Research in Developmental Disabilities 135 (2023) 104452

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

Child-therapist interaction features impact Autism treatment response trajectories

Giulio Bertamini [a,b,c,d,*], Silvia Perzolli [a], Arianna Bentenuto [a], Eleonora Paolizzi [a], Cesare Furlanello [e,f], Paola Venuti [a]

[a] Laboratory of Observation, Diagnosis, and Education (ODFLab), Department of Psychology and Cognitive Science, University of Trento, 38068 Rovereto, Italy
[b] Department of Child and Adolescent Psychiatry, Pitie-Salpetriere University Hospital, Sorbonne University, 75013 Paris, France
[c] Data Science for Health (DSH), Bruno Kessler Foundation, 38123 Trento, Italy
[d] Institute for Ingelligent Systems and Robotics (ISIR), Sorbonne University, 75005 Paris, France
[e] Orobix Life Sciences, 24121 Bergamo, Italy
[f] HK3Lab, 38068 Rovereto, Italy

A R T I C L E I N F O

Keywords:
Autism therapy
Child-therapist interaction
Quantitative methods
Observational research
Predictive modeling
Treatment response trajectories
Longitudinal outcomes
Developmental profile

A B S T R A C T

Background: Identifying mechanisms of change in Autism treatment may help explain response variability and maximize efficacy. For this, the child-therapist interaction could have a key role as stressed by developmental models of intervention, but still remains under-investigated.
Aims: The longitudinal study of treatment response trajectories considering both baseline and child-therapist interaction features by means of predictive modeling.
Methods and Procedures: N = 25 preschool children were monitored for one year during Naturalistic Developmental Behavioral Intervention. N = 100 video-recorded sessions were annotated with an observational coding system at four time points, to extract quantitative interaction features.
Outcomes and Results: Baseline and interaction variables were combined to predict response trajectories at one year, and achieved the best predictive performance. The baseline developmental gap, therapist's efficacy in child engagement, respecting children's timing after fast behavioral synchronization, and modulating the interplay to prevent child withdrawal emerged as key factors. Further, changes in interaction patterns in the early phase of the intervention were predictive of the overall response to treatment.
Conclusions and Implications: Clinical implications are discussed, stressing the importance of promoting emotional self-regulation during intervention and the possible relevance of the first period of intervention for later response.

What this paper adds?

Observational methods represent key approaches in research on child development and in clinical contexts due to their non-invasiveness. However, they still suffer from a lack of objectivity and they are often difficult to be used in combination with

* Corresponding author at: Laboratory of Observation, Diagnosis, and Education (ODFLab), Department of Psychology and Cognitive Science, University of Trento, 38068 Rovereto, Italy.
E-mail address: giulio.bertamini@unitn.it (G. Bertamini).

https://doi.org/10.1016/j.ridd.2023.104452
Received 12 December 2022; Received in revised form 18 January 2023; Accepted 1 February 2023
Available online 14 February 2023
0891-4222/

computational techniques. Measuring treatment response is fundamental in research on Autism intervention to understand inter-individual response variability and identifying predictors, mediators, moderators, and mechanisms of change. Developmental models of intervention stress the importance of interpersonal aspects playing an active part during therapy. In fact, intervention always happens and is mediated by the child-therapist interpersonal relationship. However, despite the well-known importance of these aspects in child development, e.g. infant-mother emotional communication, they are still under-investigated in the context of Autism intervention. In this work we employed observational and quantitative approaches on a longitudinal sample of preschool autistic children during one year of Naturalistic Developmental Behavioral Intervention. We used a quantitative observational coding system to extract child-therapist interaction features from video-recorded sessions of intervention at different time-points. Through a predictive modeling pipeline, we identified both baseline and interaction variables and tested to which extent they were able to predict treatment response at 1-year. We employed a developmental outcome measure which can numerically inform individual response trajectories. This effort may provide important indications in terms of key aspects of the intervention trough quantitative investigation. It could also inform clinicians about optimal *therapeutic paths* to maximize treatment efficacy in terms of developmental trajectories, bridging research and clinical practice. The quantitative methods employed in this paper, the robust variable selection process, and the investigation of trajectories of change through predictive modeling represent a valuable approach to study Autism treatment.

1. Introduction

1.1. Research on Autism Intervention

In the last decades, research on Autism treatment dramatically increased, with a consistent amount of Randomized Controlled Trials (RCTs) that fostered the efficacy of different models of intervention (French & Kennedy, 2018; pp. 8, 1282; Sandbank et al. (2020). Longitudinal investigations shed new light on factors associated with a more favorable prognosis, highlighting key elements like starting intervention as soon as possible (MacDuffie et al., 2021, the presence of precursors of language such as communicative gestures (Laister et al., 2021, and intentional vocalizations (McDaniel et al., 2021. This amount of knowledge allowed the definition of clinical best practices that consist of very early screening and diagnosis, in order to employ timely intervention (Whitehouse et al., 2021. Intervening early does not simply imply the possibility of exploiting neural plasticity at its most (Towle et al., 2020, but it also means promoting and supporting child development before accumulating significant delays in developmental milestones that impact cognitive and social functioning (McDonald et al., 2020. In fact, autistic children show a specific set of alterations in social communication and social interaction that eventually impact their developmental profile (Hazlett et al., 2017. In other words, experience-dependent relationship-mediated learning, which spontaneously happens during Typical Development (TD), often results hindered by Autism Spectrum Disorder (ASD) core symptoms (W.M. Association, 2013; A.P. Association, 2013; Piven et al., 2017. Therefore, going through developmental tasks actually represents a hard challenge for these children that, without additional support, face barriers in learning from everyday contexts (e.g. during play) (O'Keeffe & McNally, 2021. However, several studies outlined the existence of diverse developmental trajectories of autistic children (Klintwall et al., 2015. Despite the high variability, the need for methodological improvements, the under-investigation of mechanisms of change, and the difficulty in recruiting large clinical samples of children, research highlighted the progresses that individual children may achieve (French & Kennedy, 2018; pp. 8, 1282; Sandbank et al. (2020). Due to intervention, children may narrow the gap between their mental and chronological age (Klintwall et al., 2015, displaying more homogeneous developmental profiles and significant improvements in language, social communication, and adaptive functioning (Fuller & Kaiser, 2020. In terms of symptom severity, the effects of intervention are more tangible with respect to the areas of social affect, i.e., the cluster A of DSM-5 (W.M. Association (2013); A.P. Association (2013), whereas the area of restricted repetitive behaviors and interests, i.e., the cluster B, seems to be overall more stable (Tiede & Walton, 2019. Interestingly, despite at a preliminary level, some studies also investigated changes at the brain level, arguing a possible tendency towards a neural functioning that resembles TD. Other studies suggested that children may differentially respond to intervention also at the neural level, highlighting the importance of personalized treatment (Stavropoulos, 2017.

1.2. Challenges in measuring treatment response

Despite the promising results, there is still a long way to go for research on Autism intervention, with many areas that require further attention and effort. First, it is difficult to identify and define treatment response: different dimensions may be differentially affected by intervention (e.g. developmental outcomes, cognitive and executive functioning, symptom severity, social and adaptive functioning, everyday life autonomies, as well as parent perspectives). There is also the possibility that an intervention produces subtle changes that may not be identified with common outcome measures, or that require more time to be detectable (Grzadzinski et al., 2020; Stavropoulos, 2017. Secondly, measuring treatment response itself poses several challenges. Outcomes can be proximal-distal and context-bound or generalized (Yoder et al., 2013. Further, the vast majority of outcome measures employed in both research and clinical practice were originally developed for diagnostic purposes or involve measures standardized on a population with TD. These measures are often focused on diagnostic stability and test-retest reliability, and hence not necessarily suitable to detect changes over time (Pijl et al., 2018).

