaviors and deliver appropriate therapy scenarios. This evolution enhanced the overall maturity of the SAR field for children with NDDs, still with ASD children as the primary focus. More recently, studies have broadened the scope of SARs to include other NDDs such as ADHD, DCD, and intellectual disabilities, highlighting an increasing emphasis on inclusive therapeutic environments [18]. Furthermore, some studies have co-developed robotic interfaces and training scenarios, conducting multi-session interventions that demonstrated meaningful improvements in the learning performance of children with NDDs [19]. While the therapeutic benefits of SARs are well documented, most existing systems still rely mainly on behavioral observation or facial expression analysis to assess emotional engagement [20]. However, these modalities often fail to fully capture internal arousal states, especially in children who have non-autistic NDDs or comorbid conditions accompanying ASD.

B. Physiological Sensing in Human Robot Interaction

To improve the objectivity of emotional assessment, recent research has incorporated physiological measures such as heart rate variability, skin temperature, and EDA [9], [11]. EDA, in particular, provides a sensitive index of sympathetic activation and has been used to infer emotional arousal during human–robot interaction. For example, Vacaru et al. evaluated wearable devices for monitoring physiological arousal during robot-delivered mindfulness exercises [21], showing that EDA reflects stress-related physiological changes. Although Lutin et al. [22] investigated stress responses in collaborative robot contexts rather than social robots, their work addressed user stress, which is one of the key factors in social robot interaction. They demonstrated that tonic EDA can serve as a meaningful physiological indicator of perceived stress during human robot collaboration. In the context of SARs, Welch et al. employed physiological signals and machine learning algorithms to assess attentiveness during robot-assisted social skills training for children with ASD [23]. While their study highlighted the importance of developing individualized predictive systems, the focus was confined to ASD populations. These studies collectively demonstrate the growing feasibility

of integrating physiological sensing into SAR frameworks. Nevertheless, the interpretation of EDA remains methodologically challenging: it reflects general physiological activation rather than discrete emotional categories. Thus, EDA should be analyzed in task-related contexts or as a proxy for engagement and arousal dynamics, rather than as a direct indicator of specific emotions [10].

C. Research Gap and Present Approach

Existing SAR studies have mainly focused on ASD-specific applications and have often treated EDA as a categorical emotion signal. Few studies have explored how arousal patterns differ across diagnostic groups or how task-related fluctuations in EDA can inform robot behaviors. Furthermore, many prior systems relied on partial synchronization of devices or Wizard-of-Oz (WoZ) control, limiting tight integration between behavioral and physiological data [24].

To address these limitations, the present study introduces a timestamp-synchronized multimodal system for robot-mediated interaction that logs robot behavior, therapist input, facial analysis outputs, and wearable physiological signals using unified UTC-based timestamps. Multimodal frameworks have been widely studied in human robot interaction to capture internal states by combining physiological and behavioral signals. For example, Green and Iqbal collected EDA, skin temperature, blood volume pulse, and facial expressions to examine trust in a robot partner and reported statistically significant differences in physiological and facial indicators across self-reported trust levels [25]. Similarly, Perugia et al. developed a statistical engagement model using behavioral and physiological sensing, including EDA, to assess engagement in individuals with dementia during play activities [26]. In addition, Can et al. proposed multimodal deep learning architectures leveraging wearable signals such as EDA, skin temperature, and blood volume pulse for stress and emotion recognition in daily-life settings [27]. However, the primary objective of these studies has been to infer or classify internal states such as trust, engagement, or affect at the population level. At the same time, approaches that use EDA for emotion-related inference vary substantially in methodological standardization and often face challenges in reliably capturing physiological indices of emotional arousal in children with NDDs.

In this context, the present study extends prior SAR research by examining task-related EDA responses in children with heterogeneous NDD profiles rather than focusing exclusively on ASD populations. Because the emotion expression training task consists of multiple robot-generated expressive phases presented sequentially within a single session, a time-series perspective is adopted to explore how physiological arousal elicited during one phase relates temporally to subsequent phases. To this end, VAR and IRF analyses are employed as descriptive tools to investigate cross-phase arousal dynamics at the single-case level. Therefore, the contribution of this study lies not in the general concept of multimodal synchronization itself, but in its concrete implementation within a small-scale educational setting involving therapist participation and its application to exploratory, single-case temporal analysis of arousal dynamics rather than internal state inference or classification. Physiological signals were analyzed offline, and no real-time arousal feedback was provided during interaction. Furthermore, discrete emotion classification was intentionally avoided; instead, the study emphasizes descriptive, individualized analysis of arousal patterns, and all findings are interpreted at an exploratory level.

III. Methodology

A. Participants and Protocol

Participants were recruited from an after-school facility for children with NDDs. Parents were informed of the study's objectives and procedures in writing, and the facility's therapists selected participants based on the children's condition on the day of the experiment. To ensure consistency of the experimental environment and minimize stress caused by unfamiliar settings, the same therapists and caregivers from the facility participated throughout the sessions. A total of six children with neurodevelopmental disorders (three boys and three girls; mean age = 7 years) participated in the study. Detailed participant characteristics are presented in Table I.

TABLE I: Participant Characteristics

Participant	Sex	Age (years)	Diagnosis	Notes
Boy 1	M	6	ASD	Comorbid ADHD
Boy 2	M	7	ASD	—
Boy 3	M	8	ASD	—
Girl 1	F	7	PDD	—
Girl 2	F	7	DCD	—
Girl 3	F	6	ID	—

The interaction framework was based on the Theory of Mind approach [28], which is widely applied in occupational therapy to promote emotional and social understanding. The task was designed to help children observe and interpret the robot's expressive behaviors (posture, vocal tone) rather than actual emotional states. After these cues were interpreted, the children were asked to identify the corresponding expressive condition and imitate it. This approach aims to facilitate key aspects of Theory of Mind development, including empathy, self–other distinction, and understanding others' emotions.

Fig. 1 shows the experimental setup and device arrangement. Three webcams were used to record the entire session, capture the participant's frontal view, and recognize facial expressions. The room contained two tables: one for the NAO robot, keypad, and emotional expression cards, and another for the cards only. Participants sat between the two tables and adjusted their position depending on the activity. The NAO robot was controlled in both autonomous and manual (keypad operated) modes. Only the child and the therapist were present in the experimental room.

