Temporal Dynamics of Child-Adult Conversations Across Social Contexts and Tasks: A Comparative Study of Autistic and Typically Developing Children

Christopher Cox*[1], Riccardo Fusaroli*[1,2], Roberta Rocca[1], Arndis Simonsen[1,3], Sunghye Cho[4], Maggie Rose Pelella[4], Mark Liberman[4], Christopher Cieri[2], Azia Knox[6], Aili Hauptmann[7], Alison Russell[8], Alison Hulick[7], Kimberly Tena[7], Monica Madya[7], Hannah Adeniyi[5], Grace Hicks[5], James Fiumara[2], Kevin Walker[2], Ani Nenkova[8], Jennifer Uzokwe[7], Christopher Chatham[9], Judith Miller[7], Juhi Pandey[7], Robert T. Schultz[7], Julia Parish-Morris[2,7]

[1]Interacting Minds Center and Department of Linguistics and Cognitive Science, Aarhus University
[2]Linguistic Data Consortium, University of Pennsylvania
[3]Psychosis Research Unit, Aarhus University Hospital
[4]Department of Linguistics, University of Pennsylvania
[5]Department of Psychology, University of Pennsylvania
[6]Department of Linguistics and Cognitive Science, University of Delaware
[7]Center for Autism Research, Children's Hospital of Philadelphia
[8]Ohio University Hearing, Speech, and Language Clinic
[9]Computer and Information Science, University of Pennsylvania
[10]Hoffmann-La Roche Ltd

2024

Authors marked with * share joint first-authorship. Correspondence to: Christopher Cox
(chris.mm.cox@gmail.com)

Abstract

Verbal interactions involve fast inter-turn response latencies and require reciprocal behaviours that depend on an elaborate set of socio-cognitive skills. In this study, we investigated how the temporal dynamics of child-adult interactions changed according to the experimental activity, the familiarity of the interlocutor, individual socio-cognitive differences among children, information flow, and turn-by-turn dynamics. The procedure involved 28 autistic children (16 male; mean age=10.8 ± 3.2 years) and 20 age-matched typically developing children (8 male; mean age=9.6 ± 3 years). Each child participated in 7 face-to-face, task-orientated conversations with their caregivers (330 sessions) as well as 7 more affiliative telephone conversations alternately with their caregivers (144 sessions, 60 typical development sessions) and a previously unfamiliar experimenter (192 sessions, 112 autism sessions). By modelling inter-turn response latencies in multi-level Bayesian location-scale models, we found that autistic children produced faster response latencies, took shorter pauses and exhibited more overlaps than than typically developing children, especially when conversing with the experimenter. The temporal dynamics of turn-taking changed according to task demands and social context and were influenced by individual differences in the socio-cognitive, linguistic and motor skills of children as well as by turn-by-turn flows of information and interpersonal adjustment. We argue that these results call for a re-evaluation of turn-taking as immersed in a multi-dimensional social context that creates different affordances for communicative patterns.

1. Introduction

Vocal exchange is the primary way that humans navigate their social world. The rapid vocal exchanges that characterise conversation are complex feats of coordination. The modal response latencies of interactants in a typical adult conversation tend to be short, approximately 59-200 ms (Dingemanse & Liesenfeld, 2022; Levinson, 2016; Stivers et al., 2009), which is at the absolute limits of human reaction time and too rapid to rely on planning and production after the end of the turn (Magyari et al., 2014; Magyari & De Ruiter, 2012). People must therefore proactively prepare their contingent responses, detect when the likely endpoint of the other interactant's turn is, decide when to deliver their response, and predict how the other interactant will react in real time (Dale et al., 2013; Magyari & De Ruiter, 2012; Magyari et al., 2014). While adults draw upon their linguistic, interpersonal and world knowledge to plan and predict with such efficiency (Ford et al., 1996; Magyari & De Ruiter, 2012; Magyari et al., 2014), it remains an open question how children balance this complex juggling act and which skills and components are involved. In this paper, we investigate how turn-taking dynamics are influenced by the social context of the interaction, individual differences in socio-cognitive skills, turn-by-turn information flows as well as moment-to-moment dynamics of interpersonal exchanges in both typical and atypical development.

The fundamental timing structures of turn-taking emerge early in ontogeny (Hilbrink et al., 2015; Levinson, 2016), have deep evolutionary roots (Pika et al., 2018; Ravignani & de Reus, 2019; Levinson, 2006, 2016; Verga et al., 2023) and exhibit similar structures across languages (Stivers et al., 2009; Dingemanse & Liesenfeld, 2022). Turn-taking has been argued to be part of - and party derivable from - a package of fundamental capacities that enable human social behaviour (Levinson, 2006, 2016). Accordingly, turn-taking has important social implications. Shorter inter-turn latencies facilitate efficient information exchange (Fusaroli et al., 2014; Templeton et al., 2022, 2023) and foster a sense of connection and shared motivation (C. Dideriksen, Christiansen, Tylén, et al., 2023; Fusaroli et al., 2014, 2019; Fusaroli & Tylén, 2016; Konvalinka & Roepstorff, 2012; Tylén et al., 2013). Adults in verbal interactions with shorter inter-turn response latencies, for example, report enjoying their conversations more (Templeton et al., 2022), while longer response latencies in conversations between strangers – but not friends – produce awkwardness and shifts in conversational topics (Templeton et al., 2023). Because the central timing structures of turn-taking – that is, avoiding excessive overlaps and minimizing latency between turns – is such a foundational precursor for social interactions, it becomes crucial to understand how these interactional skills are grounded in a network of other individual differences – especially in individuals with social atypicalities like autism.

The inter-turn response latencies of autistic children have frequently been described as delayed in the literature (Nguyen et al., 2022), and many empirical studies corroborate this finding across different ages, languages and interactive contexts (W. Fay & Schuler, 1980; Heeman et al., 2010; Ochi et al., 2019; Warlaumont et al., 2014). However, a growing number of studies finds the opposite pattern of results – faster turn-taking in autistic 2.5-4.5 year-old children in spontaneous interactions compared to typically developing controls (Fusaroli et al., 2024), no difference between autistic 10-year-old girls and typically developing children (Cho et al., 2023), and no clear difference between autistic adults and controls (Wehrle et al., 2023a). Qualitative studies also show that autistic children generally do not display difficulties in conversational turn-taking (Ochs et al., 2004; Bottema-Beutel et al., 2018). This is possibly due to caregivers learning to produce scaffolding sequences to enable precisely contingent responses (Bottema-Beutel et al., 2018; Bottema-Beutel et al., 2022) – just as caregivers do in interactions with typically developing children who also exhibit frequent delays and non-contingent responses (Dunn & Shatz, 1989; Ervin-Tripp, 1982). Interestingly, limited generalizability is a common phenomenon when investigating other acoustic dimensions of verbal interactions – for example, speech rate, pitch, and turn length in atypical populations (Thurber & Tager-Flusberg, 1993; Fosnot & Jun, 1999; Nadig & Shaw, 2012,?; Bone et al., 2015; McCann & Peppé, 2003; Fusaroli et al., 2017, 2022; Rybner et al., 2022).

The limited generalizability of findings on child-adult interactions in atypical populations are likely caused by both between-study heterogeneity in methodological choices and heterogeneity across contexts of the phenomena studied (Yarkoni, 2022). As to the former, most studies of child-adult turn-taking finding slower responses in autism exclude overlaps (i.e., negative response latencies) (Nguyen et al., 2022). Besides ignoring between 25% and 50% of conversational turns (Dingemanse & Liesenfeld, 2022; Ervin-Tripp, 1979; Garvey & Berninger, 1981), the choice of excluding overlaps is also non-trivial in terms of conceptual and statistical assumptions introduced in the analysis and their consequences (for a full discussion, see methods section and Supplementary Materials). Indeed, the two studies explicitly including overlaps found that response latencies for the autism group were equivalent or faster than those of the typical development group (Fusaroli et al., 2024; Wehrle et al., 2023b), contrary to those excluding overlaps. Analogously, while turn-taking involves a non-trivial amount of long pauses skewing the distribution of latencies on the right (long tail), there is no well motivated consensus on how to deal with them - from ignoring them, to excluding them as outlier at varying thresholds, to explicitly accounting for long tails in the model.

Different treatment will very likely result in different findings even on the same data. Further, the limited generalizability of turn-taking findings may also be due to response latencies systematically varying as a function of context and subpopulation (contextual and population heterogeneity). For instance, the degree of familiarity of the interlocutor (Doherty-Sneddon et al., 2013), the specific experimental task (Casillas & Frank, 2017; Forgeot d'Arc et al., 2020), and the socio-demographic and clinical characteristics of the sample (Fusaroli et al., 2024; Cho et al., 2023) may all critically affect turn-taking (Nguyen et al., 2022). Response latencies in this sense may not be a one-size-fits-all individual trait. Rather, different social contexts might result in different patterns of turn-taking behaviors, and differences between autistic and typically developing children might only be apparent in specific contexts. Crucially, this can be highly informative of the mechanisms at play.

To improve our understanding of how turn-taking dynamics systematically change across these sources of heterogeneity and extract insights from new dimensions in the methodological landscape, the current study leverages a large corpus of child-adult interactions with autistic and typically developing children. This large corpus of child-adult interactions allows us to evaluate the generalizability of empirical results across multiple social contexts and experimental sessions and to assess the extent to which overlaps and pauses play a role in turn-taking dynamics. The following section articulates important components in the temporal dynamics of turn-taking and motivates how we tackle them in our research questions.

1.1. Components of Turn-Taking

Response latencies in verbal conversations are often conceived as the expression of an individual trait, an individual specific tempo in responding and as such the result of individual differences in cognitive profiles and developmental trajectories (Fusaroli et al., 2024; Donnelly & Kidd, 2021). However, conversations are a deeply contextual phenomenon: a relaxed chat and an oral questioning do not have the same tempo and different individuals might react differently to different contexts. Not least, response latencies vary on a turn-by-turn basis, as a function of what is said and of the tempo the interlocutor is adopting (Warlaumont et al., 2014; Tylén et al., 2013; Fusaroli et al., 2014, 2019). We construe the temporal dynamics of turn-taking as a window into this complex network of likely interacting factors. We argue that a nuanced perspective on verbal interactions requires consideration of the following five components: i) conversational task, ii) interlocutor familiarity, iii) socio-cognitive, linguistic and motor skills, iv) turn-by-turn flow of information, and v) turn-by-turn dynamics. Each of these components is motivated in the following paragraphs.

1.1.1 Conversational Task and Interlocutor Familiarity

Verbal interactions do not happen in a vacuum, but are an intrinsic part of joint activities(Fusaroli et al., 2014; Fusaroli & Tylén, 2012). Different activities involve affordances and constraints for the conversations supporting them (Olsen & Tylén, 2023; Parish-Morris et al., 2019; Cola et al., 2022; Dale, 2015). For instance, a joint task that requires children and parents to tell each other where to place stickers on a blank picture may demand a focus on precise information transfer as well as require interlocutors to allow each other sufficient time to place the stickers before responding. In contrast, the situational demands of a spontaneous conversation about holidays would not require as much precision and likely require fewer long pauses to ensure efficient communicative patterns C. Dideriksen, Christiansen, Tylén, et al. (2023); C. Dideriksen, Christiansen, Dingemanse, et al. (2023); C. R. Dideriksen et al. (2019).

Equally important is who we engage in joint activities. Chatting with a long-known friend is very different from chatting with somebody just met at a party, and discussing which toy to bring at a show-and-tell is very different if doing so with a peer or a with a new teacher. Indeed, adults interacting with friends take longer to respond compared to when interacting with strangers, and their occasional long pauses are perceived more positively (Templeton et al., 2022, 2023). In child-adult interactions, the familiarity of the interlocutor has likewise been shown to change how both autistic and typically developing children interact (Doherty-Sneddon et al., 2013; Dawson et al., 2002; Forgeot d'Arc et al., 2020).

Turn-taking patterns in child-adult conversations, then, represent an interaction between the structural demands of the task and the longer history of interaction with an interlocutor (their familiarity). The extent to which the temporal dynamics of verbal interactions is invariant across these dimensions is an open question. Accordingly, the study design of the current paper allowed us to evaluate how different interactional affordances emerging from the task and familiarity of the interlocutor produce differences between the autism and typical development group in their turn-taking behaviours (cf., Research Question 1 & 2 below). Crucially, each of the contexts was measured in multiple weekly sessions, to disentangle heterogeneity due to session by session variation from that due to systematic contextual variation.

1.1.2 Individual Socio-Cognitive Differences

The challenges posed by various contextual demands and interlocutors require a complex network of skills to be tackled. For instance, promptly responding requires the ability to integrate semantic, syntactic, and pragmatic information to anticipate an imminent termination of a turn (Magyari & De Ruiter, 2012), that is, linguistic skills. Smooth turn taking also requires navigating the social world: inferring others' intentions, anticipating what they might do next, planning behaviors so that the other can read our intentions, etc. Heesen & Fröhlich (2022); Levinson (2006); Waade et al. (2023). Indeed, the interaction engine hypothesis argues that such core human socio-cognitive skills underlie our ability to interact, including take-turns in effective ways Levinson (2016). Motor development also changes the nature of children's interactions with the world around them, making them better at promptly orienting towards information sources, and promptly anticipating and reacting to them. This has crucial downstream effects on their own communicative abilities Iverson (2010); West & Iverson (2021), as well as on the contingent responses from caregivers Karasik et al. (2014); Tamis-LeMonda & Adolph (2013). Individual differences among children can thus alter how the children themselves interact with caregivers but also change the input that children receive from caregivers, as exemplified by studies showing that adults adapt to the demands of their addressees – whether they be non-native speakers (Uther et al., 2007; Piazza et al., 2022), children (Cox et al., 2023; Hilton et al., 2022), children with autism (Neimy et al., 2017; Warlaumont et al., 2014; Leezenbaum et al., 2014), infants with cochlear implants Dilley et al. (2020); Kondaurova et al. (2013), or pets (Jardat et al., 2022; Panneton et al., 2023).

Crucially, different contextual and interlocutor demands might engage these skills to a higher or lower degree. For instance, familiar interlocutors might be easier to anticipate and respond to, while the ability to promptly take turns with unfamiliar ones might be more highly dependent on one's socio-cognitive skills Bigelow (1999); Jaffe et al. (2001). In other words, assessing the generalizability and variation of how individual differences relate to response latencies across contexts is essential to understand their role.

Accordingly, in this study, we incorporated rich measures of individual differences to examine how socio-cognitive, linguistic and motor differences among children might change the temporal dynamics of interactions across social contexts with different demands (cf., Research Question 3 below).

1.1.3 Information Flow

Conversational demands vary on a turn-by-turn basis in the amount of information shared, the effort required to parse what has been said, and the effort required to plan the next utterance. Indeed, adult inter-turn response latencies depend on the predictability of the content discussed (Bögels et al., 2015), and preschool children produce shorter inter-turn response latencies with more predictable ends to questions (Lindsay et al., 2019; Casillas & Frank, 2017). Intuitively, children having to respond to a complex, open-ended question (e.g., "What are your thoughts on the story we just read?") will likely require longer response times compared to a predictable, closed-ended question (e.g., "How many apples are there?").