1.3. Towards "precision therapy" in Autism intervention

Investigating outcome measures does not necessarily allow us to foster our understanding of the process of intervention. Beyond baseline prognostic factors, process research remains fundamental to unravel mechanisms of action, mediators, and moderators, eventually understanding the “hows” and “whys” of Autism intervention (Vivanti et al., 2018. A common thread in research on different empirically validated models is represented by significant inhomogeneities in the response. Such variability makes it difficult to compare different groups and suggests the presence of stratifications with different trajectories (Kasari et al., 2018. In fact, some children respond well to intervention, but other children show either lower levels of improvement, a stable trend, or even a worsening one, possibly in response to the same model of intervention. This variability may be due to some baseline variables in terms of biological, cognitive, and symptom severity factors that may pose specific challenges to the intervention. However, research often highlights mixed evidence, depicting a more complex situation (Klinger et al., 2021. For this, an increasing amount of research is pointing out the importance of individualized intervention for treatment response (Kasari et al., 2018; Wetherby et al., 2018. Given the extreme heterogeneity of ASD, it could be further argued that different children may have specific needs, requiring to scaffold and design specific *therapeutic paths* to enhance their response to treatment, instead of one-size-fits-all solutions. Current perspectives on ASD intervention highlight a continuum across different dimensions, including levels of scaffolding, the importance of exploiting child-initiated episodes, intentionality, engagement, intrinsic motivation, and shared pleasure that characterize developmental models of treatment (Frost et al., 2020; Vivanti & Zhong, 2020. Another interesting argument concerns the hypothesis that patient-clinician match plays a role in treatment response based on features of both partners (Goldstein et al., 2020. This may be particularly relevant giving the interpersonal nature of Autism intervention.

1.4. The child-therapist interplay as process aspect

Despite the specificities related to different models of treatment, intervention always happens in the context of the child-therapist interplay, and it is constantly mediated by their dyadic interpersonal relationship. Naturalistic Behavioral Developmental Interventions (NDBIs, (Vivanti & Zhong, 2020) stress the importance of child-therapist interaction patterns and dynamics in terms of structuring, child and dyadic engagement, fostering child motivation, therapist sensitivity, responsiveness, transition handling, and affect sharing (Vivanti et al., 2020. Moreover, evidence underlines the importance of employing manualized interventions that include fidelity assessments and treatment monitoring over time to be more effective (Zitter et al., 2021. In the context of child development, research pointed out the fundamental role of the caregiver-child relationship to guarantee better developmental outcomes. In particular, infant-caregiver emotional communication Feldman (2007), (2017), interpersonal synchrony Leclère et al. (2014), and mutual attunement Provenzi et al. (2018) represent key elements. At the interaction level, these constructs refer to the dynamic, multimodal and reciprocal adaptation of the temporal structure of behaviors between interactive partners Delaherche et al. (2012). These aspects were also investigated with respect to the patient-clinican relationship in psychotherapy research Koole and Tschacher (2016). However, these constructs are still under-investigated in the context of Autism intervention with preschool children, mainly due to the lack of specific instruments able to measure different aspects of interaction behaviors. The vast majority of observational instruments focus on broad and general constructs that are difficult to objectively quantify. Despite operational definitions and standardized training procedures for observational coders, these measures may still suffer from a lack of objectivity. In addition, observational grids are difficult to administer, and there is an increasing need for tailored instruments that could be employed in clinical practice of everyday interventions. In fact, despite RCTs being fundamental for the employment of evidence-based practices (Hume et al., 2021, their external validity could be improved, as would their translational applicability to the clinical setting (Pijl et al., 2018. Indeed, observational research has shown to produce results that seem to be coherent with RCTs (Anglemyer et al., 2014 and may be helpful to employ non-invasive instruments for treatment monitoring and evaluation (Vandenbroucke, 2008. In the way of precision medicine, this translational effort may represent steps towards the implementation of precision approaches in clinical psychology, improving the paradigm of early diagnosis and early intervention (Dawson & Sapiro, 2019; Jacob et al., 2019.

1.5. Aims and hypotheses

The aim of this study was the preliminary investigation of the role of the child-therapist interaction on developmental outcomes during intervention by employing an observational quantitative approach. More in general, we wished to shed light on the longitudinal predictive relation between treatment response trajectories, i.e., changes over time of developmental outcomes, and features of the child-therapist interaction, i.e., the way by which the therapist interacts with the child during their exchange. We argued that, by quantitatively measuring the child-therapist interaction patterns it was possible to extract features related to the underlying more general constructs of interpersonal synchrony and mutual attunement. To objectively quantify the exchange, we measured the quantity and the quality of the child-therapist interaction patterns by considering both the structural and functional aspects of their dyadic exchange. We considered frequencies and proportions of behaviors, as well as their durations and latencies. In addition to the behavioral patterns, latencies and durations may be particularly related to the onset of interpersonal synchrony, and especially mutual attunement when referred to sequences of reciprocal behaviors. In our analysis, key descriptors traced the different kinds of events that characterized the interaction, e.g. the therapist-initiated attempts to start a social routine or play activity, their rate of success, the child’s intentionality signals, and the modalities by which the interplay can be terminated or interrupted (Bertamini et al., 2021. We modeled treatment response considering a combination of baseline predictors and interaction variables. Coherently with literature, we included clinical measures in terms of developmental profile (including different developmental domains) and symptoms severity as baseline variables. We also included the age at intake and treatment intensity (Towle et al., 2020. We hypothesized that interaction variables may have a significant role in modeling treatment response, independently from baseline predictors. We further studied whether the integration of interaction features improved predictive performance. The analysis also aimed to identify new options for modulating child-therapist interplay in order to optimize intervention. More specifically, in terms of research hypothesis, we expected to observe that: .

1. An early intake, lower developmental delay, and lower symptom severity guarantee better developmental trajectories (Towle et al., 2020; Zachor & Ben-Itzchak, 2017
2. Child-therapist mutual attunement and behavioral synchronization predict response trajectories to some extent (Koole & Tschacher, 2016; Mayo & Gordon, 2020
3. The combination of baseline variables and interaction features significantly improves outcome prediction

We explored these hypotheses by means of linear regression models after a three-step variable selection process combining Machine Learning (ML) and statistical approaches.

2. Material and methods

2.1. Participants

N = 25 children (M chronological age = 38.5 months (10.2) [23–57]; M developmental age = 26.2 months (7.3) [14–45]) with a diagnosis of ASD (N = 22 males) were monitored for about one year (M = 15.2 (4.9) months) during early NDBI at the Lab X, a clinical and research center of the Department of Psychology and Cognitive Science of the University of X, specialized in the functional diagnosis of neurodevelopmental disorders and early individualized developmental intervention focused on Autism. At first, children undertook a complete functional evaluation using gold-standard procedures for the assessment of neurodevelopmental disorders, followed by a personalized intervention (2–4 h/week) with trained therapists. After about one year, the children were monitored by independent clinicians to assess the evolution of the developmental, cognitive, and symptoms severity profiles over time. The ASD diagnosis was conducted following the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5, (W.M. Association (2013); A. P. Association (2013)) criteria, and confirmed by the administration of gold standard instruments. The sample was formed based on the database of clinical data of the Lab. The socioeconomic status of the families (SES), assessed through the four-factor index of social status, indicated a middle status for the participants (Hollingshead, 1975. This study included a total of N = 17 therapists. In N = 14 cases, there was a therapist change at some time point during intervention, due to logistic or clinical reasons. Nonetheless, all therapists followed the same protocols and objectives, and received the same training and supervision. Clinical reasons for therapist change include the need for the child to generalize the acquired skills. This is typical when a high degree of mutual attunement is rapidly acquired by the dyad, with a substantial progression in developmental objectives, with a lack of generalization to different situations, people, and contexts, i.e., proximal and context-bound. Details about the model of intervention employed by the lab are reported in Supplementary Materials. Data were collected between 2015 and 2021.