Prior to the start of the experimental session, the E4 wristband was fitted on the participant's non-dominant wrist approximately 10 minutes in advance. This lead-in period was used to allow sensor stabilization and to obtain a stable baseline before task onset. To ensure consistent skin contact and to reduce the risk of signal loss due to loosening during

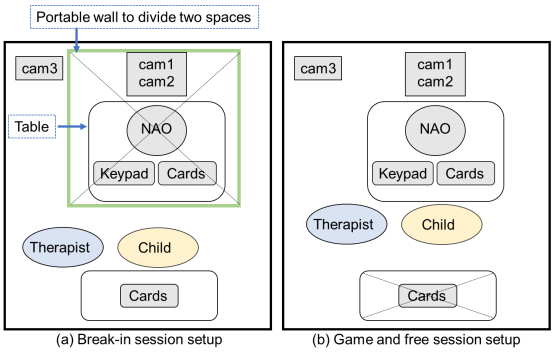


Fig. 1: Experimental setup: (a) break-in phase with an unseen robot; (b) robot-expression guessing game and free play phases with visible robot.

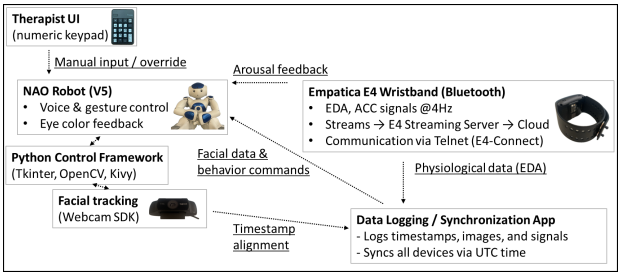


Fig. 2: Overall architecture of the multimodal robot-mediated interaction system integrating the NAO robot, facial tracking, therapist control, and physiological sensing via the E4 wristband. All devices were synchronized via a Python-based UTC timestamp framework; data analysis was performed offline.

movement, the wristband was secured with a fabric sports wristband placed over the device.

The experiment lasted for 30 minutes and consisted of three sessions: break-in, robot-expression guessing game, and free play, each lasting approximately 10 minutes. During the break-in phase, the robot was hidden behind a movable wall (Fig. 1(a)), while the therapist and child played with emotional expression cards (used purely as task labels rather than indicators of participants' subjective emotions). This period was essential for familiarization and for relaxing the child before the main task. During the robot-expression guessing phase, the wall was removed, and the main session began (Fig. 1(b)). In this phase, the NAO robot presented predefined expressive conditions using posture and vocal tone. The child was asked to identify and imitate the robot's expressive condition. The roles were then alternated so that the robot attempted to recognize and infer the child's displayed expression. In the final free play phase, the child interacted spontaneously with the robot, and the therapist provided assistance when necessary to control the robot's actions. Four predefined robot-generated expressive conditions (Anger, Sadness, Joy, and Surprise) were presented as distinct task phases, hereafter referred to as expressive phases. These labels describe the robot's programmed behaviors, not the participant's subjective or physiological emotional state. Importantly, the present study does not infer or classify children's discrete emotional states from behavior or physiology. All analyses treat EDA strictly as an indicator of autonomic arousal, and the labels "Anger," "Sadness," "Joy," and "Surprise" refer only to robot-generated expressive phases rather than participants' subjective emotions.

The study protocol was approved by the Chubu University Ethics Committee, and all procedures were conducted in full accordance with institutional ethical guidelines.

B. System Overview and Device Synchronization

The robot-mediated interaction system consisted of a NAO robot and several applications designed for its operation and data management. The robot used in this study was the NAO V5, a small humanoid robot capable of both verbal and non-verbal interaction. Owing to these characteristics, NAO has been widely adopted in educational and support contexts for children [29]. The system also included several auxiliary components: a control application developed in Python using the NAOqi framework, a software development kit (SDK) for facial recognition, a numeric keypad that allowed the therapist to manually control the robot, a data logging application, and an application for acquiring physiological data from the E4 wristband.

The system integrated these multimodal components to enable synchronized behavioral and physiological recording. The NAO control application primarily managed the robot's behavior during the robot-expression guessing game task through automated control based on facial recognition outputs. In this setup, the NAO robot operated by default in an autonomous control mode, in which facial recognition results were used to detect participants' expressions and trigger robot-generated expressive behaviors, thereby advancing the interaction scenario. When necessary, the therapist could override the autonomous behavior and switch the robot to a manual control mode using a numeric keypad. Facial recognition performance was not evaluated and was not treated as a perception accuracy metric; its role was limited to managing interaction flow rather than assessing recognition performance. The numeric keypad enabled therapist intervention when required but was not used as the primary control mechanism.

The data logging application was developed in Python using the Tkinter, OpenCV, and Kivy libraries. This application collected time information based on Coordinated Universal Time (UTC) and stored both images and corresponding timestamps in a designated folder. The experimental interface utilized a library capable of managing multiple functions concurrently. The E4 wristband application, written in Python, required a Bluetooth connection to a PC prior to initialization. Once connected to the E4 Streaming Server, the wristband began acquiring physiological signals, which were subsequently stored on E4-Connect, a cloud-based service provided by Empatica. Communication between the E4 wristband app and E4-Connect was established via Telnet. Physiological signal data were continuously collected, categorized, and recorded in text files, with each entry accompanied by timestamps in both UTC and Japan Standard Time (JST).

Figure 2 illustrates the overall architecture of the multimodal robot-mediated interaction system. The framework integrates

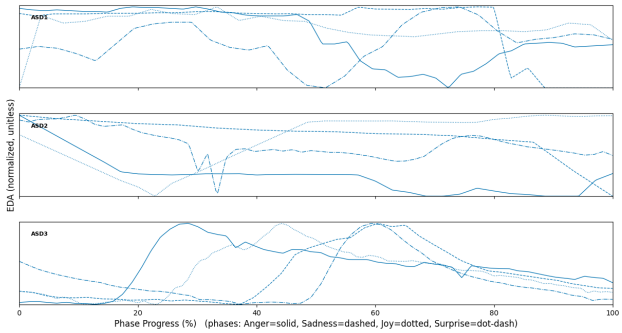
the NAO robot, facial tracking, therapist control input, and physiological sensing into a synchronized data-logging environment. Although the system architecture allows real-time arousal feedback from the E4 wristband to the NAO robot, this feature was not activated in the present study; instead, the EDA data were recorded solely for offline analysis to explore physiological arousal patterns as a preliminary investigation. To provide visual feedback to the therapist, the robot's eye color reflected the current output of the facial recognition system. Following Plutchik's theory of emotion [30], specific colors were assigned to core emotional states: yellow for Surprise, blue for Sadness, red for Anger, and green for Joy.