In this study, we attempt to capture the interplay between turn-by-turn information flow and temporal dynamics across different tasks and social contexts. In particular, we quantify the effort required to process an utterance in terms of how predictable it is from previous turns (the more predictable, the less effort needed). Analogously, we quantify the effort needed to plan an utterance in terms of how predictable the utterance to be produced next by the listener compared to previous turns. We then assess whether these measures have reliable associations with inter-turn response latencies (cf., Research Question 4 below).

1.1.4 Turn-by-Turn Dynamics

Beyond pure exchanges of information, conversations are also a dance in which interlocutors might come to share a common tempo of well-timed exchanges Wilson & Wilson (2005); Fusaroli et al. (2014). Indeed, there is evidence of strong correlations between between child and caregiver response latencies (Nguyen et al., 2022). However, these findings are based only on summary estimates (the average response latency across the whole conversation) and cannot say much about the dynamics of co-adaptation on a turn-by-turn basis, nor about how different contextual and interlocutor demands, or individual differences, might influence such co-adaptation (Fusaroli et al., 2024). Further, vocal interactions likely involve dynamic ebb-and-flow waves of mutual convergence and scaffolding that takes place during local transactions (Ritwika et al., 2020; Warlaumont et al., 2014; Pouw & Holler, 2022), but little is known about how this unfolds in child-adult interactions.

Accordingly, in our study we include an analysis of these temporal dependencies: how child and caregiver adjust to each other's - and their own - previous latencies on a turn-by-turn basis across contexts, as well as how temporally extended these co-adjustments are. By investigating this split-second world of local interactions, we can better approach the interpersonal and mutually negotiated dynamics underlying turn-taking in child-adult interactions (cf., Research Question 5 below).

1.2. Research Questions

- **Research Question 1**: Are response latencies an individual trait, broadly invariant across measurements, or does the social context (conversational task and familiarity of the interlocutor) matter?
- **Research Question 2**: Does autism involve - on average - slower response latencies - once explicitly accounting for long pauses and overlaps? How is the finding affected by context?
- **Research Question 3**: Are differences in turn-taking behaviour grounded in individual differences (motor skills, linguistic skills, social cognition, social awareness and social motivation)? How are the findings affected by context?
- **Research Question 4**: Is turn-taking behaviour grounded in the turn-by-turn information flows (predictability of previous and current utterance)? How is the finding affected by context?
- **Research Question 5**: Is turn-taking behaviour grounded in the turn-by-turn temporal dynamics of the interaction? How is the finding affected by context?

By directly addressing these different components of turn-taking, we can begin to disentangle their complex interplay and provide a more nuanced understanding of individual and group differences in turn-taking dynamics. For instance, children with higher than average social skills might be better at detecting and integrating information from social cues when interacting with unfamiliar interlocutors – and this may also determine the contexts for which we would expect a difference in response latencies between autistic and typically developing children. We thus aim to provide a nuanced construal of how interacting individuals co-construct and co-regulate turn-taking dynamics that emerge from complex interactions among these components.

2. Methods

2.1. Participants

The study involved 28 autistic children (10 female) and 20 typically developing children (12 female) children, matched at the group level on age, IQ and self-reported race (see **Table 1**). The autistic children were previously diagnosed with Autistic Disorder or Pervasive Developmental Disorder - Not Otherwise Specified (PDD-NOS) by physicians or psychologists. Participants were recruited via an electronic health record system at a large hospital-based academic medical center in the U.S., flyers and inquiry forms, and internal databases. All participants underwent the "Telephone Screening Interview" developed by the researchers at the Center for Autism Research to ascertain inclusion and exclusion criteria (such as autism diagnostic history, family history, or upcoming changes to intervention/medication). The inclusion criteria were as follows: all participants had i) to be between 6 and 18 years of age, ii) to have English as their first language, iii) to be verbally fluent (consistent with chronological age), iv) to have an IQ score above 75, and v) autistic participants had to have a current Social and Communication Questionnaire (SCQ) score above 10, and typically developing participants had to have an SCQ score of below 11, where higher scores represent a higher degree of social communication impairment (Berument et al., 1999; Chesnut et al., 2017). The exclusion criteria were as follows: i) known genetic or neurological condition that impacts neurodevelopment or vocal production/language, ii) a history of expressive language deficits or severe neurological injury likely to affect expressive language and communication behavior, iii) extreme prematurity (<32 weeks), iv) diagnosis of hearing impairment or cochlear implant, v) plan to begin or change medication or intervention during study duration, and vi) typically developing participants could not have first-degree family members with autism.

Table 1: *: Descriptive statistics of the population. IQ stands for Wechsler Abbreviated Scales of Intelligence – 2nd Edition (WASI II). SCQ stands for Social and Communication Questionnaire. SRS stands for Social Responsiveness Scales – 2nd Edition. CELF stands for Clinical Evaluation of Language Fundamentals - 5th edition. VAB stands for Vineland Adaptive Behavior - 3rd edition. The five last rows display details about the turns for each of the groups under investigation. Median and Mean Turn Length refers to the average difference in seconds between the start- and endpoint of children's turns across conditions.*

Variable	Autism Group	Typical Development
Sample Size	28 (16 male, 2 non-binary)	20 (8 male)
Age (years)	10.82 (3.14)	9.6 (3.02)
IQ	112.88 (13.26)	118.16 (13.15)
SCQ	16.68 (16.68)	1.2 (1.2)
SRS (Raw)	77.57 (25.99)	12.75 (9.13)
SRS (T-score)	67.77 (9.92)	42.5 (3.44)
Language (CELF)	107.67 (13.15)	109.53 (11.4)
Social Cognition (SRS)	13.71 (5.46)	1.8 (2.21)
Social Motivation (SRS)	12.36 (6.11)	3 (2.13)
Motor (VAB)	82.54 (3.82)	85.25 (2.07)
Social Awareness (SRS)	10.39 (3.51)	2.5 (2.04)
Number of Sessions	190	140
Number of Turns	21733	14628
Average Turns Per Session	27.7	26.3
Median Turn Length (sec)	2.87 (8.22)	2.69 (9.93)
Mean Turn Length (sec)	5.69 (8.22)	5.97 (9.93)

2.2. Procedure

2.2.1 Data Collection

Participant data was collected over the course of several telephone sessions. After an information call and having provided consent, the participants and their families underwent a screening questionnaire and the Social and Communication Questionnaire (Berument et al., 1999; Chesnut et al., 2017). Subsequent assessment relied on the following tools:

- Wechsler Abbreviated Scale of Intelligence (WASI-II) (Wechsler, 2008; McCrimmon & Smith, 2013)
- Clinical Evaluation of Language Fundamentals (CELF), 5th edition (Wiig et al., 2013; Denman et al., 2017)
- Vineland Adaptive Behavior Scales (VABS), 3rd edition - the motor scale in particular (S. et al., 2016; Farmer et al., 2020)
- Social Responsiveness Scale (SRS), 2nd edition - the social cognition, social motivation and social awareness scales in particular (Constantino, 2012; Bruni, 2014).

After evaluation according to the above assessment tools, participants underwent phone-based recording sessions – approximately once per week over a period of seven weeks. During the first session with the child, staff remained on the line to facilitate tasks with the parent and child, while in following sessions children and parents completed the tasks on their own. The participants participated in two speech recording tasks: i) a matching game and ii) prompted conversations. In the matching game, the child and parent each received a sticker sheet and a blank picture and were instructed to take turns in telling the other person where to place the stickers. The aim of the game for the child and parent was to have the pictures look the same, despite them not being able to see each other's sticker sheet. In the prompted conversations, children were given guided prompts (e.g. planning a birthday party, or food preferences) and interacted alternately with their caregivers in even weeks and with an initially unfamiliar experimenter in odd weeks. These two conversational tasks were inspired by traditionally used task-oriented (such as the referential game map task) and conversational tasks (C. Dideriksen, Christiansen, Tylén, et al., 2023; N. Fay et al., 2018), but adjusted for repeated engaging conversational interactions within the context of this study. The total number of sessions and turns for the respective groups can be viewed in **Table 1**.

2.2.2 Recording and Transcribing Turns

To record participants' speech we relied on an automated phone bank collection system - supporting single and dual speaker modes - originally developed by the University of Pennsylvania Linguistic Data Consortium and adapted for use with autistic participants. The speech recordings were transcribed by trained annotators using a web-based transcription tool with a built-in speech activity detector function. The first pass of the data involved annotators assessing and correcting the timecodes for the automatically identified speech segment boundaries and at the same time transcribing the audio files. Conversational turns were identified as sequences of speech by one interlocutor without interruption from the other, thus generating an ABABAB structure. Fully overlapping speech between interlocutors occurred occasionally (e.g., during some backchannels). In order to preserve the sequential nature of the transcript, such cases were interleaved within the current speaker's utterance at the first pause longer than 1 second, while leaving the timecodes unchanged. The impact of different definitions of conversational turns has been assessed and the findings are robust to changes in definition (Fusaroli et al., 2023a, 2023b). We ensured that there would be 20% double-coded (or "dual") transcripts to permit calculation of inter-rater reliability. Once the first pass was completed, the data annotators began their second pass through the transcripts for quality control. None of the transcribers controlled the quality their own work, and all transcribers/data annotators were blind to the diagnostic status of the participant.

2.2.3 Measuring turn-taking

Inter-turn response latencies were calculated by subtracting the onset of an utterance from the offset of the previous utterance. This yielded both negative (overlaps) and positive response latencies (gaps), including a non trivial amount of long response latencies (>1s). In previous research, overlaps have often been excluded as failures of communication, or as qualitatively different from gaps; and long pauses have either been included but ignored (i.e., using likelihood functions without long right tails), or excluded as outliers at various thresholds changing from study to study. In this study we choose to include both overlaps and long pauses and to explicitly account for them in the statistical modeling. From a statistical perspective, removing negative latencies from the analysis arbitrarily truncates an otherwise continuous distribution and this violates assumptions of the models commonly used in the literature (e.g. Gaussian likelihoods with normal residuals, see also Heldner et al 2010), introducing statistical bias on measures of central tendency and variance. Further, a hard threshold at 0ms is unnecessarily arbitrary. An overlap of 10ms and a gap of 10ms arguably require the same precise anticipation and shared rhythm, and the same for 20 and 30ms and so on. It is therefore unclear where a binary threshold should be placed to capture qualitatively different phenomena, even if we accepted there are qualitatively different phenomena. Third, as mentioned above overlaps constitute a substantial portion of conversations: between 25% and 50% of conversational utterances are estimated to start before the previous interlocutor is done with speaking Dingemanse & Liesenfeld (2022), and this is argued to be even more frequent in conversations involving children Ervin-Tripp (1979); Garvey & Berninger (1981). Excluding overlaps effectively eliminates substantial portions of a conversation, without qualifying whether they actually are qualitatively different, nor considering the biases that this exclusion might introduce. Just like gaps can mean many different things – from thoughtful engagement of the interlocutor to a desire to avoid a topic, and from cognitive difficulties to comfort in being together Templeton et al. (2023); Templeton & Wheatley (2023), so do overlaps, and their functions are often in continuity with those of gaps. Indeed, overlaps play complex roles in conversation, far from being "just an interruption". Overlaps might indicate shared attention and mutual understanding, as in backchannels, or completion of the interlocutor's sentence; or it might signal disagreement and the need for potential repair C. Dideriksen, Christiansen, Tylén, et al. (2023). Overlaps can indeed play a crucial role in building and maintaining rapport Bryant et al. (2016); Bryant

(2020); Templeton et al. (2022); Cummins (2019) and it has likely many other underexplored productive functions. To assuage concerns of comparability with previous studies, we will also conduct analyses excluding overlaps and discuss differences in findings. Analogously, removing long pauses introduces artificial and somewhat arbitrary dichotomization of response latencies without a closer look at discontinuity in function. Accordingly, we will employ statistical models accounting for long right tails in the latency distribution.

2.2.4 Statistical Models

The distributional properties of response latencies in turn-taking produce unique modelling challenges (see **Figure 1**). Turn-taking distributions involve negative latencies from overlapping speech, a large centre of mass just above zero, and positive skew from occasional long pauses between turns. To accommodate these distributional properties in our models, we built Bayesian multi-level models with ex-Gaussian likelihoods across all research questions. This form of likelihood combines a Gaussian distribution with an exponential distribution, allowing for the flexibility to represent the central tendency together with the positive skew often observed in empirical turn-taking data. Specifically, in our Bayesian framework, the ex-Gaussian likelihood was parameterized by three parameters: the mean (μ) and variance (σ) of the Gaussian component, as well as the rate (β) of the exponential tail.

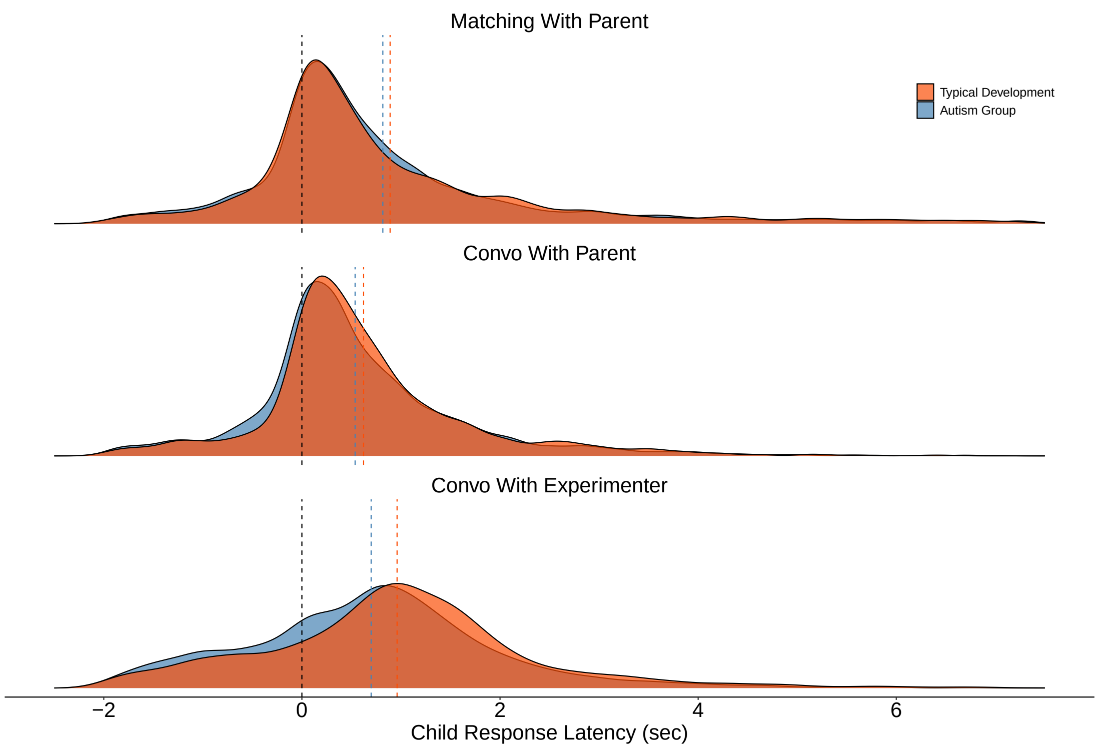

Figure 1: *Panel of plots to demonstrate the distributional properties of child response latencies across different combinations of contexts and familiarity. The black dashed line marks a response latency of zero, the orange dashed line the mean in the typical development group, and the blue dashed line the mean of the autism group. As can be gleaned from the plots, response latencies involve both negative response latencies (overlapping speech), a large centre of mass just above zero, as well as positive skew (long pauses between turns)*

To investigate research questions 1 and 2, we modeled each of the parameters — the mean (μ), and variance (σ) of the Gaussian component, as well as the rate (β) of the exponential tail — as conditioned on diagnostic group and conversational context (matching game vs. conversations with parents vs. conversations with experimenters). We allowed the estimates to vary within each combination of these factors – similar to a full interaction model – to assess the generalizability of effects across the tasks and familiarity of the interlocutor. We included varying slopes in the model to allow for the estimates across conversational contexts to vary within individual children and as a function of visit. Variance was partially pooled by child within their respective diagnostic group to accommodate the potential for between-group differences in heterogeneity in mean and variance. To investigate the test-retest reliability of our measures of turn-taking, we extracted estimates of mu, sigma and beta (mean and variance of the Gaussian component, exponential tail) by visit and diagnostic group. We then calculated a correlation matrix of the estimates within each diagnostic group across the 7 sessions and assessed both average correlation and the impact of distance on the correlation estimate. More information about the specification of our models, including priors, model checks and control analyses can be found in **Model Information and Quality Checks** in Supplementary Materials.