2.2. Procedures

All the procedures employed for this study were in accordance with the last version of the Declaration of Helsinki (A.P. Association (2013); W.M. Association (2013), and were approved by the Research Ethics Board of the University of X (Protocol number: 2020–042). Children's developmental outcomes were assessed through the Griffiths Mental Development Scales - Edition Revised (GMDS-ER, (Luiz et al., 2006). Symptoms severity was evaluated by the Autism Diagnostic Observation Schedule - Second Edition (ADOS-2, (Lord et al., 2012). Details about the clinical measures employed in this study are reported in Supplementary Materials. For each participant, four sessions of intervention video-recorded by bird's eye cameras were extracted. The first one was selected immediately after the diagnostic assessment (T0), another after 3–4 months (T1), a mid-session after 6–8 months (T2), and the last available session before the monitoring assessment after 12–16 months (T3). A total of N = 100 videos were extracted. Sessions were annotated by a trained expert observer using the Interaction Coding System (ICS), a validated observational quantitative coding system for the child-therapist interaction annotation during continuous time-event sequential micro-coding. It showed good validity and reliability, especially with the child involvement and responsiveness dimensions of the Emotional Availability Scales (EAS), a standardized instruments extensively used in research (Biringen, 2008 [Anonymous ref]. The ICS allows the characterization of the child-therapist interactions in terms of Units of Interaction (UIs, i.e., the precursor phase before the beginning of the actual interplay characterized by thearpist's or child's proposals or child's intentionality signals) and a Shared Activity state (SA, i.e., the actual interplay where both the social partners are engaged in a reciprocal social play activity or routine). The code allows the automatic extraction of quantitative behavioral descriptors of the interaction in terms of frequencies, proportions, latencies, durations, and succes rates. More details about the ICS are included in Supplementary Materials. The observer was masked to children's clinical profiles and coded the different videos in random order. The coding window was set to 20 min, extracted from the middle of the session to avoid the initial -and potentially more stressful- moments of interaction and the last ones, where the child may be more tired or face more challenges. The duration of the intervention sessions usually ranged between 60 and 90 min. The videos were annotated using BORIS (Behavioral Observation Research Interactive Software), an open source tool for behavioral observational research (Friard & Gamba, 2016).

2.3. Statistical analysis

The raw annotations produced by the BORIS coding interface were first preprocessed by a pipeline of Python scripts to automatically compute the interaction feature set, aggregated at the session-level. All statistical analysis and model development were performed using the R environment for statistical computing. For statistical inference, data were tested for normality, i.e., Shapiro-Wilk normality test and homogeneity of variances, i.e., Levene's test. When appropriate, paired Student's t-tests were employed to test for longitudinal changes in outcome or process measures. Otherwise, paired Wilcoxon-Mann-Whitney signed-rank with continuity correction tests were used. Effect sizes and correlations were computed using Cohen's *d* and Pearson's product-moment correlation coefficient (r). A Bayes Factor (BF) analysis was performed to improve statistical inference. In this study, we aimed at predicting treatment response in terms of Learning Rate (LR, (Klintwall et al., 2015). The LR, i.e., *the ratio between the variation in Developmental age-equivalents over time and the time elapsed between the assessments* represent a particularly suitable metric to assess changes over time with respect to TD trajectories. It is also useful in presence of differences in the time elapsed between the assessments, as well as in baseline chronological ages. LRs are computed with the following formula:

$$LR = \frac{DevAge_{\mathrm{TPost}} - DevAge_{\mathrm{TPre}}}{TPost - TPre}. \tag{1}$$

For detailed information about the LR see Supplementary Materials. We aimed at employ predictive models by means of linear regression including baseline clinical variables, interaction variables extracted at different time points through the ICS, and the combination of the two. We computed interaction descriptors at the four time points, as well as the variations between them. Variations were calculated using the formula:

$$Var_{\mathrm{T1\text{-}T0}} = \frac{Desc_{\mathrm{T1}} - Desc_{\mathrm{T0}}}{Desc_{\mathrm{T0}}} \tag{2}$$

An analytic pipeline for model selection and comparison was employed based on different evaluation metrics. Models were tested for assumptions in terms of homoscedasticity, i.e., studentized Breusch-Pagan test, normality of residuals, Cook's distance, and autocorrelation, i.e., Durbin-Watson test. Data were standardized by a robust normalization based on the 10th and the 90th percentile difference to account for possible outliers. Multicollinearity was addressed by computing the Variance Inflation Factor (VIF) Akinwande et al. (2015). To take into account the small sample size, the maximum number of predictors for candidate models was set to N = 3 (Austin & Steyerberg, 2015. The initial set of variables included in the analysis was selected by a 3-level pipeline combining ML and statistical techniques.

First, we employed a feature selection based on correlations between interaction variables. Specifically, after computing the correlation matrix we selected pairs of features with $|r| >= 0.55$, as a conservative threshold. Coherently with our predictive aims, we did not merge correlated features, but we excluded the features prioritizing variables measured earlier in time and variations between time points. We also prioritized baseline variables over interaction features, and sub-quotients with respect to more general scores to investigate more specific aspects.

As a second step, we employed a ML model, i.e., Random Forest (RF), a decision tree-based approach using bootstrap and random subset of features. RF can handle up to more than 100 features and can be used to assess variable importance through different metrics. The correlation-based selected features were used to fit a RF model. Variable importance was assessed through & IncMSE and permutation importance, and the common features in the first N = 10 most important variables were retained for the final selection step. We also included in the analysis three variables to control for the effect of treatment intensity, therapist, and therapist change. We considered both therapist change between T0 and T3 (N = 14), and a more conservative variable that differentiate between clinical path were all the time-points involved always the same therapist (N = 12). Finally, we included therapist change in the first period of intervention (N = 7), i.e. between T0 and T1.

Ultimately, we performed correlations between the remaining predictors and the dependent variable to select the most suitable predictors to build linear regression models. To handle multiple comparisons, we applied the Benjamini-Hochberg correction with False Discovery Rate (FDR) = 0.1, as recommended by research literature (Li & Barber, 2018). Candidate linear regression models were formed starting from baseline variables, to both evaluate their impact on treatment response and to account for them as covariates in more complex ones. The best model was selected based on multiple indexes to compare different evaluation metrics: p-value, Adjusted-R^2, Bayesian Information Criterion (BIC), and BF. We employed the Corrected Akaike Information Criterion (AICc) to account for the reduced sample size. Model comparison was performed by ANOVA with F-test when appropriate, i.e., for nested models, and considering all the other evaluation metrics. Finally, the resulting candidate models were included in a Cross-Validation (CV) step to assess model performance estimators. Given the reduced sample size, we employed a Leave-One-Out CV (LOOCV) to assess predictive performance, evaluated by Minimum Average Error (MAE) and Root Mean Square Error (RMSE). We also employed Cross-Validated R^2 (CV-R^2). Finally, we bootstrapped the models using N = 1000 replicates to compute 95% confidence intervals in order to assess the stability of parameter estimates and model evaluation metrics. We particularly focused on Adjusted-R^2, as suggested by research evidence in terms of informativeness and interpretability (Chicco et al., 2021).

3. Results

3.1. Longitudinal changes

Descriptive statistics and changes over time in developmental outcome measures and symptom severity pre- and post- intervention are reported in Supplementary Materials. The average LR was (M = 0.94 (0.49) [0.22–2.20]).

Children showed significant changes during intervention mainly in the areas of language and social affect. Also, the rate between mental and chronological age significantly increased.

3.1.1. Variable selection

The full variable selection process is described in detail in Supplementary Materials. Table 1 summarizes the final variable selection with Benjamini-Hochberg correction for multiple comparisons. The correlation matrix is reported in Table 2. Fig. 1 reports trajectories of change of the selected predictors. Fig. 2 reports paired relationships between predictors and the outcome variable (LR). RF predictor importance excluded the effect of therapist and therapist change at all the time points as relevant features for the response to treatment. N = 4 candidate predictors were selected as significantly correlated with the LR and were considered to build the three models. We added one more interaction predictor to the analysis, LATENCY_TPCA.T1, i.e., the mean latency between a therapist proposal and the child acceptance. We included this additional predictor, despite not being significantly correlated with the outcome variable, for two reasons: (i) it has been selected by RF importance as one of the most important predictors; and (ii) it is actually a clinically relevant parameter to be considered.

3.2. Aim 1: baseline feature models

Considering only baseline variables, i.e., the Developmental/Chronological age ratio (RATE_DC) and the Chronological age (AGE), see Table 1, the procedure selected the univariate model M1 for the Developmental/Chronological age ratio (RATE_DC), with b = 0.62 ($t = 4.81$; $p < 0.001$); the intercept did not reach statistical significance (b = 0.16; t = 1.83; p = 0.080). The model was significant ($F(2,22) = 23.16$; $p < 0.001$; AICc = 17.99) and explained a fair proportion of the observed variance (Adj-$R^2 = 0.48$). Cross-validation indicated a slight reduction in R^2 (MAE = 0.26; RMSE = 0.33; CV-$R^2 = 0.41$). The bootstrapped model resulted stable (see Supplementary Material). The model satisfied all the assumptions.

A more complex bivariate model including the chronological age term (AGE) resulted non-significantly different from the simpler one (F = 3.317; p = 0.082).