C. Data analysis

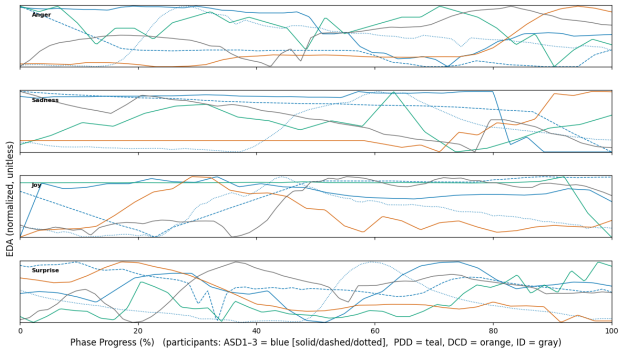
During the robot-expression guessing game session, we categorized the EDA data according to the robot-generated expressive phases (Anger, Sadness, Joy, Surprise). These labels represent task-defined expressive conditions rather than indicators of participants' subjective emotions. Initially, we inspected phase segmented EDA trajectories for the three ASD cases as an illustrative subset, without performing diagnosis level comparisons. To mitigate potential discrepancies arising from variation in the number of samples across phases, we calculated the minimum segment length for each participant and truncated all segments accordingly. Normalization procedures were applied to facilitate comparison across participants. The EDA data for each session were normalized using either min–max or offset normalization, methods widely used in physiological comparison and affective computing [31]. This transformation converted all values to nonnegative numbers and enabled a more straightforward comparison of relative changes.

Following normalization, additional metrics were computed to quantitatively characterize variability in participants' arousal responses within each robot-defined expressive phase. In addition to mean EDA and peak-to-peak change (ΔEDA), two complementary variability measures were included: the coefficient of variation (CV) and the root mean square of successive differences (RMSSD) [32]. The CV was defined as the ratio of the standard deviation to the mean EDA and was used to describe relative variability while accounting for differences in absolute arousal magnitude across participants. This measure enabled an exploratory comparison of response stability among participants with substantially different baseline arousal levels. RMSSD was calculated based on successive differences between adjacent EDA samples and was used as an index of short-term fluctuation and temporal instability in arousal responses. In this context, it provides a descriptive summary of individual response dynamics under identical task conditions. All variability metrics were computed separately for each robot-generated expressive phase. No statistical comparisons across diagnostic groups or task conditions were performed.

To investigate cross-phase dynamics of arousal, we employed a VAR model and visualized the IRF. The VAR model, widely used in time-series econometrics and affective dynamics research [33], [34], [35], was implemented using the statsmodels library in Python [36]. This allowed us to explore how perturbations in the arousal signal during one expressive phase influenced subsequent arousal patterns in other phases over time, without inferring discrete emotional states.



(a) Three ASD participants: within-case trajectories across expressive task phases.



(b) All six participants: normalized trajectories by expressive task phase.

Fig. 3: Within-participant EDA patterns across robot-defined expressive task phases. (a) ASD participants. (b) All participants; color indicates diagnostic group and line style distinguishes ASD individuals.

IV. Results

For clarity, the term "expressive phase" refers to a task defined experimental phase characterized by a specific robot-generated expressive behavior, namely Anger, Sadness, Joy, or Surprise. Here, EDA is interpreted as a measure of physiological arousal elicited by these task conditions rather than discrete emotional states.

A. Within-Participant Arousal Patterns Across Task Phases

Figure 3 summarizes within-participant EDA trajectories across the four robot-generated expressive phase, while the summary statistics in Table II clarify the absolute magnitude and distributional spread of EDA responses across participants, allowing a more precise comparison of within-phase variability than is available from the normalized trajectories alone. For visualization purposes, the raw EDA signals were synchronized with task timestamps, smoothed using a 1-second moving-average filter to reduce noise, and segmented by expressive phase. Because the duration of each expressive phase varied across participants, each segment was rescaled

to a common 0–100% progress axis. Because wearable EDA recorded during active child robot interaction is susceptible to motion artifacts, the segmented EDA signals were visually inspected for abrupt, nonphysiological spikes consistent with movement disturbances; no automated artifact rejection was applied.

Panel (a) in Figure 3 shows individual trajectories for the three ASD participants, plotted using a consistent blue color, with line style encoding expressive phase (Anger = solid, Sadness = dashed, Joy = dotted, Surprise = dot–dash). Patterns differed markedly across children, including gradual increases, sharp drops, and sustained plateaus. Even for the same phase, the duration and amplitude of arousal varied by child, indicating heterogeneous task-related arousal dynamics at the single-case level. In particular, the Joy (dotted) expressive condition showed strong responses, although not uniformly the most pronounced across all participants. Specifically, for ASD1, EDA remained at relatively high levels during the Joy (dotted) condition, while in the Anger (solid) and Surprise (dot–dash) conditions, it fluctuated considerably throughout the session. During the Sadness (dashed) condition, EDA initially stayed high but then dropped sharply. Overall, the Joy (dotted) condition showed the longest duration of elevated arousal, indicating that physiological activation during the Joy expressive phase was sustained for a longer period. ASD2 maintained high EDA levels in both the Joy (dotted) and Sadness (dashed) conditions. In the Anger (solid) condition, EDA exhibited a gradual downward trend over the latter portion of the session, whereas Surprise (dot–dash) showed a modest overall decrease with minor fluctuations. For ASD3, all phases followed a similar overall pattern, showing mid-session increases and subsequent declines, although the initial EDA levels differed across conditions.