To investigate the role of individual differences in socio-cognitive, linguistic and motor development (Research Question 3) we added scaled scores of SRS (Social Cognition, Social Motivation, Social Awareness), CELF, and the Vineland Motor Scale to the first model. Their relation to response latencies was allowed to vary by diagnostic group, conversational context and familiarity (equivalent to a full interaction model), so we could assess the generalizability of the findings. We chose to use SRS scores of socio-cognitive skills instead of Vineland Socialization scores as in the study by Fusaroli et al. (2024), since individual abilities were measured only once during data collection. Vineland scores have been shown to be highly sensitive to small variations between weekly sessions (Chatham et al., 2018) and would therefore introduce noise in the association between underlying social skills measured at one point and response latencies measured over several sessions.

To investigate the role of linguistic predictability (Research Question 4) we calculated cosine similarity between utterance-level transformer-based embeddings of successive turns extracted using the all-mpnet-base-v2 model from the Sentence Transformers library (Reimers & Gurevych, 2020). The all-mpnet-base-v2 model is based on the Transformer architecture, specifically using the MPNet (Masked and Permuted Network) variant (Song et al., n.d.). MPNet is an advanced transformer model that combines the benefits of BERT (Bidirectional Encoder Representations from Transformers) and XLNet (a generalized autoregressive pre-training method). The embeddings generated by all-mpnet-base-v2 can be used to calculate cosine similarity between sentences. Cosine similarity measures the cosine of the angle between two vectors in a multi-dimensional space, providing a value between -1 and 1. Higher values indicate greater similarity. In this context, these cosine similarity scores between successive turns served as a measure of linguistic predictability; by examining the similarity between consecutive turns in a conversation, we can quantify how predictable one speaker's response is based on the previous turn, and then model the effect of predictability on response latencies. We incorporated cosine similarity of both the adult and child utterances as scaled fixed effects, conditioned on conversational context as in the above models, and allowed both the effect of predictability of the adult and child utterance to vary by child and as a function of visit.

To investigate the role of interpersonal adjustment (Research question 5), we built a multivariate outcome model with the response latencies of the child speaker being conditioned on their own previous response latency as well as the response latency of their adult interlocutor, and vice versa with the adult speakers. Varying effects in the two equations were modeled as correlated between children and their interlocutors. This model is thus similar to structural equation modeling approaches such as the actor-partner interdependence (Cook and Kenny, 2005) and the dyadic coupling (Helm et al., 2012) models.

To evaluate the impact of overlapping speech on response latency patterns, we conducted two control analyses. First, we re-ran the models excluding overlaps (i.e., negative latencies) to determine how the inclusion of overlaps in the main analyses potentially influenced our results. Except for specifying a truncated prior indicating no value could be below 0 to facilitate the sampling process, these control models and their specifications were identical to the primary models. Second, to assess the extent of overlap in each group for each task, we ran a multi-level logistic regression model, with the rate of overlap conditioned on diagnostic group, conversational context, and familiarity of the interlocutor, using the same varying effects structure as in the primary model.

As an overall control analysis, we created surrogate dyads to emulate the baseline turn-taking that would happen by chance if there were no contingency between adult and child responses (cf., **Table 20** in Supplementary Materials). The surrogate pairs were created by pairing the time codes of the caregiver of a different child within the same diagnostic group and visit with the timecodes of the child. This control analysis shows how our patterns of results cannot be reduced to simple differences in vocalization rates and durations but are due to the specific contingency in the interpersonal temporal sequence of vocalizations. This surrogate analysis can be found in **Model With Surrogate Pairs**.

Full details on model specification, including weakly informative priors, and model quality checks are available in the Supplementary Materials. Estimates from the models are reported as mean and 95% Credible Intervals (CI) of the posterior estimates. We calculated evidence ratios (ER) for our hypotheses in the form of the posterior probability of the directed hypothesis against the posterior probability of all the alternatives; that is, if we expected higher response latencies in autistic children, we would count the posterior samples compatible with this hypothesis and divide them by the number of posterior samples compatible with a null or negative effect. An evidence ratio of 5 thus implies that the evidence in favour of the hypothesis is 5 times greater than the evidence in favour of the alternative model or hypothesis. An evidence ratio of 'Inf' implies that all posterior samples are in the direction of the hypothesis, such that the ratio in favour becomes infinite. Throughout the results section, we compare posterior estimates between those in the experimenter and parent conversations, on the one hand, and between those in parent conversations and the matching game (conducted only with parents), on the other hand, in order to ascertain the effect of the familiarity of the interlocutor and the effect of experimental task on response latencies, respectively.

All code and data for these analyses is provided on OSF (URL).

3. Results

3.1. Context and diagnostic group (RQ1 & RQ2)

3.1.1 Average Response Latencies (mu)

The overarching findings for Research Questions 1 and 2 are shown in **Table 2** and **Figure 2**. Context matters. The children exhibited slower response latencies when conversing with the unfamiliar experimenter than with their parent in both autism (200ms [24, 378], ER = 28.1) and typical development groups (355ms [114, 596], ER = 95.2). The children also produced slower latencies when solving the matching game than when more freely conversing with their parents in both autism (296ms [139, 449], ER = 146.1) and typical development groups (296ms [133, 462], ER = 146.1).

Diagnostic group also matters. Autistic child response latencies were faster than those in the typical development group when aggregating across conditions (-143ms [-282, -3], ER = 20.7), and this effect was the strongest in the conversations with the unfamiliar experimenter (-246 [-501, 0], ER = 19), followed by the matching game (-91 [-250, 72], ER = 4.8, difference from the effect in the conversation with experimenter: -155ms [-460, 151], ER = 4.1) and the conversations with the parent (-91 [-299, 115], ER = 3.5, but this was not reliably different from the matching game 0.06ms [-231, 232], ER = 1).

Test-retest reliability of response latency estimates across the seven sessions was high within each social context, especially for the autism group in conversational tasks (see also **Figure 10** in the Supplementary Materials).

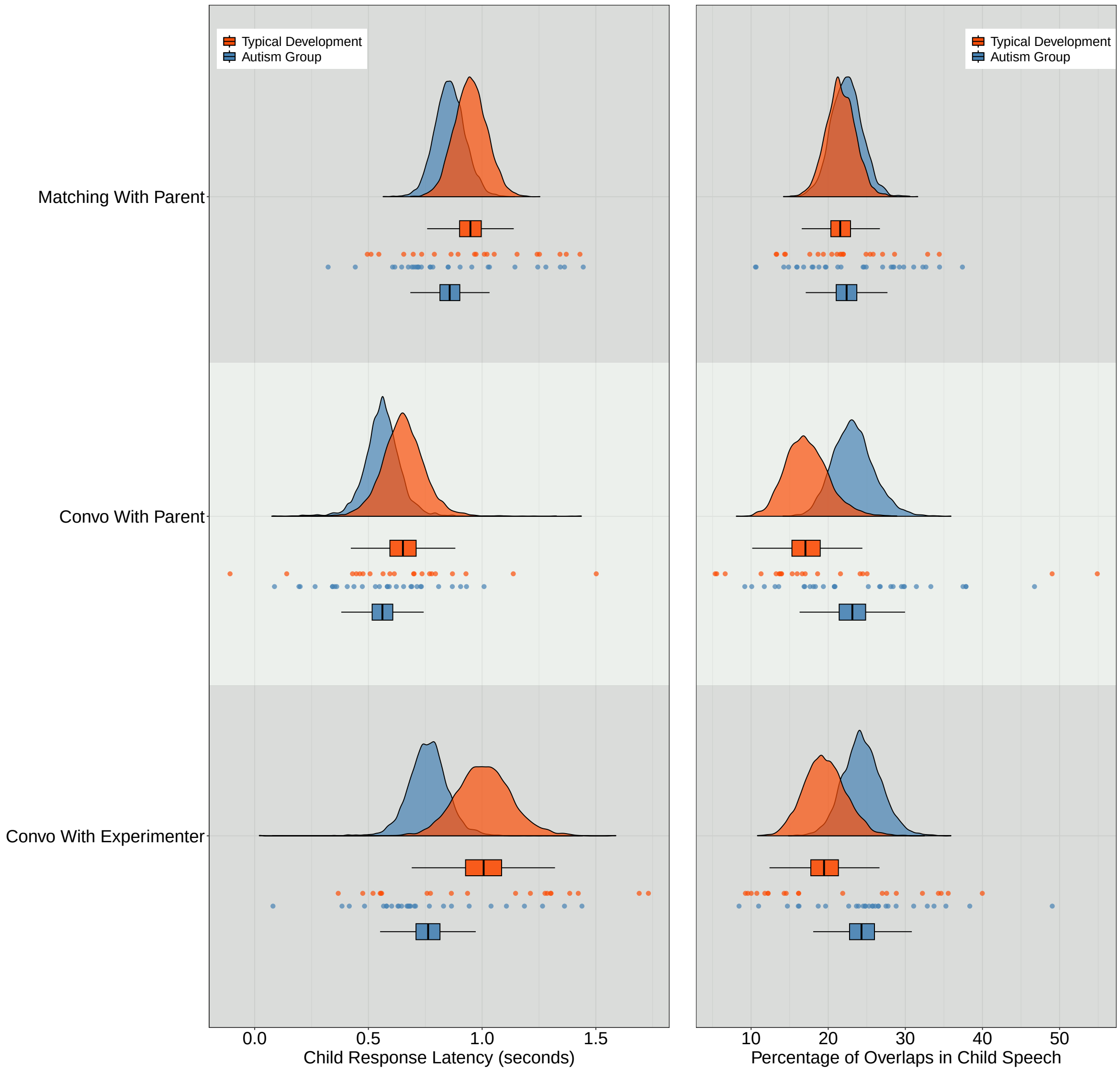

Figure 2: *Panel of plots to demonstrate model estimates for child response latencies (left) and percentage of overlaps (right) across different conversational contexts for the two diagnostic groups: Autism Group (blue) and Typical Development Group (orange). The small points represent participant-level posterior predictions from the model. The density plots and boxplots show aggregated posterior predictions for the conversational activities.*

When we look at the adult interlocutors, their response latencies showed less pronounced changes across different social contexts and no robust differences between the autism and typical development group (see **Figure 6** and **Table 6** in Supplementary Materials).

Child response latencies also differed across genders: autistic girls exhibited faster response latencies than autistic boys and produced similar response latencies to those of typically developing children (cf., Child Latency Estimates Across Gender in Supplementary Materials).

Table 2: *Posterior estimates for child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency) in milliseconds, proportion of latencies below zero (Overlap Proportion) in proportions, Variance (Sigma) in milliseconds squared, exponential component (Beta, or long pauses) on log scale, and Test-Retest Reliability across sessions in correlations. All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	858ms [752, 970]	949ms [834, 1070]
	Convo With Parent	562ms [437, 690]	653ms [501, 807]
	Convo With Experimenter	762ms [626, 901]	1008ms [811, 1211]
	Aggregate Estimate	727ms [641, 813]	870ms [766, 975]
Overlap Proportion	Matching With Parent	0.22 [0.19, 0.26]	0.22 [0.19, 0.25]
	Convo With Parent	0.23 [0.19, 0.28]	0.17 [0.13, 0.22]
	Convo With Experimenter	0.24 [0.2, 0.29]	0.2 [0.15, 0.24]
	Aggregate Estimate	0.31 [0.26, 0.36]	0.19 [0.17, 0.22]
Sigma	Matching With Parent	618ms [573, 666]	529 [483, 580]
	Convo With Parent	534ms [475, 600]	602 [525, 689]
	Convo With Experimenter	962ms [889, 1037]	899 [829, 977]
	Aggregate Estimate	682ms [647, 719]	659ms [621, 700]
Beta	Matching With Parent	0.19 [0.11, 0.27]	0.28 [0.18, 0.37]
	Convo With Parent	-0.35 [-0.47, -0.23]	-0.31 [-0.46, -0.17]
	Convo With Experimenter	-0.25 [-0.37, -0.15]	-0.18 [-0.33, -0.03]
	Aggregate Estimate	-0.14 [-0.2, -0.07]	-0.07 [-0.15, 0.01]
Test-Retest Reliability	Matching With Parent	0.42 [0.24, 0.48]	0.28 [-0.08, 0.48]
	Convo With Parent	0.83 [0.81, 0.86]	0.81 [0.8, 0.82]
	Convo With Experimenter	0.7 [0.66, 0.78]	0.61 [0.48, 0.68]

3.1.2 Frequency and impact of overlaps

Overlaps were also specifically affected by context and diagnostic group, but our findings were mostly robust to removing overlaps, (see **Table 2** and **Figure 2**).

Autistic children were more likely to overlap with their caregivers compared to typically developing childrenwhen aggregated across contexts (0.04 [0, 0.08], ER = 21.1). Conversations with parents displayed the biggest group difference (0.06 [0, 0.12], ER = 15.8) followed by conversations with the experimenter (0.05 [-0.01, 0.11], ER = 10.4, albeit not reliably different from conversations with the parent -0.01 [-0.09, 0.07], ER = 1.5), and there were no differences in the matching game (0.01 [-0.04, 0.05], ER = 1.6, which was reliably lower than conversations with parents, -0.05 [-0.12, 0.01], ER = 9.2).