3.3. Aim 2: Interaction features models

M2 included the variation between T1 and T0 in the latency required to actually start the exchange after a unit of interaction (LATENCY_SA_DELTA10), with b = 0.41 (t = 2.48; p = 0.021); and the variation between T1 and T0 in the proportional frequency of child's interruption (P_CX_DELTA10), ($b = -0.37$; $t = -2.42$; p = 0.02). The intercept in M2 was significant (b = 0.43; t = 3.51; p = 0.002). In brief, a decrease in child's interruptions and an increase in latency to start the exchange at three months were associated with better response trajectories. The model was significant ($F(2,22) = 6.718$; p = 0.005; AICc = 26.35) and explained a fair proportion of the observed variance (Adj-$R^2 = 0.32$). Cross-validation indicated a slight reduction in R^2 (MAE = 0.34; RMSE = 0.39; CV-R^2 = 0.2). The bootstrapped model resulted stable. The model satisfied all the assumptions. The indicated low multicollinearity (VIF < 1.10 for all the predictors).

The bivariate model was significantly better than both the univariate models with P_CX_DELTA10 (F = 6.146; p = 0.021) and LATENCY_SA_DELTA10 (F = 5.841; p = 0.024), respectively.

M3 included the child's latency in response to therapist proposals measured at T1 (LATENCY_TPCA.T1), with $b = -0.46$ ($t = -3.38$; p = 0.003); and the variation between T1 and T0 in the latency required to actually start the exchange after a unit of

Table 1
Final variable selection with correlation between predictors and dependent variable (LR) with Benjamini-Hochberg correction for multiple comparisons (FDR = 0.10). Variables are ordered by rank.

Variable	Description	r	p-value	Rank	Threshold	Corrected p-value
rate_DC	Developmental / Chronological age ratio	0.71	0.001	1	0.013	**0.001**
Age (months)	Chronological age	− 0.56	0.004	2	0.025	**0.015**
LATENCY_SA_DELTA10(s)	Mean latency between IUs and start of activity (variation T1-T0)	0.46	0.02	3	0.038	**0.046**
P_CX_DELTA10	Proportional frequency of child's withdrawals (variation T1-T0)	− 0.45	0.023	4	0.05	**0.046**
LATENCY_TPCA.T1(s)	Mean latency between therapist's proposals and child's acceptances (T1)	− 0.33	0.108	5	0.063	0.173
QA(Z)	Locomotor Quotient	0.30	0.152	6	0.075	0.203
LATENCY_SYNC_DELTA10 (s)	Mean latency between synchrony codes (variation T1-T0)	− 0.23	0.27	7	0.088	0.309
LATENCY_TPCA_DELTA10 (s)	Mean latency between therapist's proposals and child's acceptances (variation T1-T0)	− 0.02	0.938	8	0.1	0.938

Table 2
Correlation matrix between selected predictors. Pearson's correlation coefficient (r).

		1	2	3	4
1	**rate_DC**				
2	**Age**	**− 0.47 (p = 0.018)**			
3	**LATENCY_SA_DELTA10**	0.30	**− 0.53 (p = 0.006)**		
4	**P_CX_DELTA10**	− 0.16	0.12	− 0.11	
5	**LATENCY_TPCA.T1**	− 0.19	− 0.08	0.34	**0.46 (p = 0.020)**

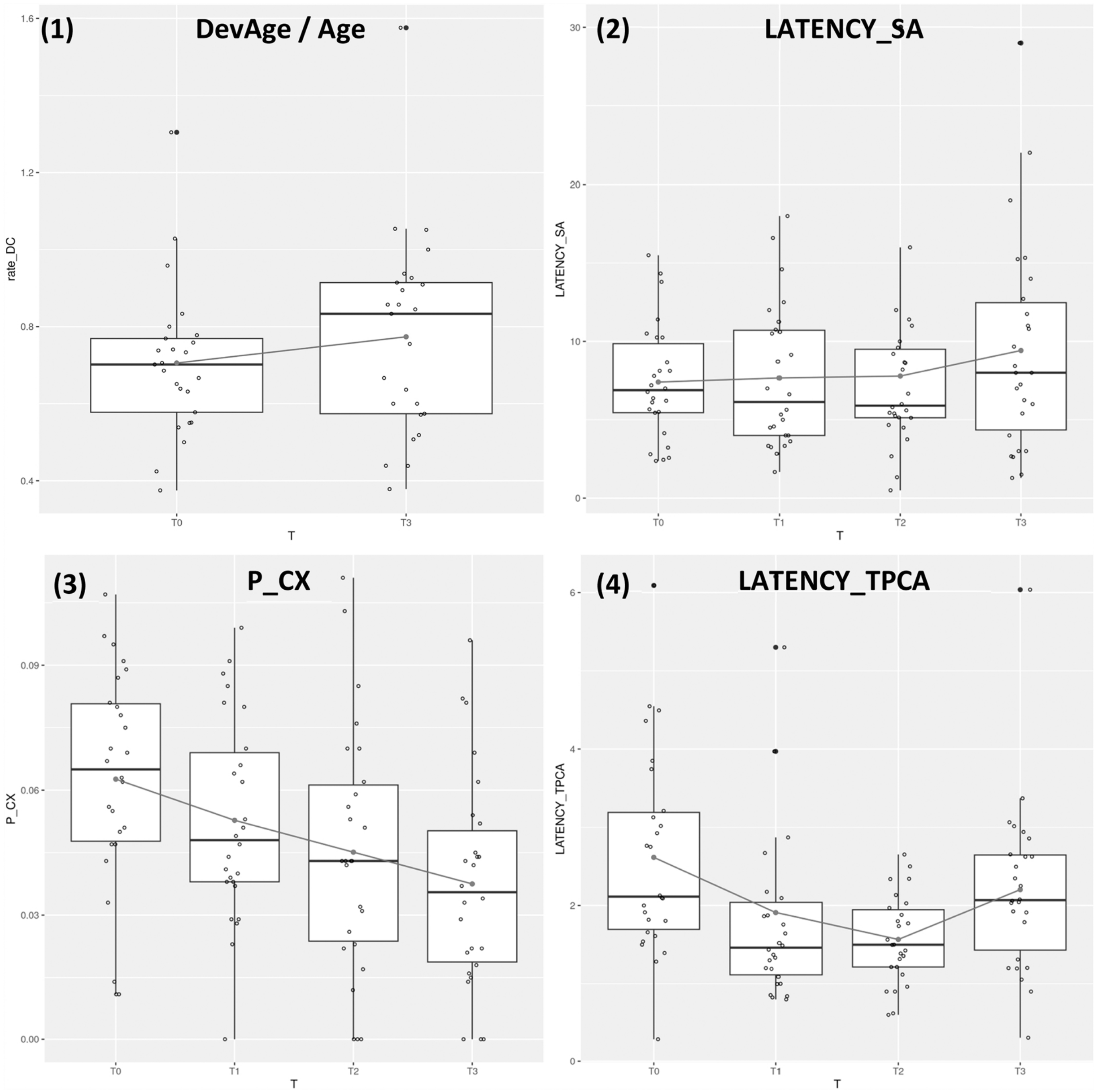

Fig. 1. Changes in main variables Trajectory between T0 and T3 of (1) Developmental / Chronological age ratio; (2) Mean latency between a Unit of Interaction and the actual beginning of the interchange; (3) Proportional frequency of child's withdrawal; (4) Mean latency between therapist's proposal and child's acceptance.