Panel (b) presents normalized EDA responses from six participants across four distinct expressive phases. A color-blind-friendly qualitative palette is used to represent diagnostic groups (ASD = blue, PDD = teal, DCD = orange, ID = gray), while line styles differentiate the three ASD participants (ASD1 = solid, ASD2 = dashed, ASD3 = dotted), and each subplot corresponds to one expressive phase. Table II summarizes the mean EDA, peak-to-peak change (ΔEDA), confidence intervals, and variability indices for each robot-defined expressive phase across all participants. To characterize response stability within each phase, 95% confidence intervals (CIs) were estimated using a bootstrap resampling approach, providing an index of distributional spread around the mean. In addition to CI-based dispersion, relative and short-term variability were quantified using the CV and RMSSD. These metrics capture complementary aspects of arousal dynamics that are not evident from normalized trajectories alone.

During the Anger expressive phases, ASD2 (blue, dashed) exhibited large overall amplitude changes (Δ EDA = 4.302), rather than a single sharp peak-and-drop pattern. The ID participant (gray, solid) showed a gradual increase followed by moderate variability in the latter half (Δ EDA = 2.057). By contrast, ASD3 (blue, dotted) exhibited a gradual rise and fall in EDA across each phase, maintaining relatively low amplitude throughout. This pattern suggests a subdued but consistent physiological responsiveness rather than a flat or unreactive profile. PDD showed consistently low mean EDA values across all conditions (Mean EDA = 0.011, 95% CI = [0.009, 0.012]), suggesting a generally reduced sensitivity or dampened physiological reactivity to these robot-generated expressive stimuli. In the Sadness phases, ASD1 (blue, solid) demonstrated a relatively low-variation pattern with a modest decline (95% CI = [0.428, 0.583]), whereas ASD2 showed greater variability, reflected in a wider confidence interval (95% CI = [6.088, 6.192]). ASD3 once again displayed a gradual increase followed by a slow decline, replicating the response pattern observed during Anger. Although the normalized plot does not show PDD at a clearly lower visual amplitude during Sadness, the summary statistics indicate that the PDD participant had low absolute EDA in this phase (Mean = 0.139, Δ = 0.042). The DCD participant (orange) showed moderately high absolute EDA levels during the Sadness phase (Mean = 3.132), with relatively small within-session variation (Δ EDA = 0.353). In the normalized plot, this appears as a mostly stable trajectory with a slight upward trend toward the latter part of the phase. In the Joy phases, ASD2 showed a rapid increase in EDA during the mid-to-late period, maintaining elevated levels thereafter, suggesting an immersive sympathetic response to this positive expressive condition (Mean EDA = 4.504, 95% CI = [3.972, 5.035], Δ EDA = 5.263). ASD3 exhibited a response trajectory similar to those in other expressive conditions, with a slow rise, a peak, and a gradual decline. ASD3 exhibited relatively high mean EDA values across all expressive conditions (Mean EDA = 4.7–6.3); however, the Δ EDA values remained within a narrow range (0.4–0.9), indicating limited variation in response magnitude across expressive phases. The ID participant displayed a sustained EDA pattern across the session, with moderate amplitude changes (Δ EDA = 1.571). In the Surprise phases, ASD2 (blue, dashed) showed pronounced amplitude changes (Δ EDA = 1.675), indicating substantial within-phase variability despite the absence of rapid oscillatory patterns. ASD3 (blue, dotted) displayed a gradual rise and fall with relatively small phasic variation (Δ EDA = 0.959), consistent with the slow, low-variation trajectories observed in the other expressive phases, even though the absolute mean EDA remained high. The ID participant (gray) exhibited moderate variability (Δ EDA = 0.914), characterized by a mid-session increase followed by continued fluctuations rather than a clear stabilization.

Using these CI estimates, participant level differences in within-phase variability became evident. ASD1 and ASD3 showed consistently narrow CIs across conditions (e.g., ASD1-Anger: 95% CI = [0.517, 0.541]; ASD3–Surprise: [6.257, 6.390]), indicating tightly clustered EDA distributions. In contrast, ASD2 exhibited comparatively wider CIs in several phases (e.g., Sadness: [6.088, 6.192]; Joy: [3.972, 5.035]), corresponding to greater dispersion around the mean. Among the non-ASD participants, PDD showed the lowest absolute EDA with narrow CIs (e.g., Sadness: [0.135, 0.144]), whereas DCD and ID displayed higher means accompanied by broader intervals (e.g., DCD–Anger: [2.780, 2.905]; ID–Joy: [4.263, 4.527]). These CI metrics quantify response stability beyond

what is visible in the normalized trajectories.

To further quantify these observed differences, Table II reports phase-specific variability indices (CV and RMSSD) for each participant. Even within the three ASD cases observed in this study, marked differences in within-phase variability were evident. ASD2 consistently showed elevated variability across multiple expressive phases, with high CV values during Anger (CV = 0.945) and Joy (CV = 0.361), as well as large RMSSD values indicating pronounced short-term fluctuations (e.g., Joy RMSSD = 0.369). In contrast, ASD3 exhibited comparatively low variability across phases despite high mean EDA levels, with CV values remaining below 0.05 in all conditions (e.g., Anger CV = 0.024; Joy CV = 0.029), reflecting stable but elevated arousal responses. ASD1 demonstrated intermediate and phase dependent patterns, with moderate variability during Sadness (CV = 0.469; RMSSD = 0.104) and lower variability in other phases. These quantitative differences reinforce the observation that task-related arousal dynamics vary substantially at the individual level, even under identical robot-generated expressive task conditions. In addition, the variability indices provide complementary information beyond mean amplitude, ΔEDA, or confidence intervals alone, highlighting individual differences in response stability and temporal fluctuation rather than diagnosis level effects.

TABLE II: Mean EDA, ΔEDA, 95% CI, and variability indices (CV, RMSSD) across robot-defined expressive phases for each participant. Note: Phase labels denote robot-defined expressive task phases and do not indicate the child's subjective emotional state.