These differences in overlaps notwithstanding, removing overlaps did not overly affect the findings on response latencies. Autistic children were still faster than the typical development group in the matching game and conversations with the experimenter, albeit not with parents (see **Table 10** in Supplementary Materials).

3.1.3 Variance in Response Latencies (Sigma)

Different contexts were related to different variance in response latencies (see **Table 2** and **Figure 2**).

Conversations with an unfamiliar experimenter displayed the highest variance for both groups (on log scale, ASD: 0.59 [0.45, 0.73], ER = Inf; TD: 0.4 [0.24, 0.56], ER = Inf). This was followed by the matching for autism (0.15 [0.01, 0.29], ER = 21.8), and conversation with parent for typical development (-0.13 [-0.29, 0.03], ER = 9.4).

Autistic children displayed a slightly higher degree of variance than typically developing children, aggregated across conditions (on log scale: 0.03 [-0.04, 0.11], ER = 3.2), with the strongest between-group difference being in the matching game (on log scale: 0.16 [0.04, 0.27], ER = 56.5) followed by the conversations with parents (on log scale: -0.12 [-0.3, 0.06], ER = 6.4) and conversations with the experimenter (on log scale: 0.07 [-0.04, 0.18], ER = 5.4). There was evidence for a more pronounced between-group difference in variance when comparing experimenter with parent conversations (on log scale: 0.19 [-0.03, 0.39], ER = 12.3) as well as when comparing the matching game to conversations with parents (on log scale: 0.27 [0.06, 0.49], ER = 59.2).

3.1.4 Long Pauses (Beta)

Different contexts elicited different prevalence of long pauses (i.e., the exponential tail) was influenced by social context and diagnostic group (see **Table 2** and **Figure 2**)

Across groups, the matching game displayed the highest tendency to long pauses followed by conversations with the unfamiliar experimenter (difference between the two conditions: ASD: 0.54 [0.41, 0.68], ER = Inf; TD: 0.59 [0.43, 0.76], ER = Inf). Conversations with parents had the lowest tendency to low pauses across groups (compared to conversations with experimenters: ASD: 0.09 [-0.07, 0.26], ER = 5.1; TD: 0.13 [-0.07, 0.33], ER = 6.6).

Autistic children produced fewer long pauses when aggregated across conditions (on a log scale: -0.07 [-0.17, 0.04], ER = 6) and during both the matching game (-0.09 [-0.21, 0.04], ER = 7) and experimenter conversations (-0.08 [-0.26, 0.11], ER = 3.2), but not during the conversations with parents (-0.04 [-0.23, 0.15], ER = 1.7). However, there was no reliable difference in group differences between contexts.

3.2. Individual Differences (RQ3)

3.2.1 Social Cognition

Children with higher social cognition skills displayed faster latencies in both the autism group (-71ms [-194, 55], ER = 5) and the typical development group (-419ms [-792, -55], ER = 30.4). The effect of higher social cognition skills was slightly stronger in the matching game compared to conversations with parents for typically developing children (-243ms [-875, 381], ER = 3), but not for the autism group (-12ms [-191, 168], ER = 1.2). Similarly, the typically developing children showed stronger effects of social cognition skills when engaging with the experimenter compared to their parents (-558ms [-1407, 307], ER = 6.3), but this was not the case for the autism group (46ms [-183, 278], ER = 1.7). The effect of social cognition was more pronounced in the typical development group than in the autism group when aggregated across conditions (349ms [-46, 738], ER = 12.7) and was highest in conversations with the experimenter (675ms [-83, 1403], ER = 13.7), followed by the matching game (301ms [-128, 729], ER = 6.9), and there were no clear group differences in conversations with parents (70ms [-546, 676], ER = 1.4). The between-group difference in the effect of social cognition skills was larger in conversations with the experimenter than with parents (604ms [-287, 1488], ER = 6.6) and slightly larger in the matching game compared to parent conversations (230ms [-421, 896], ER = 2.6).

3.2.2 Social Awareness

Higher social awareness skills produced a slowdown for the typical development group (270ms [-8, 548], ER = 17.5) but no changes in the autism group (11ms [-116, 140], ER = 1.3). The effect of social awareness skills on the response latencies of typically developing children was stronger in conversations with the experimenter compared to with parents (472ms [-107, 1077], ER = 10), whereas the autism group showed no differences (-68ms [-298, 166], ER = 2.2). Similarly, typically developing children with higher social awareness skills slowed down their response latencies more in the matching game compared to in conversations with parents (200ms [-251, 653], ER = 3.5), and this was not the case with the autism group (-36ms [-227, 150], ER = 1.7). The effect of social awareness produced a greater slowdown in typical development group compared to the autism group when aggregated across conditions (-259ms [-569, 48], ER = 11.2) and was highest in experimenter conversations (-540ms [-1095, 18], ER = 16.4) followed by the matching game (-236ms [-571, 99], ER = 7.6), and there were no clear differences in conversations with parents (-0.02ms [-470, 457], ER = 1). The between-group difference in the effect of social awareness skills was stronger in the experimenter conversations compared to parent conversations (-540ms [-1185, 85], ER = 11.9) and slightly stronger in the matching game compared to parent conversations (-236ms [-737, 263], ER = 3.8).

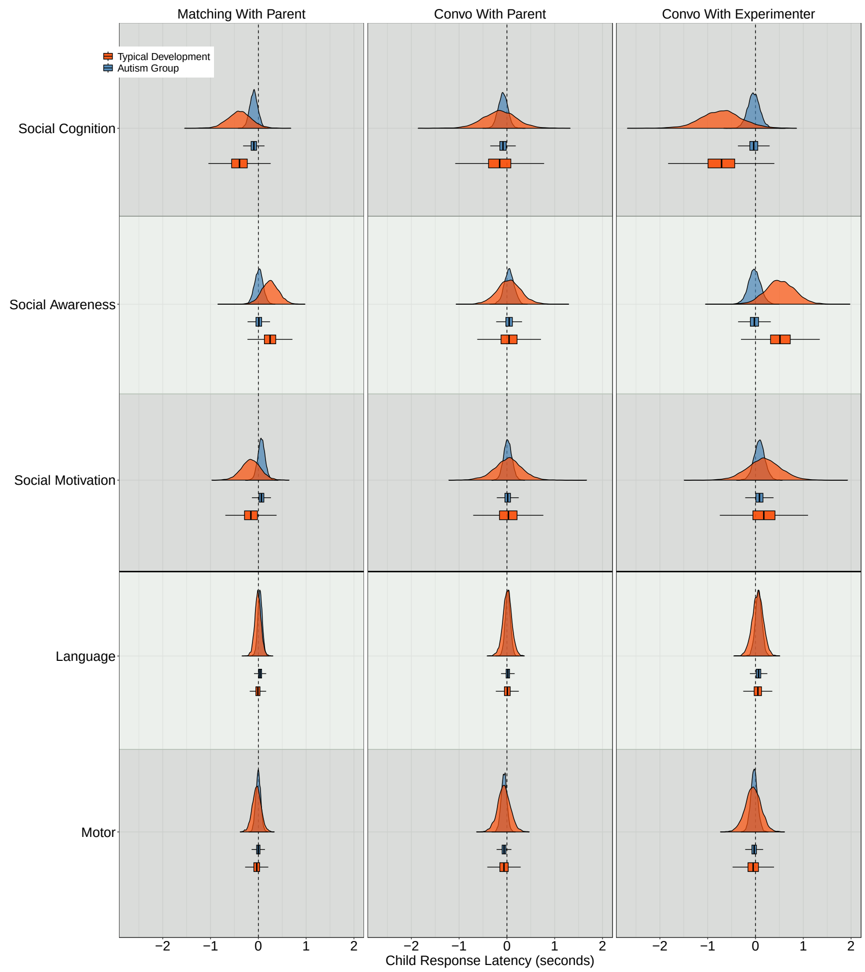

Figure 3: *Panel of plots to demonstrate how model estimates for child response latencies change with measures of individual differences. The posterior estimates denote the change in child response latency as a function of one standard deviation increase in each measure of individual differences. The three top measures (i.e., Social Cognition, Social Awareness, Social Motivation) come from the Social Responsiveness Scale, Motor Skills refer to Vineland Adaptive Behaviour Scales, and Language skills refer to the Clinical Evaluation of Language Fundamentals.*

3.2.3 Social Motivation

Children who displayed higher social motivation scores produced slightly slower response times in the autism group (55ms [-55, 166], ER = 3.9) but similar response times in the typical development group (16ms [-300, 323], ER = 1.2). The slowdown effect of social motivation skills on the response latencies of typically developing children was only slightly stronger in conversations with the experimenter compared to with parents (150ms [-544, 835], ER = 1.8), and the same was the case for the autism group (65ms [-138, 267], ER = 2.4). In contrast, typically developing children with higher social motivation skills slightly sped up their latencies in the matching game compared to the conversations with parents (-182ms [-682, 343], ER = 2.8), whereas autistic children slightly slowed down their latencies in the matching game compared to the conversations with parents (47ms [-115, 211], ER = 2.3). The effect of social motivation did not produce a robustly greater slowdown in the autism group compared to the typical development group when aggregated across conditions (39ms [-290, 377], ER = 1.4). The autism group displayed a greater slowdown compared to the typical development group in the matching game (220ms [-146, 576], ER = 5.7, but in the conversations with the experimenter, the typical development group showed a bigger slowdown compared to the autism group (-94ms [-709, 532], ER = 1.5), and there were no between-group differences in the conversations with parents (-9ms [-501, 499], ER = 1.1). The between-group difference in the effect of social motivation skills was no stronger in the experimenter conversations compared to parent conversations (-85ms [-816, 643], ER = 1.4) and only slightly stronger in the matching game compared to parent conversations (229ms [-321, 752], ER = 3.4).

3.2.4 Language Skills

Children with higher linguistic skills showed a tendency for slightly slower response times in the autism group (40ms [-28, 108], ER = 5.1) and no clear differences in the typical development group (16ms [-84, 116], ER = 1.6). The slowdown effect of linguistic skills on response latencies was only slightly stronger in the matching game compared to in conversations with parents for the autism group (17ms [-87, 118], ER = 1.6), and there were no differences for the typical development group (-17ms [-182, 146], ER = 1.3). The effect of linguistic skills on response latencies was also no stronger when conversing with the experimenter compared to parents in the autism group (46ms [-80, 170], ER = 2.7) or the typical development group (39ms [-187, 268], ER = 1.6). The effect of linguistic skills did not produce a robustly greater slowdown in the autism group compared to the typical development group neither when aggregated across conditions (24ms [-94, 149], ER = 1.7), nor within the conditions of matching game (44ms [-93, 182], ER = 2.4), conversations with experimenters (17ms [-199, 240], ER = 1.2), or conversations with parents (11ms [-170, 193], ER = 1.2). The between-group difference in the effect of linguistic skills did not change across neither conversations with the experimenter and parent (7ms [-254, 262], ER = 1) nor across tasks of matching game compared to parent conversations (33ms [-158, 225], ER = 1.6).

3.2.5 Motor Skills

Children with higher motor skills had slightly faster response latencies in the autism group (-30ms [-103, 44], ER = 3) as well as in the typical development group (-48ms [-189, 97], ER = 2.5). The speedup effect of motor skills was slightly stronger in conversations with parents compared to with experimenters for the autism group (33ms [-93, 161], ER = 2) and the typical development group (12ms [-316, 341], ER = 1.1). Similarly, the speedup effect of motor skills was stronger in conversations with parents compared to in the matching game for the autism group (59ms [-50, 170], ER = 4.6) but not for the typical development group (27ms [-208, 259], ER = 1.4). There was no evidence of between-group differences in the effect of motor skills when aggregated across conditions (18ms [-143, 175], ER = 1.4) and within the conditions of matching game (32ms [-147, 204], ER = 1.7), conversations with experimenters ($asd < tdc$: 22ms [-275, 325], ER = 1.2) and conversations with parents

(0.37ms [-239, 247], ER = 1). The extent of between-group difference in the effect of motor skills did not change neither across conversations with the experimenter and parents (21ms [-337, 374], ER = 1.2) nor across the matching game and parent conversations (32ms [-233, 290], ER = 1.4).

Table 3: *Posterior estimates for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of each of the types of skills listed below (Motor, Cognitive Skills, Language Skills, Social Awareness and Social Motivation). The estimates denote the change in child response latency as a function of one standard deviation increase in the skills of the children.*

	Skills	**Autism Group**	**Typical Development**
Social Cognition	Matching With Parent	-94ms [-233, 48]	-395ms [-804, 17]
	Convo With Parent	-82ms [-245, 78]	-152ms [-734, 436]
	Convo With Experimenter	-36ms [-245, 175]	-711ms [-1400, 5]
	Aggregate Estimate	-71ms [-194, 55]	-419ms [-792, -55]
Social Awareness	Matching With Parent	10ms [-139, 159]	246ms [-49, 550]
	Convo With Parent	46ms [-123, 217]	46ms [-373, 476]
	Convo With Experimenter	-22ms [-234, 194]	518ms [21, 1031]
	Aggregate Estimate	11ms [-116, 140]	270ms [-8, 548]
Social Motivation	Matching With Parent	65ms [-60, 191]	-156ms [-486, 179]
	Convo With Parent	18ms [-125, 162]	26ms [-468, 499]
	Convo With Experimenter	82ms [-103, 274]	176ms [-399, 757]
	Aggregate Estimate	55ms [-55, 166]	16ms [-300, 323]
Language	Matching With Parent	36ms [-42, 116]	-8ms [-117, 104]
	Convo With Parent	19ms [-71, 109]	9ms [-146, 162]
	Convo With Experimenter	65ms [-51, 180]	48ms [-148, 236]
	Aggregate Estimate	40ms [-28, 108]	16ms [-84, 116]
Motor	Matching With Parent	-1.33ms [-87, 85]	-34ms [-186, 118]
	Convo With Parent	-61ms [-158, 37]	-61ms [-288, 162]
	Convo With Experimenter	-27ms [-148, 94]	-49ms [-326, 219]
	Aggregate Estimate	-30ms [-103, 44]	-48ms [-189, 97]

As a control analysis, we ran the individual differences model on the data when excluding overlaps and found similar results for social cognition, social awareness and social motivation measures as well as no clear effects of linguistic and motor skills (cf., **Table 15** in Supplementary Materials). We also evaluated the extent to which the proportion of produced overlaps related to individual differences and found that the proportion of overlaps increases for children with higher social cognition in the autism and typical development group, decreases with higher social awareness skills for the typical development group, increases for children with higher social motivation skills in the typical development group but decreases for children with higher social motivation in the autism group (see **Figure 13** in Supplementary Materials).