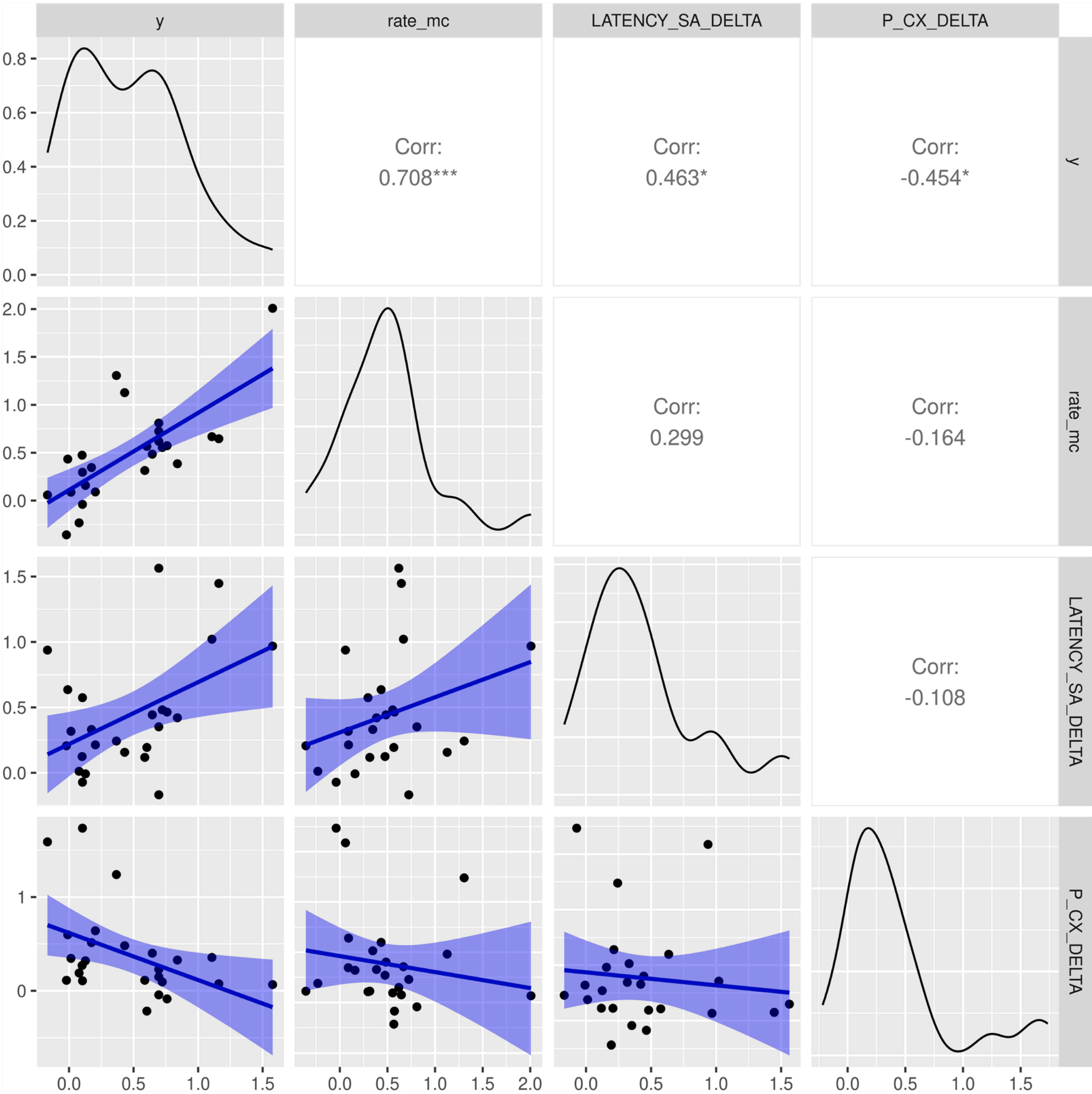

Fig. 2. Variable relationships Paired plot between independent variable predictors and the Learning Rate.

interaction (LATENCY_SA_DELTA10), with b = 0.64 (t = 4.00; p < 0.001). The intercept in M3 was also significant (b = 0.37; t = 3.77; p = 0.001). In brief, prompt responses to therapist's attempts to engage the child in the interaction after the first period of intervention, together with an increase in the latency to actually start the exchange during the first three months were associated with better response trajectories. The model was significant (F(2,22) = 10.27; p < 0.001; AICc = 21.78) and explained a fair proportion of the observed variance (Adj-R^2 = 0.44). Cross-validation indicated a slight reduction in R^2 (MAE = 0.28; RMSE = 0.34; CV-R^2 = 0.36). The bootstrapped model resulted stable. The model satisfied all the assumptions. The VIF indicated low multicollinearity (VIF < 1.15 for all the predictors).

The bivariate model was significantly better than both the univariate models with LATENCY_TPCA.T1 (F = 15.93 p < 0.001) and LATENCY_SA_DELTA10 (F = 11.42; p = 0.002).

The more complex model containing all three predictors was significantly better than M2 (F = 5.42; p = 0.030), but non-significantly better than M2 (F = 1.00; p = 0.33). Further, P_CX_DELTA10 were not significant as a predictor, and hence we included the two bivariate models.

3.4. Aim 3: integration of baseline and interaction features models

When integrating baseline variables and interaction features the first integrated model (M4) started from the strongest interaction predictor (P_CX_DELTA10, i.e., the variation in the proportional frequency of child's withdrawals). The model integrated M1 including the RATE_DC feature in our ($b = 0.57$; $t = 4.88$; $p < 0.001$) and P_CX_DELTA10 ($b = -0.313$; $t = -2.6$; $p = 0.016$). The intercept was also significant ($b = 0.31$; $t = 3.17$; $p = 0.004$). The model was significant ($F(2,22) = 17.87$; $p < 0.001$; AICc = 14.15), and explained a greater proportion of the observed variance with respect to M1, M2, and M3 (Adj-R^2 = 0.58). Cross-validation indicated a slight reduction in R^2 (MAE = 0.25; RMSE = 0.30; CV-R^2 = 0.51). The bootstrapped model resulted stable. The model satisfied all the assumptions. The VIF indicated low multicollinearity (VIF < 1.10 for all the predictors).

The second combined model (M5) integrated M1 and M2, and included the RATE_DC feature of model M1 ($b = 0.63$; $t = 5.58$; $p < 0.001$), combined with P_CX_DELTA, i.e., the variation in the proportional frequency of child's withdrawals ($b = -0.30$; $t = -2.64$; $p = 0.015$), and LATENCY_SA_DELTA, i.e., the variation between T1 and T0 in the latency required to actually start the exchange ($b = 0.25$; $t = 1.97$; $p = 0.064$). The intercept was significant ($b = 0.22$; $t = 2.22$; $p = 0.038$). The model was significant ($F(2,22) = 14.72$; $p < 0.001$; AICc = 13.12) and explained a higher proportion of the observed variance (Adj-R^2 = 0.63) than M1, M2, and M3. Cross-validation indicated a reduction in R^2 (MAE = 0.25; RMSE = 0.30; CV-R^2 = 0.52). The bootstrapped model resulted stable. The model satisfied all the assumptions. The VIF indicated low multicollinearity (VIF < 1.13 for all the predictors).

The last combined model (M6) integrated M1 and M3, and included the RATE_DC feature of model M1 ($b = 0.44$; $t = 3.73$; $p = 0.001$), combined with LATENCY_SA_DELTA ($b = 0.44$; $t = 3.16$; $p = 0.005$), and LATENCY_TPCA.T1 ($b = -0.32$; $t = -2.82$; $p = 0.010$). The intercept did not reach significance ($b = 0.19$; $t = 2.05$; $p = 0.053$) The model was significant ($F(3,21) = 15.49$; $p < 0.001$; AICc = 12.24) and explained a higher proportion of the observed variance (Adj-R^2 = 0.64) than all the other models. Cross-validation indicated a mild reduction in R^2 (CV-R^2 = 0.6). The bootstrapped model resulted stable. The model satisfied all the assumptions. The VIF indicated low multicollinearity (VIF < 1.30 for all the predictors).

The model comparison analysis in described in detail in Supplementary Materials. Fig. 3 graphically represent the predictive performance of the six presented models.