Phase	Subj	Mean EDA	ΔEDA	CI Low	CI High	CV	RMSSD
Anger	ASD1	0.529	0.137	0.517	0.541	0.083	0.010
	ASD2	1.118	4.302	0.773	1.463	0.945	0.289
	ASD3	4.776	0.424	4.748	4.805	0.024	0.027
	PDD	0.011	0.014	0.009	0.012	0.323	0.003
	DCD	2.843	0.622	2.780	2.905	0.067	0.040
	ID	4.630	2.057	4.493	4.767	0.117	0.162
Sadness	ASD1	0.506	0.661	0.428	0.583	0.469	0.104
	ASD2	6.140	0.425	6.088	6.192	0.017	0.053
	ASD3	5.829	0.581	5.774	5.884	0.032	0.040
	PDD	0.139	0.042	0.135	0.144	0.075	0.011
	DCD	3.132	0.353	3.098	3.166	0.031	0.038
	ID	2.430	0.567	2.380	2.480	0.066	0.061
Joy	ASD1	0.480	0.122	0.471	0.488	0.047	0.022
	ASD2	4.504	5.263	3.972	5.035	0.361	0.369
	ASD3	4.987	0.507	4.958	5.015	0.029	0.018
	PDD	1.349	1.204	1.255	1.442	0.180	0.175
	DCD	2.511	0.396	2.471	2.551	0.046	0.049
	ID	4.395	1.571	4.263	4.527	0.126	0.095
Surprise	ASD1	0.684	0.115	0.674	0.694	0.042	0.015
	ASD2	5.871	1.675	5.795	5.947	0.053	0.183
	ASD3	6.324	0.959	6.257	6.390	0.042	0.062
	PDD	0.198	0.091	0.191	0.205	0.122	0.014
	DCD	2.259	0.580	2.203	2.315	0.075	0.047
	ID	3.634	0.914	3.578	3.689	0.062	0.071

B. Dynamic Arousal Response Analysis Between Expressive Phases Using Impulse Response Functions

To descriptively explore how participants' physiological responses evolved across robot-generated expressive phases, we employed VAR and IRF to explore cross-phase arousal pattern similarity and transition dynamics [34]. Although the expressive-phase segments were temporally separated, the experimental design ensured clearly defined intervals for each condition and sufficient time-series length within each phase. Accordingly, each phase-specific EDA segment was resampled to a fixed length using linear interpolation to enable a descriptive comparison of cross-phase arousal patterns [37], rather than to model continuous physiological dynamics. The VAR model was fitted individually to the resampled data from each participant using the Python statsmodels library. The IRF was analyzed over 0–5 time steps to capture the immediate and short-term responses of changes in arousal during one expressive phase on subsequent phases. This range was chosen to characterize phase-linked perturbation patterns while minimizing the influence of noise or random fluctuations that may occur over longer time frames. Importantly, these models describe transitions in physiological arousal, not transitions between discrete emotional states. Because the expressive phases were presented in a fixed, non-randomized sequence, the VAR and IRF analyses were not used to infer causal relationships between phases. Instead, they were applied as descriptive tools to visualize short-term temporal dependencies and carryover patterns in physiological arousal across successive task phases at the individual level.

TABLE III: Strongest expressive-phase arousal influence per participant (IRF, 0–5 steps).

Participant	Source Phase	Target Phase	Avg IRF (0–5)
ASD1	Sadness	Sadness	1.02
ASD2	Anger	Sadness	-1.68
ASD3	Anger	Anger	1.04
DCD	Anger	Anger	1.15
ID	Joy	Joy	1.05
PDD	Joy	Anger	6.36

According to the IRF analysis, the strongest phase-to-phase arousal influence observed for each participant is summarized in Table III. In particular, large IRF magnitudes observed in individual cases (e.g., the PDD participant) should be interpreted cautiously, as they arise from single-case estimates and may be sensitive to normalization and resampling choices. For ASD1, the most prominent transition occurred from Sadness to Sadness (Avg IRF = 1.02), indicating persistence of arousal within the same expressive phase. ASD2 exhibited the strongest negative cross-phase influence from Anger to Sadness (Avg IRF = –1.68), suggesting a potential transition from a higher to a lower arousal level between two negative expressive conditions. Both ASD3 and DCD demonstrated the highest IRF values within Anger (Avg IRF = 1.04 and 1.15, respectively), reflecting a pattern of self-reinforcing arousal. In contrast, the participant with ID showed the strongest influence within Joy (Avg IRF = 1.05), implying retention of heightened arousal during this expressive phase. The PDD participant presented an unusually strong transition from Joy to Anger (Avg IRF = 6.36), potentially indicative of abrupt changes in physiological activation across phases. These findings highlight the heterogeneity of arousal transition dynamics among children with NDDs.

Note that Fig. 4 visualizes the mean IRF values averaged across all participants, whereas Table III highlights the strongest phase-to-phase influence identified for each participant individually. As a result, differences between these representations are expected, as they capture complementary

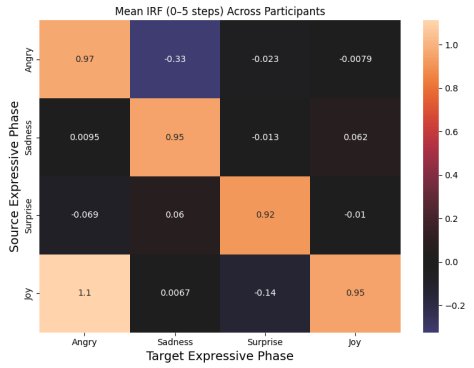


Fig. 4: Cross-phase arousal influence matrix derived from impulse response functions. Mean IRF values (steps 0–5) averaged across participants indicate influence between robot-expressed phases in physiological arousal (EDA). Warmer colors indicate stronger positive influence.

aspects of arousal-transition dynamics. Fig. 4 shows the average impulse response (IRF) values from steps 0 to 5 across all participants. Each cell represents the influence of one robot-generated expressive phase (source) on another (target) in terms of physiological arousal (EDA). A self-reinforcement pattern was observed across all expressive phases, indicating that physiological arousal elicited during each phase tended to persist rather than rapidly diminish or shift. Strong cross-phase influences, such as from Joy to Anger (1.1), were observed, whereas transitions like Surprise to Joy (–0.01) were comparatively weak. The IRF value of -0.14 for the transition from the Joy phase to the Surprise phase indicates that when the robot's expressive condition shifted from Joy to Surprise, participants' EDA responses tended to decrease or the arousal associated with Surprise was suppressed. This suggests that even when Joy occurred first, the subsequent presentation of Surprise did not evoke a strong physiological response, implying an attenuation of arousal intensity. Similarly, the negative value of –0.33 for the transition from Anger to Sadness indicates that when the robot's expressive condition shifted from Anger to Sadness, participants' physiological arousal levels decreased over time.