3.3. Information Flow

3.3.1 Predictability of Previous Adult Utterance

The predictability of an utterance had a robust influence on the response latency of the following utterance (see **Table 4** and **Figure 4**). The higher the predictability of the previous adult utterance, the faster the response latency for both the autism group (-29ms [-48, -10], ER = 118) and typical development group (-32ms [-61, -2], ER = 23.3). The children exhibited stronger effects of predictability in the parent conversations compared to experimenter conversations in the autism group (-30ms [-80, 23], ER = 4.9) and typical development group (-61ms [-140, 19], ER = 9). While the autism group showed no difference in the effects of predictability in the matching game compared to parent conversations (0.2ms [-45, 43], ER = 1.1), the typical typical development group showed stronger effects in the parent conversations compared to the matching game (54ms [-17, 123], ER = 8.9). There is no evidence of between-group differences in the effect of predictability on response latencies when aggregated across conditions (2.31ms [-34, 38], ER = 1.2). Within conditions, the autism group exhibited a slightly steeper negative slope in the matching game compared to the typical development group (-23ms [-69, 22], ER = 4.2), whereas the typical development group showed a stronger negative influence of adult utterance predictability than the typical development group in conversations with parents (31ms [-39, 100], ER = 3.4), and there were no robust group differences in conversations with experimenters (-1.05ms [-67, 65], ER = 1). The between-group differences were only slightly more pronounced in parent conversations than in experimenter conversations (32ms [-63, 128], ER = 2.4) as well as in parent conversations compared to the matching game (-53ms [-137, 29], ER = 5.8).

3.3.2 Predictability of Own Utterance

The predictability of children's own utterances also had a robust influence on their response latencies (see **Table 4** and **Figure 4**). The higher the predictability of their own utterances, the slower the response latencies for the autism group (68ms [45, 91], ER = Inf) and the typical development group (50ms [22, 78], ER = 415.7). There were stronger positive effects of predictability when conversing with the experimenter compared to with the parent for both the autism group (-121ms [-178, -63], ER = 453.5) and the typical development group (-114ms [-187, -39], ER = 130.6). The children also showed a stronger positive effect of the predictability of their own utterance when conversing with parents compared to in the matching game in the autism group (-48ms [-97, 1], ER = 17.8), but not in the typical development group (19ms [-34, 75], ER = 2.6). The influence of the predictability of children's own utterances is slightly stronger for the autism group compared to the typical development group when aggregated across conditions (18ms [-17, 55], ER = 3.9), with the strongest difference being in the matching game (-48ms [-97, 1], ER = 17.8), followed by the conversations with parents (38ms [-13, 91], ER = 8.2) and conversations with experimenters (46ms [-32, 124], ER = 5.3). There was no evidence that the between-group differences were more pronounced in conversations with parents than with experimenters (-7ms [-100, 86], ER = 1.2), but the between-group difference in slopes was more pronounced in the parent conversations compared to the matching game (-68ms [-141, 4], ER = 16.3). A control analysis excluding backchannels and interjections finds the same patterns of influence for both the adults' and children's utterances (see **Table 18** and **Table 19** in Supplementary Materials).

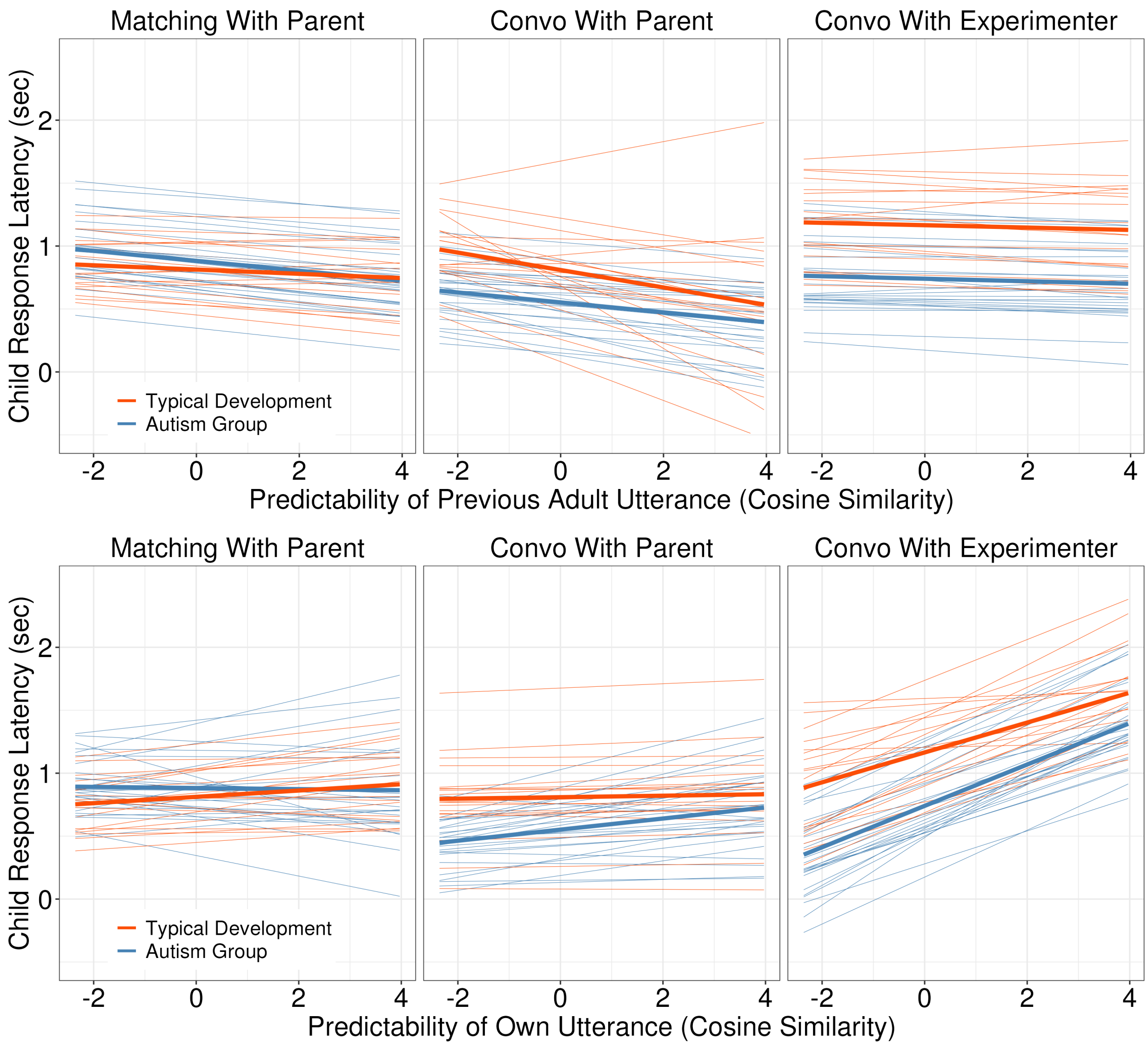

Figure 4: *Panel of plots to demonstrate how model estimates for child response latencies change with the predictability of the previous adult response latency (top) and semantic predictability of own utterance (bottom). The faded lines are posterior predictions from the model for individual child participants in the study, whereas the thicker lines are average predictions*

Table 4: *Posterior estimates for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous adult utterance and child utterance. The estimates denote the change in child response latency as a function of one standard deviation increase in the predictability of the utterance*

.

	Tasks	**Autism Group**	**Typical Development**
Predictability of Previous Adult Utterance	Matching With Parent	-39ms [-64, -13]	-16ms [-52, 23]
	Convo With Parent	-39ms [-74, -3]	-70ms [-128, -10]
	Convo With Experimenter	-10ms [-46, 28]	-9ms [-61, 45]
	Aggregate Estimate	-29ms [-48, -10]	-32ms [-61, -2]
Predictability of Own Utterance	Matching With Parent	-4.24ms [-38, 33]	25ms [-14, 67]
	Convo With Parent	44ms [8, 81]	6ms [-31, 42]
	Convo With Experimenter	165ms [122, 210]	119ms [56, 183]
	Aggregate Estimate	68ms [45, 91]	50ms [22, 78]

3.4. Turn-By-Turn Dynamics

3.4.1 Interpersonal Adjustment

The response latency of the previous utterances had a robust influence on the response latency of the following utterance (see **Table 5** and **Figure 5**). Whereas the autism group showed a positive adjustment to their interlocutors' response latencies (28ms [5, 50.08], ER = 41.9), the typical development group did not (3.03ms [-29, 35], ER = 1.3). There was strong evidence that the children in the autism group adapted more to the response latencies of their interlocutors when interacting with their parents compared to the experimenter (68ms [-0.13, 137.54], ER = 18.7), but this was not the case for the typical development group (20ms [-71, 110], ER = 1.9). There was also strong evidence that the autism group adapted more to the response latencies of their parents in conversations rather than in the matching game (-38ms [-80, 3], ER = 14.3), but again, the same was not the case for the typical development group (3.95ms [-50, 60], ER = 1.1). The autism group displayed a greater degree of interpersonal adjustment than the typical development group when aggregated across conditions (25ms [-13, 64], ER = 5.7), especially within conversations with parents (55ms [-3.54, 117.38], ER = 15) followed by the matching game (13ms [-23, 48], ER = 2.6), but there was no evidence of a between-group difference in the conversations with the experimenter (6ms [-87, 98], ER = 1.2). There was a slightly greater between-group difference in interpersonal adjustment when children were engaging with their parents than with the experimenter (49ms [-70, 160], ER = 3.4) as well as in the matching game compared to conversations with a parent (-42ms [-113, 26], ER = 5.4).

3.4.2 Self Adjustment

The response latency of the children's own previous utterances also had a robust influence on the response latency of their following utterance (see **Table 5** and **Figure 5**). The slower the response latencies of their previous utterances, the slower the current response latencies for both the autism group (35ms [16, 53], ER = 749) and typical development group (13ms [-13, 38], ER = 3.9). The children adapted slightly more to their own response latencies when interacting with their parents compared to the experimenter in the autism group (23ms [-29, 74], ER = 3.5) and in the typical development group (27ms [-42, 99], ER = 2.6). Children in the autism group adapted slightly less to their own response latencies in the matching game compared to when conversing with their parents (-22ms [-65, 22], ER = 3.9), whereas there were no differences of this sort in the typical development group (11ms [-54, 73], ER = 1.6). The autism group displayed a greater degree of self-adjustment when aggregated across conditions (22ms [-8, 54], ER = 7.3), with the strongest difference being in conversations with experimenters (36ms [-20, 91], ER = 5.8), followed by conversations with parents (32ms [-39, 101], ER = 3.7), and there were no differences in the matching game (-1.05ms [-33, 30], ER = 1.1). There was no evidence for a greater between-group difference in self adjustment when children were engaging with their parents than with the experimenter (-3.62ms [-93, 83], ER = 1.1) nor in the matching game compared to in parent conversations (-33ms [-109, 44], ER = 3.3).

Table 5: *Posterior estimates for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the latency of the previous adult utterance (Interpersonal Adjustment) and previous child response latency (Self-Adjustment). The estimates denote the change in child response latency as a function of a one-second increase in the response latency of the previous utterance.*

	Tasks	**Autism Group**	**Typical Development**
Interpersonal Adjustment	Matching With Parent	25ms [5, 45.23]	12ms [-16, 42]
	Convo With Parent	63ms [27, 101]	8ms [-42, 55]
	Convo With Experimenter	-5ms [-58, 49]	-11ms [-89, 64]
	Aggregate Estimate	28ms [5, 50.08]	3.03ms [-29, 35]
Self Adjustment	Matching With Parent	28ms [9, 46]	29ms [2.59, 54.77]
	Convo With Parent	50ms [9, 90]	18ms [-36, 77]
	Convo With Experimenter	27ms [-6, 58]	-9ms [-55, 37]
	Aggregate Estimate	35ms [16, 53]	13ms [-13, 38]

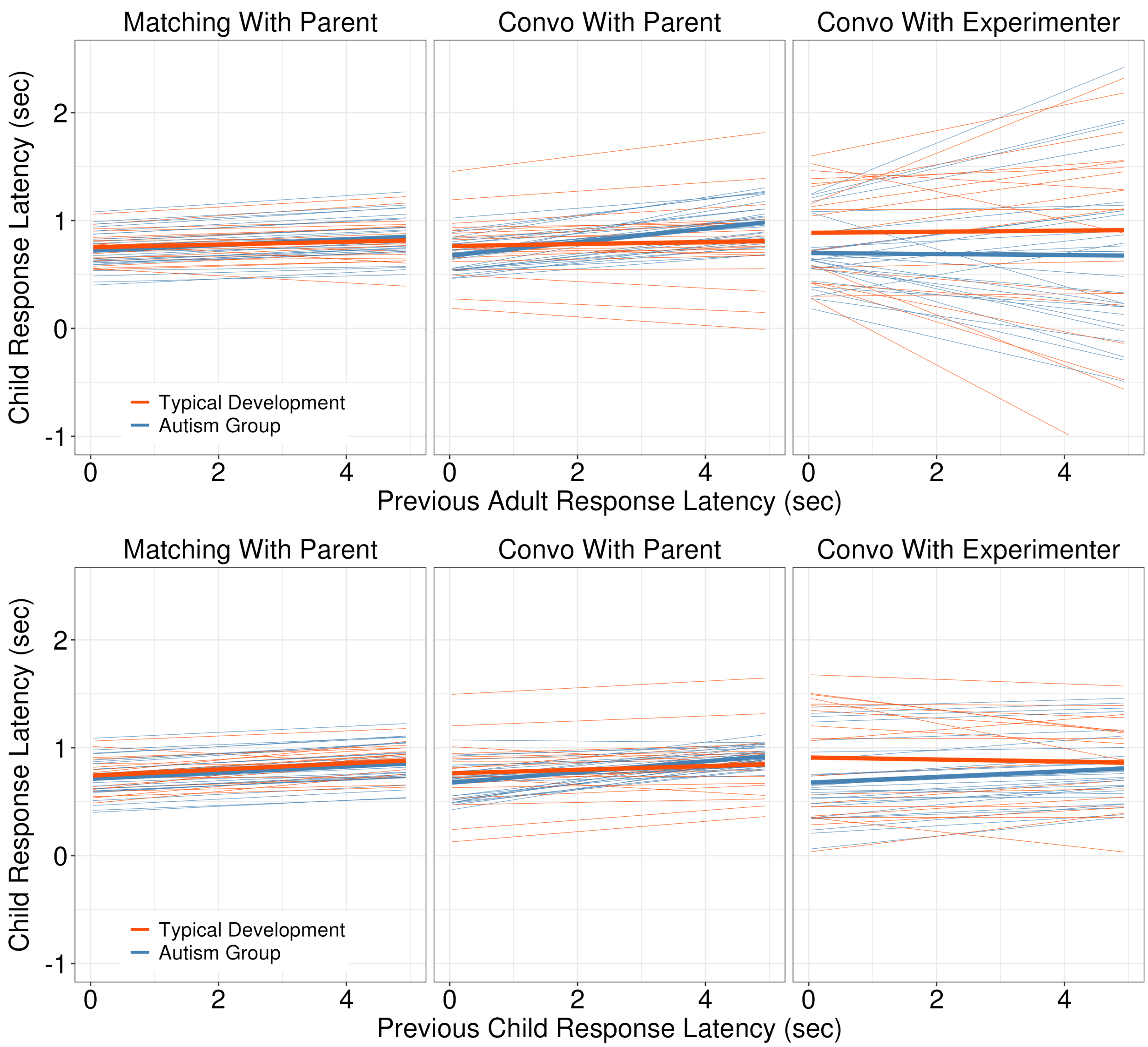

Figure 5: *Panel of plots to demonstrate how model estimates for child response latencies change with the latency of the previous adult response latency (top) and the previous child response latency (bottom).*

4. Discussion

Understanding turn-taking dynamics in child-adult conversations requires investigation of how results generalize and vary across different social contexts (tasks and familiarity), individual differences in socio-cognitive, motor and linguistic skills, and turn-by-turn exchanges (information and interpersonal adjustment). Our results showed that the test-retest reliability within social contexts for each diagnostic group remained high across the seven sessions – especially in conversational tasks – suggesting that turn-taking dynamics represent a stable, generalizable construct across repeated sessions of child-adult interactions. Autistic children produced faster and more variable inter-turn response latencies, more overlaps and shorter pauses compared to typically developing children. Individual differences among children also mattered: both autistic and typically developing children with higher cognitive and motor skills produced faster response latencies; autistic children with higher social motivation and linguistic scores – and typically developing children with higher social awareness scores - exhibited slower responses. The turn-by-turn exchange of information also produced different response latencies. The higher the predictability of the speech content in the previous parent – but not experimenter - utterance, the faster both the autistic and typically developing children responded. Similarly, the higher the predictability of the child's own utterance, the slower both groups of children responded, especially when interacting with the experimenter. Only the autism group adjusted their response times to those of the adult's preceding turn – but again, only in interactions with their parents and not the experimenter. Both groups displayed positive auto-correlation of their own successive response latencies. Together, these findings highlight how turn-taking dynamics are shaped by social context, individual differences as well as turn-by-turn flows of information. In the following sections, we contextualise these findings in relation to our Research Questions and discuss their implications for future studies.