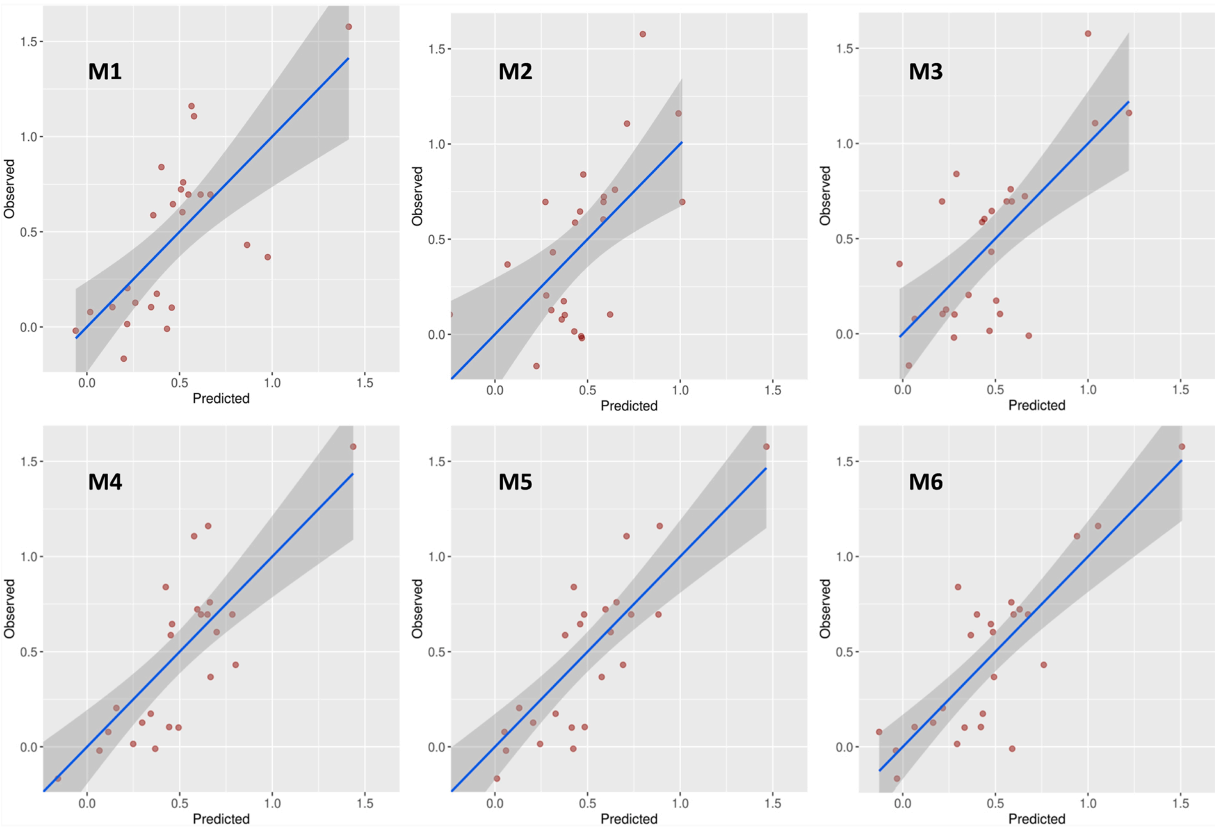

Fig. 3. Comparison between model predictive performance Predicted vs. Observed values. M1: baseline model; M2: interaction model; M3: interaction model; M4: combined model; M5: combined model; M6: combined model.

4. Discussion

The general aim of this preliminary work was the longitudinal investigation of the predictive association between interaction variables and ASD treatment response trajectories over time. We employed a quantitative observational method that produces objective indexes reflecting the structure, in terms of different types of behaviors, and the temporal dynamics of the child-therapist interaction. More specifically, we wanted to assess if and how interaction features had an impact on treatment response, and whether they could be combined with baseline variables to improve model performance and shed light on mediators of change. To our knowledge, an analytic design based on robust variable selection and the investigation of trajectories of change through predictive modeling was never employed in the context of Autism treatment. Modeling of treatment response investigated the role of interaction features first in comparison and then in combination with the clinical variables at baseline.

First, we found that the ratio between children's mental and chronological age at baseline is consistently the most important predictor of treatment response at one year. This result combines in one parameter two known features, i.e., outcomes are higher when starting the intervention early, and from a less severe clinical profile (Towle et al., 2020. In our data, profiles characterized by a lower pre-treatment developmental delay were associated with better response trajectories. Interestingly, the absolute mental age at intake and the chronological age separately were not good predictors of the outcome, if compared with their interaction. Further, with our model selection framework symptom severity appeared to only moderately impact the outcome measure. This is in line with an amount of recent literature focusing on pre-clinical interventions for conditions "at increased probability" for neurodevelopmental alterations during wide and very early screenings (MacDuffie et al., 2021; Whitehouse et al., 2021. From the clinical standpoint, this result points out the critical importance of avoiding the accumulation of developmental delays.

As a second step in our analysis, we tried to predict treatment response based on interaction variables only. M2 pointed out the importance of (1) a reduction in child's withdrawals in the first period of intervention and (2) an increase in the time needed by the child to actually start the interchange after being engaged. Further, M3 also pointed out another dynamic aspect that goes in the opposite direction. Better response trajectories were predicted by (1) the same increase in the latency to start the interaction, but associated with (2) a reduction in the latency of acceptance of therapist's proposals by the child. Especially in the firsts phases of the intervention, the therapist plays a more active role in trying to involve the child in adequate social routines based on the individual functioning and exploiting intrinsic motivation and shared pleasure. This is also in line with the developmental framework that characterizes NDBIs (Vivanti et al., 2020. Taken together, these models seem to indicate that the interaction impacts developmental outcomes and suggest that specific dyadic aspects may play a role. From a clinical standpoint, this result highlights the importance of gaining a rapid child-therapist mutual attunement, and stresses the importance of following child's timing in initiating the interplay. Even more importantly, it seems to be relevant to foster these aspects without overshadowing emotional regulation. In fact, social interplays often exceed children's abilities to self-regulate the interaction. In turn, these dynamics may actually prevent their possibility to participate in the interplay, ultimately leading to child withdrawal. Therefore, it may be critical for therapists to keep these aspects in mind during intervention, promoting children's self-regulation and supporting emotional aspects. It may also be important to improve therapists' scaffolding abilities and regulatory strategies. Further, respecting children's timing during transitions and while involving them in the interplay appears important and may represent a challenging aspect for therapists. In fact, they may mistakenly interpret children's pauses as denials or failed attempts to scaffold the interaction. Especially when preceded by a prompt response by the child, therapist's proposals may actually be effective if the child's timing is respected. Interestingly, these insights may link attunement with emotional-regulatory aspects in a complex and dynamic way, coherently with recent research in the field of bio-behavioral synchrony (Cai et al., 2018; Galbusera et al., 2019; Goldstein et al., 2020; Koole & Tschacher, 2016; Mayo & Gordon, 2020.

As a third step, we investigated the possibility to combine baseline variables and interaction measures, focusing on our hypothesis that this may significantly improve outcome prediction. Coherently with baseline and interaction only models, the first combined model (M4) pointed out that combining the baseline developmental delay in terms of Developmental/Chronological age ratio with the variation of the proportion of child's withdrawals significantly improved the baseline model (M1). Clinically, this result seems to stress the specific relevance of successfully modulating the interplay taking into account emotional regulation aspects in the first period of intervention. From our analysis, the duration of the exchanges and the absolute frequency of withdrawals did not seem to have the same importance as the variation in this proportion. This may narrow the focus down to an overall aspect of therapy, underlining the importance of experiencing a social interaction that gradually adapts to the child's developmental level, becoming more and more manageable by the child. This seems also to be in line with evidence about the importance of infant emotional communication for child development (Feldman, 2007, which remains under-investigated with respect to treatment. In turn, this may foster mutual attunement, decrease social avoidance and lay the foundations of the subsequent work on more specific and structured developmental goals and abilities. M5 introduced again the importance of respecting and following child's timing and M6 stressed the importance to pay particular attention to the first phases of the interaction. These moments may be clinically important but difficult to exploit. Therapists need to understand child's individual functioning and successfully adapt their proposals and attempts focusing on intrinsic motivation to structure a playful routine. This may happen through different intersubjective modalities, from primary intersubjectivity involving direct interaction to the more sophisticated secondary intersubjective routines involving object triangulation and triadic coordination (Vivanti et al., 2020).

In summary, the clinical interpretation of these first results seem to be coherent with a developmental framework of intervention with a key focus on emotional regulation aspects. The preliminary analysis seems to point out that the first months of intervention may be particularly relevant for treatment response trajectories. Promoting emotional self-regulation may represent a base building block of the intervention (Chetcuti et al., 2020; Green, 2006; Reyes et al., 2019. Interventions with better response trajectories seemed to be characterized by the coexistence of different aspects: .

- A developmental profile characterized by a lower degree of accumulated delay in developmental milestones at in-take
- The need to modulate the interaction preventing child's withdrawals
- The importance of the precursors of the interaction in terms of promoting a rapid attunement in response to child's attempt to engage the child while respecting child's timing

From a longitudinal perspective, the results indicate that considering just the initial evolution may be enough to estimate what will happen in the future. From a clinical standpoint, this may suggest the importance of allocating specific resources focused on this period to regularly monitor the evolution of child-therapist interaction. Finally, as a transversal and coherent aspect of our analysis, it seems to emerge the importance of working on emotional and regulatory aspects from the very beginning of the intervention, to foster dyadic attunement and to promote children's capability to participate in the interplay (Chetcuti et al., 2020; Green, 2006; Reyes et al., 2019).