V. Discussion

A. Individual Differences in Task-Related Arousal Dynamics

This study developed and tested a synchronized, multi-device system designed to explore physiological arousal responses in children with NDDs during structured robot-mediated interaction sessions in an after school support setting. The system integrated several components, including a NAO humanoid robot, facial recognition software for controlling the robot, a therapist-controlled keypad, an E4 wristband for physiological acquisition, and a data-logging interface. By synchronizing these components using unified timestamps, the system enabled temporally aligned logging of robot actions (including therapist control inputs) and sensor streams. Interpretation and analysis were performed offline after data collection. Although the platform supports multi-stream integration, the present study focused on offline analysis of EDA aligned to task timestamps and did not perform real-time adaptation during the sessions.

This design enhances temporal precision and allows reproducible analysis of the temporal associations between behavioral events and autonomic responses. Moreover, the synchronized framework presented here offers an extensible foundation for multimodal data fusion in robot-mediated interaction research. In line with prior multimodal research emphasizing the importance of temporally aligned behavioral and physiological data for understanding heterogeneity in neurodevelopmental conditions, the present system supports synchronized multimodal data logging to facilitate interpretation of arousal dynamics relative to robot behavior and experimenter input [38].

In the pilot application, the system recorded time-aligned EDA during interactions with robot-generated expressive scenarios. As shown in Section IV-A, each child exhibited distinct temporal trajectories across four expressive phases, reflecting the heterogeneous physiological response patterns observed among individuals with NDDs. In this study, EDA does not represent discrete emotion categories directly but rather serves as an indicator of sympathetic arousal levels. The labels "Anger," "Sadness," "Joy," and "Surprise" refer not to participants' subjective emotions but to the contextual expressions produced by the robot during the tasks. Therefore, observed fluctuations in EDA may be interpreted as variations in overall arousal related to attention, task engagement, or stimulus novelty, rather than emotion-specific signals. Accordingly, the present study does not infer or classify children's discrete emotional states from EDA; all analyses treat EDA strictly as a physiological arousal signal within task-defined expressive phases.

For instance, one participant (ASD3) displayed gradually increasing and then decreasing EDA curves across multiple expressive phases. Rather than suggesting a lack of emotional response, this pattern may reflect a sustained autonomic activation with limited differentiation between expressive contexts. In contrast, another participant (ASD2) exhibited rapid and frequent fluctuations, particularly during phases labeled as Surprise, indicating high moment-to-moment reactivity. This aligns with previous findings that some children with ASD display heightened sympathetic responses during emotionally demanding interactions [39]. Meanwhile, the participant with ID showed relatively stable and distinct EDA patterns across conditions, suggesting consistent yet less reactive arousal modulation. Although individuals with ID have often been excluded from psychophysiological research within the broader field of NDDs [40], recent studies report that autonomic signals, including EDA, show condition-dependent differences in stress or task contexts and can be utilized for behavioral prediction [41], [42]. When considered alongside the heterogeneous EDA patterns observed in the present study, these findings highlight the utility of single-case exploratory approaches in developmental research. Each child's physiological trajectory likely reflects a unique combination of autonomic regulation, sensory sensitivity, and task engagement, supporting the value of individualized interpretations over group-level comparisons.

Moreover, the confidence-interval estimates provide further insight into the stability of these heterogeneous patterns. Participants such as ASD1 and ASD3 showed consistently narrow CIs across expressive phases, suggesting relatively stable or low-variation arousal profiles despite differences in absolute magnitude. In contrast, ASD2 exhibited wider intervals in several conditions, indicating greater moment-to-moment variability and less stable autonomic modulation. Among the non-ASD participants, PDD demonstrated uniformly low absolute EDA with narrow CIs, whereas DCD and ID showed moderate-to-high mean levels accompanied by CI widths in the narrow-to-moderate range. These CI-based distinctions highlight that heterogeneity in autonomic responding arises not only from differences in overall arousal level but also from differences in the stability and variability of responses across expressive contexts. The observed differences in CV and RMSSD among the ASD participants further support an individual level interpretation of arousal dynamics rather than inference at the diagnostic category level.

B. Limitations and Implications

Inter-individual differences may also arise from contextual factors such as attention span, or engagement level. In addition, given the high activity level of children, motion-related noise may have partially contributed to observed EDA variations. Because the present analysis focused on exploratory single-case characterization, artifact aware preprocessing was not applied. Therefore, the potential influence of movement related artifacts cannot be fully excluded. Future studies should integrate behavioral coding and posture sensors to disentangle physiological arousal from movement-related effects. Because emotional expressive scenarios were presented sequentially rather than counterbalanced, statistical inference of causal relationships between phases or labels is not appropriate. Accordingly, IRF analysis was used solely as a technical tool to visualize short-term dependencies and residual effects between expressive phases, rather than to infer causality. Given the very small sample size ($n = 6$), statistical generalization is not possible, and all comparative analyses and summary statistics are presented in an exploratory and descriptive manner. Although applied descriptively, IRF analysis provided a structured way to examine whether arousal carries over across phases or dissipates rapidly. Such temporal carryover patterns are relevant for understanding individual differences in autonomic regulation, recovery rates, or sensitivity to sequential stimuli factors often observed in children with NDDs but rarely quantified.

From a methodological perspective, identifying individual arousal trajectories may inform the design of future experimental protocols, such as the pacing of task transitions or the ordering of expressive conditions. For example, participants exhibiting prolonged arousal may benefit from longer transition intervals between task phases, whereas participants showing rapid fluctuations may require additional stabilization periods. While the present study is limited to offline analysis and exploratory modeling, physiological signals may, in the longer term and within separate research frameworks, be investigated as potential inputs for robot behavior modulation. To better differentiate between recognition of expressive cues and emotional experience, subsequent studies may combine physiological data with additional behavioral or contextual measures obtained during interaction.