4.1. Temporal Dynamics as Shaped by Affordances from Social Context

4.1.1 Child Turn-Taking in Telephone Conversations

Our first two research questions concerned the extent to which autistic children produced faster, slower or similar inter-turn response latencies compared to typically developing children across different social contexts and with different interlocutors. Both groups of children responded with latencies of under 1 second when including overlaps and around 1.2 seconds when excluding overlaps, which conforms to other studies of child response latencies in conversations (Hilbrink et al., 2015; Nguyen et al., 2022) (although see (Fusaroli et al., 2024)). This estimate, however, is slightly higher than cross-linguistic estimates for adult-adult conversations Dingemanse & Liesenfeld (2022); Stivers et al. (2009); Wehrle et al. (2023b). The slower response latencies observed here may be partly because interactants participated in conversations over the telephone (i.e., only the matching game condition took place face-to-face). The contextual demands of a task is often reflected in conversational dynamics, including the type and timing of speaker changes (C. Dideriksen, Christiansen, Tylén, et al., 2023; Ervin-Tripp, 1979; Garvey & Berninger, 1981; Garvey, 1975). Conversations conducted on telecommunication systems produce delays in response times for adults (135ms in face–to–face conversations versus 487 ms in conversations conducted via teleconference (Boland et al., 2022)) and likely also do so for children in child-adult interactions. Despite the longer response latencies, interlocutors readily adapt to different degrees of delays in responses over the phone and exhibit robustness in measures of perceived interaction quality (Schoenenberg et al., 2014). Telephone conversations thus likely represent a specific social context with affordances that require longer response times from both children and adults (see **Table 6**). These results further highlight that verbal interactions are always situated within a dynamic social context with different task demands.

4.1.2 Distributional Properties

The distributional properties of inter-turn response latencies differed according to the experimental task and familiarity of the interlocutor. Across groups, the matching game involved a higher propensity to occasional long pauses compared to the conversations with either the parent or experimenter. We might speculate that the matching game – with its demands to find the relevant locations to place the stickers and precisely communicate these locations to one's interlocutors – would generate more frequent long pauses in the conversation. Children also produced slower and more variable response latencies as well as longer pauses when conversing with the experimenter compared to their own parent. This result may be a function of the relatively higher degree of unpredictability involved in interactions with unfamiliar interlocutors (Kimhi et al., 2014; Ying Sng et al., 2018) and the consequent lack of shared personal experiences and world knowledge of the interactants (Pickering & Garrod, 2021; Tylén et al., 2013). We may expect individual differences in socio-cognitive abilities to influence the degree of behavioural flexibility of children across social contexts with different degrees of familiarity and predictability, as discussed further below.

4.1.3 Typical Development & Autism

While it has been often argued that autistic individuals exhibit slower response latencies in conversations (Nguyen et al., 2022), recent evidence suggests similar or faster turn-taking latencies in autism compared to controls (Fusaroli et al., 2024; Wehrle et al., 2023b). This is further supported by our findings. The underlying reason for our finding, however, may be different. In Fusaroli et al. (2024), the faster response latencies were driven by larger overlaps in autistic children. In our study, excluding overlaps still results in autistic children responding faster than typically developing children in the matching game and experimenter conversations, but the group difference in the parent conversations disappears (cf., **Table 10** in Supplementary Materials). Thus, in a social context similar to Fusaroli et al. (2024) - that is, in conversations with a parent - autistic participants produced similar response latencies but more overlaps compared to typically developing children. In other social contexts, overlaps did not explain the divergence between the groups. A second control model excluding the varying beta parameter (to capture long tails) explains away the difference in the matching game but not in the conversations (cf., **Table 14** in Supplementary Materials). This suggests that autistic children were not faster (or slower) than typically developing children in general response latencies in the matching, but were less likely to produce the occasional long pause. Importantly, these differences are unlikely to be driven by the adults in the interaction: across social contexts, the parents and experimenter displayed comparable response latencies, overlap proportions, long pauses and degrees of variability (cf., **Table 6** in Supplementary Materials), also when excluding overlaps from the model (cf., **Table 11** in Supplementary Materials). A speculative explanation for these patterns could make reference to group differences in impulsivity and executive control (Hughes et al., 1994; Hill, 2004; Hlavatá et al., 2018; Geurts et al., 2014). When interacting with a parent in a casual social context of a conversation, autistic children may display difficulties in inhibiting their responses and therefore produce premature responses that overlap with their interlocutors. Similarly, in the matching task, which requires more precision and focus than a conversation, autistic children may jump in without waiting as long as typically developing children. Any explanations invoking inhibition and impulsivity should be taken as highly tentative, however, as there are mixed results regarding these components of autism (Hill, 2004; Geurts et al., 2014).

4.1.4 Social Responsivity

When controlling for overlaps and long pauses, then, the main difference between the autism and typical development group concerned their response latencies when interacting with the experimenter. Whereas typically developing children displayed slower latencies (300ms) when interacting with the experimenter, autistic children did so to a lesser degree (200ms). While little research on speaker familiarity in conversations with autistic children is available (Ying Sng et al., 2018; Nguyen et al., 2022), we can advance some speculative interpretation. Autistic children are argued to display less responsivity to new social contexts compared to typically developing children (Stefanatos & Baron, 2011). Compared to typically developing children, for example, autistic children show no difference in gaze aversion when interacting with familiar and unfamiliar interlocutors (Doherty-Sneddon et al., 2013), no difference in brain activation in response to familiar versus unfamiliar face stimuli (Dawson et al., 2002), and less flexibility in their behavioural repertoire across different social tasks (Forgeot d'Arc et al., 2020). Investigating how turn-taking dynamics change in different social contexts permits insights into how children re-adjust expectations as task demands, social contexts and interlocutors change. These group differences are likely to involve a large suite of socio-cognitive abilities (Frith & Happé, 1994). For example, executive function in autism is different to that in typical development (Russell, 1997; Hlavatá et al., 2018), and this has downstream effects on planning, mental flexibility, impulsivity and inhibition in social interactions. A more nuanced view of responsivity to new social contexts in atypical populations could also look at how interactants make different use of the common ground with changes to different facets of the social context. For example, some autistic individuals have been shown not to modify their referental descriptions in the presence of a new interactant who did not share their common ground – and that those who did adapt their referential descriptions exhibited delayed responses in the majority of autistic participants (Nadig et al., 2015). Further investigation is thus needed to examine the impact of these relatively subtle differences in communicative patterns and use of common ground in situations of unpredictability and unfamiliarity. In other words, a single underlying account of group differences in behavioural flexibility is unrealistic (Baron-Cohen et al., 1985; Devaine et al., 2014; Waade et al., 2023), and there is now a greater focus on explaining the particular symptom domains, such as restricted or repetitive behaviours, which are related to inflexibility in autistic children (Kenworthy et al., 2010; Yerys et al., 2009).

4.2. Individual Differences Matter

4.2.1 Social Cognition

In line with this more granular perspective on individual differences, we investigated the grounding of turn-taking dynamics in socio-cognitive, motor and linguistic skills. Measures of social cognition exerted the strongest effects on inter-turn response latencies: the higher the social cognition skills, the faster the response latencies, because of both increased overlapping (cf., **Figure 13**) and shorter gaps (cf., **Table 15**). These findings held for both groups, but were stronger in typically developing children. Given the cognitive challenges involved in turn-taking (Levinson, 2006, 2016; Magyari et al., 2014) and the complex information integration from semantics, syntax, and pragmatics that takes place in real time, a more developed social cognition may help children to anticipate an imminent termination and to launch their own turn faster (Magyari & De Ruiter, 2012). Social cognition measures children's ability to interpret social cues, and these findings imply that the children who were better at integrating information from pragmatic and social cues both produced faster responses and overlapped with their interlocutor more. The finding that typically developing children with higher social cognition skills produced so much faster response latencies when interacting with the experimenter – while the autistic children with higher social cognition skills did not speed up their responses with experimenters – may be further evidence of a greater degree of behavioural flexibility in the typically developing children (Stefanatos & Baron, 2011). In social contexts with a high degree of unpredictability, such as when interacting with an unfamiliar interactant, higher social cognition skills may allow for a greater degree of flexibility in adapting to the social cues of people in interactions. The precise weighting of these cues and how the integration of them relates to the individual socio-cognitive abilities of autistic and typically developing children requires further experimental and qualitative work (Fusaroli et al., 2023; Bottema-Beutel et al., 2022).

4.2.2 Social Awareness & Social Motivation

The finding that higher degrees of social awareness and social motivation led to relative slowdowns in inter-turn response latencies for the typically developing and autistic children, respectively, may relate to the notion that smooth turn-taking also requires giving space for the interlocutor to contribute to the conversation. In child-adult conversations, for example, adults tend to slow down their own response latencies to around the same response latencies as the children (Nguyen et al., 2022; Fusaroli et al., 2024) – as was the case here (see **Table 6**) – presumably to accommodate the relatively slower modal response latencies of the children and provide them with an opportunity to contribute to the interaction (Fusaroli et al., 2019). In a similar fashion, typically developing children with a greater ability to pick up on social cues may be better at timing their responses to suit those of their interlocutor. Our control analysis that conditioned the proportion of overlaps on individual differences of the children showed a marked decrease in overlap proportion with higher awareness skills for the typically developing children but not for the autistic children (cf., **Figure 13**). Similarly, for social motivation in the autistic children, we see a decrease in the proportion of overlaps with higher motivation skills and a slight increase for the typically developing children (cf., **Figure 13**). These findings imply that typically developing children who are better at picking up on social cues overlap less with their interlocutors – presumably because they spend more time monitoring the other interactant and picking up on their facial expressions and gestures (Dawson et al., 2004). Autistic children with a greater motivation to engage in social communication also produced fewer overlaps with their interlocutors, which may be a product of the children adapting their responses in ways that give space for the other interlocutor to respond. That different skills (i.e., social awareness for the typical development group and social motivation for the autism group) led to slowdowns in each population may indicate different mechanisms to achieve the same effect, but further studies and comparisons would be needed to assert a stronger version of this tentative proposition.

4.2.3 Motor & Linguistic Skills

Lastly, motor skills showed a weak negative effect on child response latencies, which we conjecture could be due to better motor control over the articulators and speech production system. Advanced motor skills may allow for smoother speaking or head control, enabling faster shifting of attention – both of which could facilitate turn-taking (van Der Fels et al., 2019). Similarly, research has provided robust evidence that turn-taking is highly dependent on learning linguistic cues to predict the end of turns (Magyari & De Ruiter, 2012; Lindsay et al., 2019), so the finding of no effect of linguistic skills on the response latencies of typically developing children – and of slightly slower response latencies for those of autistic children – may relate to the speech planning abilities of children, as discussed further in the next section.

4.3. Information Flow

4.3.1 Predictability of Previous Adult Utterance

As mentioned in the introduction, the exquisitely fast timing in turn-taking requires an anticipatory mechanism to be at play (Magyari & De Ruiter, 2012; Levinson, 2016). By measuring the predictability of successive turns in these child-adult conversations, we were able to show that both autistic and typically developing children speed up the response latencies of their turns given a higher predictability in the speech content of the preceding adult's turn. In control analyses, we showed that this effect of utterance predictability persisted even when excluding short utterances and backchannels (cf., **Tables 18 and 19**). This evidence thus corroborates findings from studies showing that prediction based solely on the content (or at least the gist) of the unfolding turn can allow the next speaker to begin planning their turn while still listening to an on-going turn (Bögels et al., 2015; Magyari & De Ruiter, 2012; Magyari et al., 2014; Pickering & Garrod, 2021) and to predict when the current turn may end (Sacks et al., 1974; Lindsay et al., 2019). We should note that this measure of utterance predictability only takes the content of the speech into account – and not other highly informative cues to the ends of turns, including phonetic (Local & Walker, 2012) and prosodic cues (Schaffer, 1983) – and future experimental studies could look into the individual contributions of contextual and prosodic information in the estimation of turn ends (Bögels & Torreira, 2021).

4.3.2 Predictability of Children's Own Utterance

There was also a slowdown of children's own response latencies according to the predictability of the child's own utterance in both groups, especially in the experimenter conversations, even while keeping the predictability of the adult's utterance constant in the model. The overall degree of utterance predictability differed slightly across social contexts, with both the child and adult utterances agreeing with each other in terms of their overall degree of predictability (cf., **Table 16 in Supplementary Materials**). What underlies the slowdown in response latencies according to the predictability on their own utterance requires further investigation, as most studies focus on how the predictability of the previous interlocutor utterance influences response latencies (Lindsay et al., 2019). We can therefore only speculate about the underlying mechanisms in the current context. Saying something predictable in the context also means that there is a greater pressure for the utterance to be correct. The longer response latencies, then, could be due to a more deliberate process of formulating the exact wording or structure of their response, or self-monitoring and considering multiple options before speaking (Levelt, 1993). The reason for their particular slowdown when interacting with an unfamiliar interlocutor could in turn be that children might anticipate that their own utterance is expected by the adult, leading them to pause and reflect to ensure accuracy or appropriateness. This anticipatory pause, then, might be a form of cognitive checking or fine-tuning the utterance before actually speaking, thus increasing the response latency.