4.1. Limitations

This study comes with several limitations, the main one being represented by reduced sample size, which we tried to compensate for by limiting the number of independent variables through an independent and controlled process of variable selection. The reduced sample size eventually prevents the generalizability of the results to a wider population and makes the models suffer from limited power. Therefore, the reproducibility of these preliminary results should be addressed as a main priority. Further, the majority of the sample consisted of male children, and hence gender-based differences could not be investigated, and results may not generalize to the female population. With respect to treatment, despite being manualized and including the regular administration of observational grids, fidelity assessment was undertaken by means of expert supervision with video-feedback, without collecting quantitative data. Finally, a consistent amount of cases experienced therapist change. To control for this, we included therapist change variables between the different time points. Those variables were ruled out by the RF importance variable selection. In addition to this, the same training, regular supervision, and adherence to a structured model may enable the maintenance of coherent interaction styles tailored to the individual patient. Future research should aim at employing this methodology on larger samples, including symptom severity as an outcome measure and the other developmental subdomains that may be differentially impacted by intervention. Further, the role of therapist should be more systematically investigated. Fidelity could also be more rigorously measured. Finally, automating the annotation procedure could greatly improve translational applicability.

Credit author statement

Giulio Bertamini, Arianna Bentenuto, Cesare Furlanello, and Paola Venuti contributed to the study conception and design. Material preparation and data collection were performed by Giulio Bertamini and Silvia Perzolli. Eleonora Paolizzi annotated the videos, as expert and trained observer. Statistical methodology was designed by Giulio Bertamini and Cesare Furlanello. Statistical analysis was performed by Giulio Bertamini. The first draft of the manuscript was written by Giulio Bertamini Silvia Perzolli. All authors commented on previous versions of the manuscript. All authors contributed to review and edit the manuscript. All authors read and approved the final manuscript. Paola Venuti and Cesare Furlanello supervised the study, for the clinical and statistical modeling parts, respectively.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ridd.2023.104452.

References

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. *Open Journal of Statistics, 05*, 754–767. https://doi.org/10.4236/ojs.2015.57075

Anglemyer, A., Horvath, H. T., & Bero, L. (2014). Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database of Systematic Reviews, 2014*, Article 000034. https://doi.org/10.1002/14651858.mr000034.pub2

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association,, 10.1176/appi.books.9780890425596.

Association, W. M. (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*, 2191–2194. https://doi.org/10.1001/jama.2013.281053

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology, 68*, 627–636. https://doi.org/10.1016/j.jclinepi.2014.12.014

Bertamini, G., Bentenuto, A., Perzolli, S., Paolizzi, E., Furlanello, C., & Venuti, P. (2021). Quantifying the child–therapist interaction in ASD intervention: An observational coding system. *Brain Sciences, 11*, 366. https://doi.org/10.3390/brainsci11030366

Biringen, Z. (2008). The Emotional Availability (EA) Scales manual: Part 1.Infancy/Early Childhood Version.

Cai, R. Y., Richdale, A. L., Uljarević, M., Dissanayake, C., & Samson, A. C. (2018). Emotion regulation in autism spectrum disorder: Where we are and where we need to go. *Autism Research, 11*, 962–978. https://doi.org/10.1002/aur.1968

Chetcuti, L., Uljarevic, M., Varcin, K. J., Boutrus, M., Wan, M. W., Green, J., Iacono, T., Dissanayake, C., Whitehouse, A. J. O., & Hudry, K. (2020). Subgroups of temperament associated with social emotional difficulties in infants with early signs of autism. *Autism Research, 13*, 2094–2101. https://doi.org/10.1002/aur.2381

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, Article e623. https://doi.org/10.7717/peerj-cs.623

Dawson, G., & Sapiro, G. (2019). Potential for digital behavioral measurement tools to transform the detection and diagnosis of autism spectrum disorder. *JAMA Pediatrics, 173*, 305. https://doi.org/10.1001/jamapediatrics.2018.5269

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal Synchrony: A Survey of Evaluation Methods across Disciplines. *IEEE Transactions on Affective Computing, 3*, 349–365. https://doi.org/10.1109/T-AFFC.2012.12

Feldman, R. (2007). Parent–infant synchrony. *Current Directions in Psychological Science, 16*, 340–345. https://doi.org/10.1111/j.1467-8721.2007.00532.x

Feldman, R. (2017). The neurobiology of human attachments. *Trends in Cognitive Sciences, 21*, 80–99. https://doi.org/10.1016/j.tics.2016.11.007

French, L., & Kennedy, E. (2018). Annual research review: Early intervention for infants and young children with, or at-risk of, autism spectrum disorder: a systematic review. *Journal of Child Psychology and Psychiatry, 59*, 444–456. https://doi.org/10.1111/jc

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution, 7*, 1325–1330. https://doi.org/10.1111/2041-210X.12584

Frost, K. M., Brian, J., Gengoux, G. W., Hardan, A., Rieth, S. R., Stahmer, A., & Ingersoll, B. (2020). Identifying and measuring the common elements of naturalistic developmental behavioral interventions for autism spectrum disorder: Development of the ndbi-fi. *Autism, 24*, 2285–2297. https://doi.org/10.1177/1362361320944011

Fuller, E., & Kaiser, A. (2020). The effects of early intervention on social communication outcomes for children with autism spectrum disorder: A meta-analysis. *Journal of Autism and Developmental Disorders, 50*, 1683–1700. https://doi.org/10.1007/s10803-019-03927-z

Galbusera, L., Finn, M., & Tschacher, W. (2019). Interpersonal synchrony feels good but impedes self-regulation of affect. *Scientific Reports, 9*, 14691. https://doi.org/10.1038/s41598-019-50960-0

Goldstein, P., Losin, E. A. R., Anderson, S. R., Schelkun, V. R., & Wager, T. D. (2020). Clinician-patient movement synchrony mediates social group effects on interpersonal trust and perceived pain. *The Journal of Pain, 21*, 1160–1174. https://doi.org/10.1016/j.jpain.2020.03.001

Green, J. (2006). Annotation: the therapeutic alliance-a significant but neglected variable in child mental health treatment studies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47*, 425–435. https://doi.org/10.1111/j.1469-7610.2005.01516.x

Grzadzinski, R., Janvier, D., & Kim, S. H. (2020). Recent developments in treatment outcome measures for young children with autism spectrum disorder (ASD). *Seminars in Pediatric Neurology, 34*, Article 100806. https://doi.org/10.1016/j.spen.2020.100806

Hazlett, H., Gu, H., Munsell, B., Kim, S., Styner, M., Wolff, J., Elison, J., Swanson, M., Zhu, H., Botteron, K., Collins, D., Constantino, J., Dager, S., Estes, A., Evans, A., Fonov, V., Gerig, G., Kostopoulos, P., McKinstry, R., Pandey, J., & Analysis, S. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature, 542*, 348–351. https://doi.org/10.1038/nature21369

Hollingshead, A. (1975). Four Factor Index of Social Status. *Yale Journal of Sociology, 8*, 21–51.