Future work will also employ repeated sessions, counterbalanced expressive sequences, and larger sample sizes to enhance reliability. The use of preregistration and permutation based multiple comparison correction is planned to strengthen the statistical foundation of single case experimental designs. In this light, the present work should be interpreted as a feasibility and characterization study of synchronized logging and single-case arousal dynamics, not as evidence of intervention efficacy. Nevertheless, several fundamental limitations must be acknowledged. Participants represented heterogeneous diagnostic categories, further constraining interpretation beyond the individual level. Moreover, the analyses were exploratory and descriptive in nature; thus, the findings should be interpreted as indicative trends rather than confirmatory results. In the absence of a control group, observed differences are interpreted as reflecting individual variability under identical task constraints rather than diagnostic categories. Despite these limitations, small-sample pilot designs are common in robot-assisted therapy research involving children with NDDs, where recruitment and sustained participation present inherent challenges [18]. Within this context, the present study serves as a proof of concept demonstration of a timestamp synchronized multimodal data logging framework for exploratory analysis of physiological arousal during robot-mediated interaction.

VI. Conclusion

This pilot study presented a synchronized multimodal system for robot-mediated interaction study and explored individual EDA patterns across robot-defined expressive phases in children with NDDs. The findings highlight both the technical feasibility of multimodal synchronization and the pronounced individual differences that emerge in physiological arousal dynamics. Interpreting EDA as a task- and context-dependent arousal signal, rather than as a direct indicator of discrete emotional categories, contributes to a more realistic and ecologically valid understanding of physiological responses during robot-assisted interaction in children with NDDs. Moreover, the technical foundation supporting synchronization and recording of multimodal data using timestamps provides an essential basis for future research exploring how physiological and behavioral signals may be incorporated into robot-mediated interaction frameworks in relation to individual child states. Although the small sample size limits the generalizability of the findings, this study offers an important empirical basis for refining arousal-informed interaction modeling and examining the role of multimodal sensing in supporting exploratory research on robot-mediated interaction with children with NDDs in larger-scale research. Continued efforts to integrate physiological signal detection, adaptive robotic behaviors, and personalized feedback mechanisms will support the development of more responsive and individualized interaction systems.

REFERENCES

[1] A. Kouroupa, K. R. Laws, K. Irvine, S. E. Mengoni, A. Baird, and S. Sharma, "The use of social robots with children and young people on the autism spectrum: A systematic review and meta-analysis," *Plos one*, vol. 17, no. 6, p. e0269800, 2022.

[2] B. Scassellati, H. Admoni, and M. Matarić, "Robots for use in autism research," *Annu. Rev. Biomed. Eng.*, vol. 14, pp. 275–294, 2012.

[3] W. Wang, J. Xiao, and L. Diao, "The effects of robots on children with autism spectrum disorder: A meta-analysis," *J. Autism Dev. Disord.*, pp. 1–12, 2025.

[4] M. J. Maenner and et al, "Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2020," *MMWR Surveill. Summ.*, vol. 72, no. 2, pp. 1–14, 2023.

[5] O. Rudovic, J. Lee, L. Mascarell-Maricic, B. W. Schuller, and R. W. Picard, "Measuring engagement in robot-assisted autism therapy: A cross-cultural study," *Front. Robot. AI*, vol. 4, p. 36, 2017.

[6] A. Bonarini, F. Garzotto, M. Gelsomini, M. Romero, F. Clasadonte, and A. N. Ç. Yilmaz, "A huggable, mobile robot for developmental disorder interventions in a multi-modal interaction space," in *Proc. IEEE RO-MAN*, pp. 823–830, IEEE, 2016.

[7] M. E. Márquez-Caraveo, R. Rodriguez-Valentin, V. Perez-Barron, R. A. Vazquez-Salas, J. C. Sánchez-Ferrer, F. De Castro, B. Allen-Leigh, and E. Lazcano-Ponce, "Children and adolescents with neurodevelopmental disorders show cognitive heterogeneity and require a person-centered approach," *Sci. Rep.*, vol. 11, no. 1, p. 18463, 2021.

[8] X. Liu and M. Potmesil, "A review of research on the development of inclusive education in children with special educational needs over the past 10 years: a visual analysis based on citespace," *Front. Educ.*, vol. 9, p. 1475876, 2025.

[9] W. Boucsein, *Electrodermal activity*. Springer, 2012.

[10] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," in *Handbook of Psychophysiology* (J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, eds.), pp. 159–181, Cambridge, U.K.: Cambridge Univ. Press, 3rd ed., 2007.

[11] D. Erol Barkana, K. D. Bartl-Pokorny, H. Kose, A. Landowska, M. Milling, B. Robins, B. W. Schuller, P. Uluer, M. R. Wrobel, and T. Zorcec, "Challenges in observing the emotions of children with autism interacting with a social robot," *Int. J. Soc. Robot.*, pp. 1–16, 2024.

[12] M. O. Ribas, M. Micai, A. Caruso, F. Fulceri, M. Fazio, and M. L. Scattoni, "Technologies to support the diagnosis and/or treatment of neurodevelopmental disorders: A systematic review," *Neurosci. Biobehav. Rev.*, vol. 145, p. 105021, 2023.

[13] L. J. Wood, A. Zaraki, B. Robins, and K. Dautenhahn, "Developing kaspar: a humanoid robot for children with autism," *Int. J. Soc. Robot.*, vol. 13, pp. 491–508, 2021.

[14] R. E. Simut, J. Vanderfaeillie, A. Peca, G. Van de Perre, and B. Vanderborght, "Children with autism spectrum disorders make a fruit salad with probo, the social robot: an interaction study," *J. Autism Dev. Disord.*, vol. 46, pp. 113–126, 2016.

[15] H. Kozima, C. Nakagawa, and Y. Yasuda, "Children–robot interaction: a pilot study in autism therapy," *Prog. Brain Res.*, vol. 164, pp. 385–400, 2007.

[16] S. Shamsuddin, H. Yussof, L. I. Ismail, S. Mohamed, F. A. Hanapiah, and N. I. Zahari, "Initial response in hri-a case study on evaluation of child with autism spectrum disorders interacting with a humanoid robot nao," *Procedia Eng.*, vol. 41, pp. 1448–1455, 2012.