Hume, K., Steinbrenner, J., Odom, S., Morin, K., Nowell, S., Tomaszewski, B., Szendrey, S., McIntyre, N., Yücesoy-Özkan, S., & Savage, M. (2021). Evidence-based practices for children, youth, and young adults with Autism. *Third Generation Review Journal of Autism and Developmental Disorders, 51*, 4013–4032. https://doi.org/10.1007/s10803-020-04844-2

Jacob, S., Wolff, J., Steinbach, M., Doyle, C., Kumar, V., & Elison, J. (2019). Neurodevelopmental heterogeneity and computational approaches for understanding autism. *Translational Psychiatry, 9*, 63. https://doi.org/10.1038/s41398-019-0390-0

Kasari, C., Sturm, A., & Shih, W. (2018). SMARTer approach to personalizing intervention for children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 61*, 2629–2640. https://doi.org/10.1044/2018_jslhr-l-rsaut-18-0029

Klinger, L., Cook, M., & Dudley, K. (2021). Predictors and moderators of treatment efficacy in children and adolescents with autism spectrum disorder. *Journal of Clinical Child & Adolescent Psychology, 50*, 517–524. https://doi.org/10.1080/15374416.2020.1833735

Klintwall, L., Eldevik, S., & Eikeseth, S. (2015). Narrowing the gap: Effects of intervention on developmental trajectories in autism. *Autism: the International Journal of Research and Practice, 19*, 53–63. https://doi.org/10.1177/1362361313510067

Koole, S., & Tschacher, W. (2016). Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance. *Frontiers in Psychology, 7*, 862. https://doi.org/10.3389/fpsyg.2016.00862

Laister, D., Stammler, M., Vivanti, G., & Holzinger, D. (2021). Social-communicative gestures at baseline predict verbal and nonverbal gains for children with autism receiving the Early Start Denver Model. *Autism, 25*, 1640–1652. https://doi.org/10.1177/1362361321999905

Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why synchrony matters during mother-child interactions: A systematic review. *PLoS ONE, 9*, Article e113571. https://doi.org/10.1371/journal.pone.0113571

Li, A., & Barber, R. F. (2018). Multiple testing with the structure-adaptive benjamini–hochberg algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 81*, 45–74. https://doi.org/10.1111/rssb.12298

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation. (schedule-second edition (ados-2 ed.)* (p. 2012). Los Angeles: Western Psychological Services.

Luiz, D., Barnard, A., Knoesen, N., Kotras, N., Horrocks, S., McAlinden, P. and O'Connell, R. (2006). Griffiths Mental Development Scales—Extended Revised: Two to Eight. Years: Administration Manual. Hogrefe.

MacDuffie, K. E., Estes, A. M., Harrington, L. T., Peay, H. L., Piven, J., Pruett, J. R., Wolff, J. J., & Wilfond, B. S. (2021). Presymptomatic detection and intervention for autism spectrum disorder. *Pediatrics*, 147. https://doi.org/10.1542/peds.2020-032250

Mayo, O., & Gordon, I. (2020). In and out of synchrony—Behavioral and physiological dynamics of dyadic interpersonal coordination. *Psychophysiology, 57*, 13574.

McDaniel, J., Brady, N., & Warren, S. (2021). Effectiveness of responsivity intervention strategies on prelinguistic and language outcomes for children with autism spectrum disorder: A systematic review and meta-analysis of group and single case studies. *Journal of Autism and Developmental Disorders*, 1–34. https://doi.org/10.1007/s10803-021-05331-y

McDonald, N., Senturk, D., Scheffler, A., Brian, J., Carver, L., Charman, T., Chawarska, K., Curtin, S., Hertz-Picciotto, I., Jones, E., Klin, A., Landa, R., Messinger, D., Ozonoff, S., Stone, W., Tager-Flusberg, H., Webb, S., Young, G., Zwaigenbaum, L., & Jeste, S. (2020). Developmental Trajectories of Infants With Multiplex Family Risk for Autism: A Baby Siblings Research Consortium Study. *JAMA Neurology, 77*, 73–81. https://doi.org/10.1001/jamaneurol.2019.3341

O'Keeffe, C., & McNally, S. (2021). A systematic review of play-based interventions targeting the social communication skills of children with autism spectrum disorder in educational contexts. *Review Journal of Autism and Developmental Disorders*, 1–31. https://doi.org/10.1007/s40489-021-00286-3

Pijl, M., Rommelse, N., Hendriks, M., Korte, M., Buitelaar, J., & Oosterling, I. (2018). Does the brief observation of social communication change help moving forward in measuring change in early autism intervention studies? *Autism: the International Journal of Research and Practice, 22*, 216–226. https://doi.org/10.1177/1362361316669235

Piven, J., Elison, J., & Zylka, M. (2017). Toward a conceptual framework for early brain and behavior development in autism. *Molecular Psychiatry, 22*, 1385–1394. https://doi.org/10.1038/mp.2017.131

Provenzi, L., diMinico, G. S., Giusti, L., Guida, E., & Müller, M. (2018). Disentangling the dyadic dance: Theoretical, methodological and outcomes systematic review of mother-infant dyadic processes. *Frontiers in Psychology*, 9. https://doi.org/10.3389/fpsyg.2018.00348

Reyes, N., Pickard, K., & Reaven, J. (2019). Emotion regulation: A treatment target for autism spectrum disorder. *Bulletin of the Menninger Clinic, 83*, 205–234. https://doi.org/10.1521/bumc.2019.83.3.205

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, K., Feldman, J., Crank, J., Albarran, S., Raj, S., Mahbub, P., & Woynaroski, T. (2020). Project AIM: Autism intervention meta-analysis for studies of young children. *Psychological Bulletin, 146*, 1–29. https://doi.org/10.1037/bul0000215

Stavropoulos, K. (2017). Using neuroscience as an outcome measure for behavioral interventions in Autism spectrum disorders (ASD): A review. *Research in Autism Spectrum Disorders, 35*, 62–73. https://doi.org/10.1016/j.rasd.2017.01.001

Tiede, G., & Walton, K. (2019). Meta-analysis of naturalistic developmental. *Behavioral interventions for Young Children with Autism Spectrum Disorder Autism, 23*, 2080–2095. https://doi.org/10.1177/1362361319836371

Towle, P., Patrick, P., Ridgard, T., Pham, S., & Marrus, J. (2020). Is earlier better? the relationship between age when starting early intervention and outcomes for children with autism spectrum disorder: A selective review. *Autism Research and Treatment*. , Article 7605876. https://doi.org/10.1155/2020/7605876

Vandenbroucke, J. P. (2008). Observational research, randomised trials, and two views of medical science. *PLOS Medicine, 5*, Article e67. https://doi.org/10.1371/journal.pmed.0050067

Vivanti, G., Bottema-Beutel, K., & Turner-Brown, L. (2020). *Clinical Guide to Early Interventions for Children with Autism* (Eds.). Berlin/Heidelberg, Germany: Springer, 10.1007/978-3-030-41160-2.

Vivanti, G., Kasari, C., Green, J., Mandell, D., Maye, M., & Hudry, K. (2018). Implementing and evaluating early intervention for children with Autism: Where are the gaps and what should we do? *Autism Res, 11*, 16–23. https://doi.org/10.1002/aur.1900

Vivanti, G., & Zhong, H. (2020). Naturalistic Developmental Behavioral Interventions for Children with Autism. In G. Vivanti, K. Bottema-Beutel, & L. Turner-Brown (Eds.), *Clinical Guide to Early Interventions for Children with Autism* (pp. 93–130). Cham: Springer International Publishing, 10.1007/978-3-030-41160-2_6.

Wetherby, A., Woods, J., Guthrie, W., Delehanty, A., Brown, J., Morgan, L., Holland, R., Schatschneider, C., & Lord, C. (2018). Changing developmental trajectories of toddlers with Autism spectrum disorder: strategies for bridging research to community practice. *Journal of Speech, Language, and hearing research: JSLHR, 61*, 2615–2628. https://doi.org/10.1044/2018_JSLHR-L-RSAUT-18-0028

Whitehouse, A., Varcin, K., Pillar, S., Billingham, W., Alvares, G., Barbaro, J., Bent, C., Blenkley, D., Boutrus, M., Chee, A., Chetcuti, L., Clark, A., Davidson, E., Dimov, S., Dissanayake, C., Doyle, J., Grant, M., Green, C., Harrap, M., Iacono, T., & Hudry, K. (2021). Effect of preemptive intervention on developmental outcomes among infants showing early signs of Autism: A randomized clinical trial of outcomes to diagnosis. *JAMA Pediatrics, 175*, Article 213298. https://doi.org/10.1001/jamapediatrics.2021.3298

Yoder, P. J., Bottema-Beutel, K., Woynaroski, T., Chandrasekhar, R., & Sandbank, M. (2013). Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders. *Evidence-Based Communication Assessment and Intervention, 7*, 150–174. https://doi.org/10.1080/17489539.2014.917780

Zachor, D., & Ben-Itzchak, E. (2017). Variables affecting outcome of early intervention in autism spectrum disorder. *Journal of Pediatric Neurology, 15*, 129–133. https://doi.org/10.1055/s-0037-1601444

Zitter, A., Rinn, H., Szapuova, Z., Avila-Pons, V. M., Coulter, K. L., Stahmer, A. C., Robins, D. L. and Vivanti, G. (2021). Does treatment fidelity of the early start denver model impact skill acquisition in young children with autism? Journal of autism and developmental.10.1007/s10803–021-05371–4.