[17] F. Lecciso, A. Levante, R. A. Fabio, T. Caprì, M. Leo, P. Carcagnì, C. Distante, P. L. Mazzeo, P. Spagnolo, and S. Petrocchi, "Emotional expression in children with asd: A pre-study on a two-group pre-post-test design comparing robot-based and computer-based training," *Front. Psychol.*, vol. 12, p. 678052, 2021.

[18] M. Pivetti, S. Di Battista, F. Agatolio, B. Simaku, M. Moro, and E. Menegatti, "Educational robotics for children with neurodevelopmental disorders: A systematic review," *Heliyon*, vol. 6, no. 10, 2020.

[19] J. Zou, S. Gauthier, H. Pellerin, T. Gargot, D. Archambault, M. Chetouani, D. Cohen, and S. M. Anzalone, "R2c3, a rehabilitation robotic companion for children and caregivers: the collaborative design of a social robot for children with neurodevelopmental disorders," *Int. J. Soc. Robot.*, vol. 16, no. 3, pp. 599–617, 2024.

[20] O. Rudovic, J. Lee, M. Dai, B. Schuller, and R. W. Picard, "Personalized machine learning for robot perception of affect and engagement in autism therapy," *Sci. Robot.*, vol. 3, no. 19, p. eaao6760, 2018.

[21] S. V. Vacaru, L.-P. Lau, K. Frederiks, P. S. Sterkenburg, and E. Barakova, "Selecting optimal wearables for measuring physiological arousal in robot-delivered mindfulness-based exercises," *Adv. Robot.*, vol. 38, no. 19-20, pp. 1364–1377, 2024.

[22] E. Lutin, S. A. Elprama, J. Cornelis, P. Leconte, B. Van Doninck, M. Witters, W. De Raedt, and A. Jacobs, "Pilot study on the relationship between acceptance of collaborative robots and stress," *Int. J. Soc. Robot.*, vol. 16, no. 6, pp. 1475–1488, 2024.

[23] K. C. Welch, R. Pennington, S. Vanaparthy, H. M. Do, R. Narayanan, D. Popa, G. Barnes, and G. Kuravackel, "Using physiological signals and machine learning algorithms to measure attentiveness during robot-assisted social skills intervention: A case study of two children with autism spectrum disorder," *IEEE Instrum. Meas. Mag.*, vol. 26, no. 3, pp. 39–45, 2023.

[24] L. D. Riek, "Wizard of oz studies in hri: a systematic review and new reporting guidelines," *J. Hum.-Robot Interact.*, vol. 1, no. 1, pp. 119–136, 2012.

[25] H. N. Green and T. Iqbal, "Examining physiological response and facial expression as indicators of trust in a robot partner," *ACM Trans. Human-Robot Interact.*, vol. 15, no. 2, Article 32, pp. 1–24, 2025.

[26] G. Perugia, M. Díaz-Boladeras, A. Catala-Mallofré, E. I. Barakova, and M. Rauterberg, "Engage-dem: a model of engagement of people with dementia," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 926–943, 2020.

[27] Y. S. Can, M. Benouis, B. Mahesh, and E. André, "Application of multimodal self-supervised architectures for daily life affect recognition," *IEEE Trans. Affect. Comput.*, vol. 16, no. 3, pp. 2454–2465, 2025.

[28] S. Baron-Cohen, A. M. Leslie, and U. Frith, "Does the autistic child have a "theory of mind"?," *Cognition*, vol. 21, no. 1, pp. 37–46, 1985.

[29] A. W. Kushniruk, W. Kuang, and E. M. Borycki, "The nao robot in healthcare and education: A scoping review," *Stud. Health Technol. Inform.*, vol. 323, pp. 374–378, 2025.

[30] R. Plutchik, "A general psychoevolutionary theory of emotion," in *Theories of emotion*, pp. 3–33, Elsevier, 1980.

[31] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, 2005.

[32] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[33] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer, 2005.

[34] T. Krone, C. J. Albers, P. Kuppens, and M. E. Timmerman, "A multivariate statistical model for emotion dynamics.," *Emotion*, vol. 18, no. 5, p. 739, 2018.

[35] L. F. Bringmann, E. Ferrer, E. L. Hamaker, D. Borsboom, and F. Tuerlinckx, "Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model," *Multivariate Behav. Res.*, vol. 53, no. 3, pp. 293–314, 2018.

[36] S. Seabold and J. Perktold, "Statsmodels: econometric and statistical modeling with python.," *SciPy*, vol. 7, no. 1, pp. 92–96, 2010.

[37] M. Halicki and T. Niedzielski, "A new approach for hydrograph data interpolation and outlier removal for vector autoregressive modelling: a case study from the odra/oder river," *Stochastic Environ. Res. Risk Assess.*, vol. 38, no. 7, pp. 2781–2796, 2024.

[38] L. Pang, X. Zhao, L. Zhao, J. Li, F. Kuo, H. Wang, and C. Liu, "Multimodal data analysis for autism spectrum disorder in children: State of the art and trends," *EngMedicine*, vol. 3, no. 1, p. 100117, 2026.

[39] R. M. Fenning, S. A. Erath, J. K. Baker, D. S. Messinger, J. Moffitt, B. R. Baucom, and A. K. Kaeppler, "Sympathetic-parasympathetic interaction and externalizing problems in children with autism spectrum disorder," *Autism Res.*, vol. 12, no. 12, pp. 1805–1816, 2019.

[40] J. M. Moffitt, J. K. Baker, R. M. Fenning, S. A. Erath, D. S. Messinger, S. M. Zeedyk, S. A. Paez, and S. Seel, "Parental socialization of emotion and psychophysiological arousal patterns in children with autism spectrum disorder," *Res. Child Adolesc. Psychopathol.*, vol. 49, no. 3, pp. 401–412, 2021.

[41] R. Simons, R. Koordeman, P. de Looff, R. Otten, *et al.*, "Physiological measurements of stress preceding incidents of challenging behavior in people with severe to profound intellectual disabilities: Longitudinal study protocol of single-case studies," *JMIR Res. Protoc.*, vol. 10, no. 7, p. e24911, 2021.

[42] S. de Vries, F. van Oost, H. Smaling, N. de Knegt, P. Cluitmans, R. Smits, and E. Meinders, "Real-time stress detection based on artificial intelligence for people with an intellectual disability," *Assist. Technol.*, vol. 36, no. 3, pp. 232–240, 2024.