4.4. Turn-By-Turn Adjustment

Finally, we assessed whether children would slow down, when the caregiver adopted a slower tempo on turn-by-turn basis, while controlling for the child's own dynamics. Compared to the typically development children, autistic children to a greater extent adapted to the response latencies of their interlocutor – but only with their parent and not the experimenter. In adult-adult telephone conversations, Pouw & Holler (2022) found negative interpersonal auto-correlations, arguing that there is a counter-adjustment mechanism at the dyadic level that emerges from the joint nature of human conversational dynamics. In the child-adult telephone conversations here, our finding of a positive interpersonal correlation between successive turns may instead be indicative of waves of slow and fast turn-taking based on the situational demands of the task Fusaroli et al. (2024). The finding that both self-adjustment and interpersonal adjustment displayed positive effects here also accords with the adults displaying similarly positive auto-correlation structures to those of the children, especially the experimenters in interactions with autistic children **Table 9** and **Figure 9**. Further work on turn-by-turn adaptations in conversation is needed to disentangle whether these results are a product of methodological differences (e.g. Pouw & Holler (2022) did not model concomitant self-adjustment), social context (familiar versus unfamiliar interlocutors) or population (child versus adult & autism versus typical development). Further research should also better assess whether additional time-scales of adjustment are at play and how they might interact (Pouw et al., 2021).

5. Conclusion

At the heart of turn-taking lies mutual interaction with the aim of sharing social interest and communicating information (Harrist & Waugh, 2002; Schertz et al., 2018). The back-and-forth dynamics of child-adult vocal exchanges can straightforwardly be extended to other social contexts. These co-produced structures can also appear, for example, when a child stacks cups to make a tower, and the caregiver stacks another set of cups to make a tower, and the actions are repeated. In these back-and-forth interactions, each interactant creates social affordances for the other to respond without necessarily relying on prompts from the interlocutor. Here we built a framework of crucial components to understanding turn-taking in the context of individual differences, task demands, interlocutor familiarity, and turn-by-turn information flow and dynamics. This framework provides novel insights and strongly situates turn-taking within a more dynamic perspective on social interactions. We hope that these findings can contribute to a reevaluation of how we analyse conversational turn-taking and create stronger theories about turn-taking in diverse populations.

Reproducibility Statement

All code and materials required to reproduce this research are publicly available and documented on github (`LINK`) and OSF (`LINK`).

Acknowledgments

We would like to acknowledge and thank all participating families for being involved with this research project. We would also like to thank audiences at the Max Planck Institute in Nijmegen, University of Melbourne and the International Society for Autism Research for insightful feedback on early versions of this manuscript.

References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *The British Journal of Psychiatry*, *175*(5), 444–451.

Bigelow, A. E. (1999). Infants' sensitivity to imperfect contingency in social interaction. *Early social cognition: Understanding others in the first months of life*, 137–154.

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific reports*, *5*(1), 12881.

Bögels, S., & Torreira, F. (2021). Turn-end estimation in conversational turn-taking: The roles of context and prosody. *Discourse Processes*, *58*(10), 903–924.

Boland, J. E., Fonseca, P., Mermelstein, I., & Williamson, M. (2022). Zoom disrupts the rhythm of conversation. *Journal of Experimental Psychology: General*, *151*(6), 1272.

Bone, D., Black, M. P., Ramakrishna, A., Grossman, R. B., & Narayanan, S. S. (2015). Acoustic-prosodic correlates of'awkward'prosody in story retellings from adolescents with autism. In *Interspeech* (pp. 1616–1620).

Bottema-Beutel, K., Crowley, S., & Kim, S. Y. (2022). Sequence organization of autistic children's play with caregivers: Rethinking follow-in directives. *Autism*, *26*(5), 1267–1281.

Bottema-Beutel, K., Lloyd, B., Watson, L., & Yoder, P. (2018). Bidirectional influences of caregiver utterances and supported joint engagement in children with and without autism spectrum disorder. *Autism Research*, *11*(5), 755–765.

Bruni, T. P. (2014). Test review: Social responsiveness scale–second edition (srs-2). *Journal of Psychoeducational Assessment*, *32*(4), 365-369. doi: 10.1177/0734282913517525

Bryant, G. A. (2020). Evolution, structure, and functions of human laughter. In K. Floyd & R. Weber (Eds.), *The handbook of communication science and biology* (pp. 63–77). Routledge.

Bryant, G. A., Fessler, D. M., Fusaroli, R., Clint, E., Aarøe, L., Apicella, C. L., . . . Zhou, Y. (2016). Detecting affiliation in colaughter across 24 societies. *Proceedings of the National Academy of Sciences*, *113*(17), 4682–4687.

Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language*, *92*, 234–253.

Chatham, C., Taylor, K., Charman, T., Liogier D'Ardhuy, X., Eule, E., Fedele, A., . . . del Valle Rubido, M. (2018). Adaptive behavior in autism: minimal clinically important differences on the vineland-II. *Autism Research*, *11*(2), 270–283.

Chesnut, S. R., Wei, T., Barnard-Brak, L., & Richman, D. M. (2017). A meta-analysis of the social communication questionnaire: Screening for autism spectrum disorder. *Autism*, *21*(8), 920–928.

Cho, S., Cola, M., Knox, A., Pelella, M. R., Russell, A., Hauptmann, A., ... Schultz, R. T. (2023). Sex differences in the temporal dynamics of autistic children's natural conversations. *Molecular Autism*, *14*(1), 13. doi: https://doi.org/10.1186/s13229-023-00545-6

Cola, M., Zampella, C. J., Yankowitz, L. D., Plate, S., Petrulla, V., Tena, K., ... Parish-Morris, J. (2022). Conversational adaptation in children and teens with autism: Differences in talkativeness across contexts. *Autism Research*, *15*(6), 1090–1108.

Constantino, J. N. (2012). *Social responsiveness scale*. Western Psychological Services.

Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2023). A systematic review and bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, *7*(1), 114–133.

Cummins, F. (2019). *The ground from which we speak: Joint speech and the collective subject*. Cambridge Scholars Publishing.

Dale, R. (2015). An integrative research strategy for exploring synergies in natural language performance. *Ecological Psychology*, *27*(3), 190–201.

Dale, R., Fusaroli, R., Duran, N. D., & Richardson, D. C. (2013). The self-organization of human interaction. In *Psychology of learning and motivation* (Vol. 59, pp. 43–95). Elsevier.

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child development*, *73*(3), 700–717.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental psychology*, *40*(2), 271.

Denman, D., Speyer, R., Munro, N., Pearce, W. M., Chen, Y.-W., & Cordier, R. (2017). Psychometric properties of language assessments for children aged 4–12 years: A systematic review. *Frontiers in psychology*, *8*, 1515. doi: 10.3389/fpsyg.2017.01515

Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of mind: did evolution fool us? *PloS One*, *9*(2), e87619.

Dideriksen, C., Christiansen, M. H., Dingemanse, M., Højmark-Bertelsen, M., Johansson, C., Tylén, K., & Fusaroli, R. (2023). Language-specific constraints on conversation: Evidence from danish and norwegian. *Cognitive Science*, *47*(11), e13387.

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General*, *152*(3), 864.

Dideriksen, C. R., Fusaroli, R., Tylén, K., Dingemanse, M., & Christiansen, M. H. (2019). Contextualizing conversational strategies: Backchannel, repair and linguistics alignment in spontaneous and task-oriented conversations. In *The 41st annual meeting of the cognitive science society* (pp. 261–267).

Dilley, L., Lehet, M., Wieland, E. A., Arjmandi, M. K., Kondaurova, M., Wang, Y., ... Bergeson, T. (2020). Individual differences in mothers' spontaneous infant-directed speech predict language attainment in children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *63*(7), 2453–2467.

Dingemanse, M., & Liesenfeld, A. (2022). From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings of the 60th annual meeting of the association for computational linguistics* (pp. 5614–5633).

Doherty-Sneddon, G., Whittle, L., & Riby, D. M. (2013). Gaze aversion during social style interactions in autism spectrum disorder and williams syndrome. *Research in developmental disabilities*, *34*(1), 616–626.

Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Development*, *92*(2), 609–625.

Dunn, J., & Shatz, M. (1989). Becoming a conversationalist despite (or because of) having an older sibling. *Child development*, 399–410.

Ervin-Tripp, S. (1979). Children's verbal turn-taking. *Developmental pragmatics*, 391–414.

Ervin-Tripp, S. (1982). Ask and it shall be given unto you: Children's requests. *Contemporary Perceptions of Language: Interdisciplinary Dimensions*, 235–245.

Farmer, C., Adedipe, D., Bal, V., Chlebowski, C., & Thurm, A. (2020). Concordance of the vineland adaptive behavior scales, second and third editions. *Journal of intellectual disability research*, *64*(1), 18–26.

Fay, N., Walker, B., Swoboda, N., Umata, I., Fukaya, T., Katagiri, Y., & Garrod, S. (2018). Universal principles of human communication: Preliminary evidence from a cross-cultural communication game. *Cognitive Science*, *42*(7), 2397–2413.

Fay, W., & Schuler, A. L. (1980). *Emerging language in autistic children.* University Park Press.

Ford, C. E., Fox, B. A., & Thompson, S. A. (1996). Practices in the construction of turns: The "tcu" revisited. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *6*(3), 427–454.

Forgeot d'Arc, B., Devaine, M., & Daunizeau, J. (2020). Social behavioural adaptation in autism. *PLoS computational biology*, *16*(3), e1007700.

Fosnot, S. M., & Jun, S. (1999). Prosodic characteristics in children with stuttering or autism during reading and imitation. In *Proceedings of the 14th international congress of phonetic sciences* (pp. 1925–1928).

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition*, *50*(1-3), 115–132.

Fusaroli, R., Cox, C. M. M., Weed, E., Szabó, B. I., Fein, D., & Naigles, L. (2024). The development of turn-taking skills in typical development and autism. *PsyArxiv*. doi: https://doi.org/10.31234/osf.io/5ap6u

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2022). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in american and danish autistic children. *Autism Research*, *15*(4), 653–664.

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). Is voice a marker for autism spectrum disorder? a systematic review and meta-analysis. *Autism Research*, *10*(3), 384–407.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.

Fusaroli, R., & Tylén, K. (2012). Carving language for social coordination: A dynamical approach. *Interaction studies*, *13*(1), 103–124.

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive science*, *40*(1), 145–171.

Fusaroli, R., Weed, E., Fein, D., & Naigles, L. (2019). Hearing me hearing you: Reciprocal effects between child and parent language in autism and typical development. *Cognition*, *183*, 1–18.

Fusaroli, R., Weed, E., Rocca, R., Fein, D., & Naigles, L. (2023). Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development. *Cognition*, *236*, 105422.

Garvey, C. (1975). Requests and responses in children's speech. *Journal of child language*, *2*(1), 41–63.

Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse processes*, *4*(1), 27–57.

Geurts, H. M., van den Bergh, S. F., & Ruzzano, L. (2014). Prepotent response inhibition and interference control in autism spectrum disorders: Two meta-analyses. *Autism Research*, *7*(4), 407–420.

Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental review*, *22*(4), 555–592.

Heeman, P. A., Lunsford, R., Selfridge, E., Black, L. M., & Van Santen, J. (2010). Autism and interactional aspects of dialogue. In *Proceedings of the SIGDIAL 2010 conference* (pp. 249–252).

Heesen, R., & Fröhlich, M. (2022). *Revisiting the human 'interaction engine': comparative approaches to social action coordination* (Vol. 377) (No. 1859). The Royal Society.

Hilbrink, E., Casillas, M., & Lammertink, I. (2015). Twelve-month-olds differentiate between typical and atypical turn-timing in conversation. In *Workshop on infant language development (wild).*

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in cognitive sciences*, *8*(1), 26–32.

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., ... others (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, *6*(11), 1545–1556.

Hlavatá, P., Kašpárek, T., Linhartová, P., Ošlejšková, H., & Bareš, M. (2018). Autism, impulsivity and inhibition a review of the literature. *Basal Ganglia*, *14*, 44–53.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*(4), 477–492.

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of child language*, *37*(2), 229–261.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the society for research in child development*, *66*(2), 1–149.

Jardat, P., Calandreau, L., Ferreira, V., Gouyet, C., Parias, C., Reigner, F., & Lansade, L. (2022). Pet-directed speech improves horses' attention toward humans. *Scientific Reports*, *12*(1), 4297.

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental science*, *17*(3), 388–395.

Kenworthy, L., Case, L., Harms, M. B., Martin, A., & Wallace, G. L. (2010). Adaptive behavior ratings correlate with symptomatology and iq among individuals with high-functioning autism spectrum disorders. *Journal of autism and developmental disorders*, *40*, 416–423.

Kimhi, Y., Shoam-Kugelmas, D., Agam Ben-Artzi, G., Ben-Moshe, I., & Bauminger-Zviely, N. (2014). Theory of mind and executive function in preschoolers with typical development versus intellectually able preschoolers with autism spectrum disorder. *Journal of autism and developmental disorders*, *44*, 2341–2354.

Kondaurova, M. V., Bergeson, T. R., & Xu, H. (2013). Age-related changes in prosodic features of maternal speech to prelingually deaf infants with cochlear implants. *Infancy*, *18*(5), 825–848.

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in human neuroscience*, *6*, 215.

Leezenbaum, N. B., Campbell, S. B., Butler, D., & Iverson, J. M. (2014). Maternal verbal responses to communication of infants at low and heightened risk of autism. *Autism*, *18*(6), 694–703.

Levelt, W. J. (1993). *Speaking: From intention to articulation.* MIT press.

Levinson, S. (2006). Cognition at the heart of human interaction. *Discourse studies*, *8*(1), 85–93.

Levinson, S. (2016). Turn-taking in human communication–origins and implications for language processing. *Trends in Cognitive Sciences*, *20*(1), 6-14.

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological science*, *30*(4), 504–515.

Local, J., & Walker, G. (2012). How phonetic features project more talk. *Journal of the International Phonetic Association*, *42*(3), 255–280.

Magyari, L., Bastiaansen, M. C., De Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. *Journal of Cognitive Neuroscience*, *26*(11), 2530–2539.

Magyari, L., & De Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in psychology*, *3*, 376.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, *38*(4), 325–350.

McCrimmon, A. W., & Smith, A. D. (2013). Review of the wechsler abbreviated scale of intelligence, second edition (wasi-ii). *Autism*, *31*(3), 337-341.

Nadig, A., Seth, S., & Sasson, M. (2015). Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology*, *6*, 1888.

Nadig, A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of autism and developmental disorders*, *42*, 499–511.

Neimy, H., Pelaez, M., Carrow, J., Monlux, K., & Tarbox, J. (2017). Infants at risk of autism and developmental disorders: Establishing early social skills. *Behavioral Development Bulletin*, *22*(1), 6.

Nguyen, V., Versyp, O., Cox, C., & Fusaroli, R. (2022). A systematic review and bayesian meta-analysis of the development of turn taking in adult–child vocal interactions. *Child Development*, *93*(4), 1181–1200.

Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PloS one*, *14*(12), e0225377.

Ochs, E., Kremer-Sadlik, T., Sirota, K. G., & Solomon, O. (2004). Autism and the social world: An anthropological perspective. *Discourse studies*, *6*(2), 147–183.

Olsen, K., & Tylén, K. (2023). On the social nature of abstraction: cognitive implications of interaction and diversity. *Philosophical Transactions of the Royal Society B*, *378*(1870), 20210361.

Panneton, R., Cristia, A., Taylor, C., & Christine, M. (2023). Positive valence contributes to hyperarticulation in maternal speech to infants and puppies. *Journal of Child Language*, 1–11.

Parish-Morris, J., Pallathra, A. A., Ferguson, E., Maddox, B. B., Pomykacz, A., Perez, L. S., … Brodkin, E. S. (2019). Adaptation to different communicative contexts: an eye tracking study of autistic adults. *Journal of neurodevelopmental disorders*, *11*, 1–10.

Piazza, G., Martin, C. D., & Kalashnikova, M. (2022). The acoustic features and didactic function of foreigner-directed speech: A scoping review. *Journal of Speech, Language, and Hearing Research*, *65*(8), 2896–2918.

Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press.

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B*, *285*(1880), 20180598.

Pouw, W., & Holler, J. (2022). Timing in conversation is dynamically adjusted turn by turn in dyadic telephone conversations. *Cognition*, *222*, 105015.

Pouw, W., Proksch, S., Drijvers, L., Gamba, M., Holler, J., Kello, C., … Wiggins, G. A. (2021). Multilevel rhythms in multimodal communication. *Philosophical Transactions of the Royal Society B*, *376*(1835), 20200334.

Ravignani, A., & de Reus, K. (2019). Modelling animal interactive rhythms in communication. *Evolutionary Bioinformatics*, *15*, 1176934318823558.

Reimers, N., & Gurevych, I. (2020, 11). Making monolingual sentence embeddings multilingual using knowledge distillation. In *Proceedings of the 2020 conference on empirical methods in natural language processing.* Association for Computational Linguistics. Retrieved from https://arxiv.org/abs/2004.09813

Ritwika, V., Pretzer, G. M., Mendoza, S., Shedd, C., Kello, C. T., Gopinathan, A., & Warlaumont, A. S. (2020). Exploratory dynamics of vocal foraging during infant-caregiver communication. *Scientific reports*, *10*(1), 10469.

Russell, J. E. (1997). *Autism as an executive disorder.* Oxford University Press.

Rybner, A., Jessen, E. T., Mortensen, M. D., Larsen, S. N., Grossman, R., Bilenberg, N., ... others (2022). Vocal markers of autism: Assessing the generalizability of machine learning models. *Autism Research*, *15*(6), 1018–1030.

S., S. S., V., C. D., & A., S. C. (2016). *Vineland adaptive behavior scales, third edition.* Pearson.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). Linguistic society of america. *Language*, *50*(1), 696–735.

Schaffer, D. (1983). The role of intonation as a cue to turn taking in conversation. *Journal of phonetics*, *11*(3), 243–257.

Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2018). Mediating parent learning to promote social communication for toddlers with autism: Effects from a randomized controlled trial. *Journal of autism and developmental disorders*, *48*, 853–867.

Schoenenberg, K., Raake, A., & Koeppe, J. (2014). Why are you so slow?–misattribution of transmission delay to attributes of the conversation partner at the far-end. *International journal of human-computer studies*, *72*(5), 477–487.

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (n.d.). Mpnet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*, *33*, 16857–16867.

Stefanatos, G. A., & Baron, I. S. (2011). The ontogenesis of language impairment in autism: a neuropsychological perspective. *Neuropsychology Review*, *21*, 252–270.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... Yoon, K.-E. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592.

Tamis-LeMonda, C. S., & Adolph, K. E. (2013). Social referencing in infant motor action. In B. D. Homer & C. S. Tamis-LeMonda (Eds.), *The development of social cognition and communication* (pp. 145–164). Psychology Press.

Templeton, E. M., Chang, L. J., Reynolds, E. A., Cone LeBeaumont, M. D., & Wheatley, T. (2022). Fast response times signal social connection in conversation. *Proceedings of the National Academy of Sciences*, *119*(4), e2116915119.

Templeton, E. M., Chang, L. J., Reynolds, E. A., Cone LeBeaumont, M. D., & Wheatley, T. (2023). Long gaps between turns are awkward for strangers but not for friends. *Philosophical Transactions of the Royal Society B*, *378*(1875), 20210471.

Templeton, E. M., & Wheatley, T. (2023). Listening fast and slow. *Current Opinion in Psychology*, *53*, 101658. Retrieved from `https://www.sciencedirect.com/science/article/pii/S2352250X23001033` doi: https://doi.org/10.1016/j.copsyc.2023.101658

Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental disorders*, *23*(2), 309–322.

Tylén, K., Fusaroli, R., Bundgaard, P. F., & Østergaard, S. (2013). Making sense together: A dynamical account of linguistic meaning-making. *Semiotica*, *2013*(194), 39–62.

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak e-ng-li-sh? a comparison of foreigner-and infant-directed speech. *Speech communication*, *49*(1), 2–7.

van Der Fels, I. M., Smith, J., De Bruijn, A. G., Bosker, R. J., Königs, M., Oosterlaan, J., ... Hartman, E. (2019). Relations between gross motor skills and executive functions, controlling for the role of information processing and lapses of attention in 8-10 year old children. *PloS one*, *14*(10), e0224219.

Verga, L., Kotz, S. A., & Ravignani, A. (2023). The evolution of social timing. *Physics of Life Reviews*, *46*, 131–151.

Waade, P. T., Enevoldsen, K. C., Vermillet, A.-Q., Simonsen, A., & Fusaroli, R. (2023). Introducing tomsup: Theory of mind simulations using python. *Behavior Research Methods*, *55*(5), 2197–2231.

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324.

Wechsler, D. (2008). Wechsler adult intelligence scale–fourth edition (wais-iv). *APA PsycTests*. doi: https://doi.org/10.1037/t15169-000

Wehrle, S., Cangemi, F., Janz, A., Vogeley, K., & Grice, M. (2023a). Turn-timing in conversations between autistic adults: Typical short-gap transitions are preferred, but not achieved instantly. *Plos one*, *18*(4), e0284029.

Wehrle, S., Cangemi, F., Janz, A., Vogeley, K., & Grice, M. (2023b). Turn-timing in conversations between autistic adults: Typical short-gap transitions are preferred, but not achieved instantly. *Plos one*, *18*(4), e0284029.

West, K. L., & Iverson, J. M. (2021). Communication changes when infants begin to walk. *Developmental science*, *24*(5), e13102.

Wiig, E. H., Secord, W. A., & Semel, E. (2013). *Clinical evaluation of language fundamentals: CELF-5*. Pearson.

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic bulletin & review*, *12*, 957–968.

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, e1.

Yerys, B. E., Wallace, G. L., Sokoloff, J. L., Shook, D. A., James, J. D., & Kenworthy, L. (2009). Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. *Autism Research*, *2*(6), 322–333.

Ying Sng, C., Carter, M., & Stephenson, J. (2018). A systematic review of the comparative pragmatic differences in conversational skills of individuals with autism. *Autism & Developmental Language Impairments*, *3*, 2396941518803806.

6. Supplementary Materials

The Supplementary Materials include results adult response latencies, test-re-test reliability models, control analyses and extra model results mentioned in the main manuscript as well as model information and quality checks:

- Results for Adults
  - Response Latencies
  - Individual Differences
  - Utterance Predictability
  - Turn-By-Turn Adjustment
- Test-Re-Test Reliability
- Control Analyses
  - Child Latency With No Overlaps
  - Adult Latency With No Overlaps
  - Child Latency Estimates Across Gender
  - Model Estimates With No Sigma and No Beta
  - Model Estimates With Sigma But No Beta
  - Overlaps according to Individual Differences
  - Individual Differences With No Overlaps
  - Predictability Model With No Backchannels
  - Predictability Model With Short Utterances Removed
  - Model With Surrogate Pairs
- Model Information and Quality Checks
  - Choice of Priors
  - Prior and Posterior Predictive Checks
  - Prior-Posterior Update Plots

6.1. Results on Variability for Response Latency

6.1.1 Child Standard Deviation

The standard deviation from group-level estimates was smaller in the autism group than in the typical development group across conditions (*asd* < *td*: -71 [-145, -2], ER = 21.6) and within each condition with the exception of the matching game (matching game, *asd* < *td*: 0.92 [-92, 93], ER = 0.9, parent conversations: *asd* < *td*: -72 [-176, 24], ER = 8, experimenter conversation: *asd* < *td*: -142 [-176, -8], ER = 24.5). Both the autism group and typical development group display greater degree of between-child variability when interacting with the experimenter than with their parents (Autism, *Experimenter* > *Parent*: 59 [-29, 150], ER = 6.4; Typical Development, *Experimenter* > *Parent*: 129 [-6.47, 281.64], ER = 15.9). The groups showed only slightly more pronounced differences in variability when in conversations with the experimenter compared to with their parents ($Experimenter_{asd-td} < Parent_{asd-td}$: -70 [-244, 96], ER = 2.9). The autism group showed no difference in between-child variability when comparing across the matching game and parent conversations (*Matching Game* < *Parent Conversation*: 17 [-61, 96], ER = 1.8), but the typical development group showed a higher group-level standard deviation in parent conversations compared to in the matching game (*Matching Game* < *Parent Conversation*: -56 [-157, 39], ER = 5). There was thus a greater difference in between-subjects variability across groups in the parent conversations compared to the matching game ($MatchingGame_{asd-td} < ParentConversation_{asd-td}$: -73 [-202, 51], ER = 5).

6.1.2 Change in Sigma With Visit

The variance component (i.e., sigma) exhibited a slight decrease with each visit for both the autism and typical development group when aggregated across conditions (Autism Group, $SigmaVisit_{asd} < 0$ on log scale: -0.01 [-0.02, 0.01], ER = 3.2; Typical Development, $SigmaVisit_{td} < 0$ on log scale: -0.01 [-0.03, 0.01], ER = 4.7). The autism group showed decreases in sigma in the matching game and experimenter conversations but an increase in variance in the parent conversations across visits (matching game, *asd* < *0*: -0.03 [-0.05, -0.01], ER = 60.7; parent conversations, *asd* > *0*: 0.02 [-0.01, 0.06], ER = 8.2; experimenter conversations *asd* < *0*: -0.02 [-0.04, 0], ER = 9.7). Similarly, the typical development group showed no clear decreases in the matching game or experimenter conversations but a strong decrease in the parent conversations (matching game, *td* < *0*: 0 [-0.02, 0.03], ER = 1.5; parent conversations, *td* < *0*: -0.04 [-0.09, 0], ER = 15.8; experimenter conversations *td* > *0*: 0.01 [-0.01, 0.03], ER = 2.4). Sigma exhibited a similar magnitude of decrease for the typical development and autism group across conditions (*autism* > *typical development*: 0 [-0.02, 0.03], ER = 1.6); however, while the autism group showed a stronger decrease in sigma with visit in the matching game (*asd* < *td*: -0.03 [-0.07, 0], ER = 13.4) and experimenter conversations (*asd* < *td*: -0.02 [-0.06, 0.02], ER = 4.8), the typical development group exhibited a stronger decrease in sigma in parent conversations (*autism* > *typical development*: 0.02 [-0.02, 0.06], ER = 3.5). The autism group showed a greater decrease in sigma in the experimenter conversations than in the parent conversations (*Experimenter* < *Parent*: -0.04 [-0.08, 0], ER = 22.4). The opposite was true for the typical development group: a greater decrease in sigma in the parent rather than in the experimenter conversations across visits

(*Experimenter* > *Parent*: 0.05 [0, 0.1], ER = 20.6). There were stronger between-group differences in the extent to which sigma changed across visits in parent conversations than in experimenter conversations ($Experimenter_{asd-td} < Parent_{asd-td}$: -0.09 [-0.15, -0.03], ER = 150.5). The autism group showed similar decreases in sigma in the matching game and parent conversations (*Matching Game* > *Parent Conversations*: 0.01 [-0.02, 0.04], ER = 2.8) and the same was the case for the typical development group (*autism* > *typical development*: 0 [-0.03, 0.04], ER = 1.1), but there was evidence for a smaller between-group difference in the matching game compared to the parent conversations (*autism* > *typical development*: -0.1 [-0.17, -0.04], ER = 146.1).

6.1.3 Variability Across Visits

Variability across visits was similar for both groups when aggregated across conditions ($VisitSD_{asd} < VisitSD_{td}$: 11 [-82, 102], ER = 1.5) and within each condition (matching game: $SubjectSD_{asd} < SubjectSD_{td}$: -21 [-117, 76], ER = 1.9, parent conversation: $SubjectSD_{asd} < SubjectSD_{td}$: 7.32 [-175, 193], ER = 1.2, experimenter conversation: $SubjectSD_{asd} < SubjectSD_{td}$: 3.87 [-176, 170], ER = 0.9). The groups showed no difference in the strength of the difference in variability across visits when in conversations with the experimenter compared to with their parents ($SubjectSD_{asd-td} < SubjectSD_{asd-td}$: -3.45 [-277, 256], ER = 1) and no difference in the extent of the variability difference between parent conversations and matching game ($SubjectSD_{asd-td} < SubjectSD_{asd-td}$: 14 [-199, 225], ER = 1.3).

6.2. Results for Adults

6.2.1 Response Latencies

As shown in **Figure 6** and **Table 6**, adults in the autism group were only slightly faster than those in the typical development group both when aggregating across conditions ($asd < td$: -59ms [-196, 80], ER = 3.2) and considering each condition separately (matching game, $asd < td$: -64 [-239, 116], ER = 2.5; parent conversations, $asd < td$: -60 [-311, 183], ER = 2; experimenter conversations, $asd < td$: -52 [-321, 210], ER = 1.7). As shown in **Table 11** in the Supplementary Materials, when removing overlaps, adults in the autism group remained slightly faster than the typical development group ($asd < td$: -97ms [-182, -8], ER = 26.6). Parents in both the autism and typical development group exhibited slower response latencies compared to the experimenters (Autism, $Parent < Experimenter$: -78 [-303, 140], ER = 0.4; Typical Development, $Parent < Experimenter$: -86 [-364, 193], ER = 0.4), and there were stronger differences between the groups in parent conversations compared to experimenter conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: 8 [-352, 364], ER = 0.9). Parents in both groups also showed slightly slower response latencies in the matching game compared to in conversations (Autism, $Parent\ Conversation < Matching\ Game$: 50 [-138, 233], ER = 2; Typical Development, $Parent\ Conversation < Matching\ Game$: 54 [-180, 273], ER = 2), but here the extent of group difference remained comparable across the two conditions $ParentConversation_{asd-td} < MatchingGame_{asd-td}$: -3.99 [-293, 291], ER = 0.9).

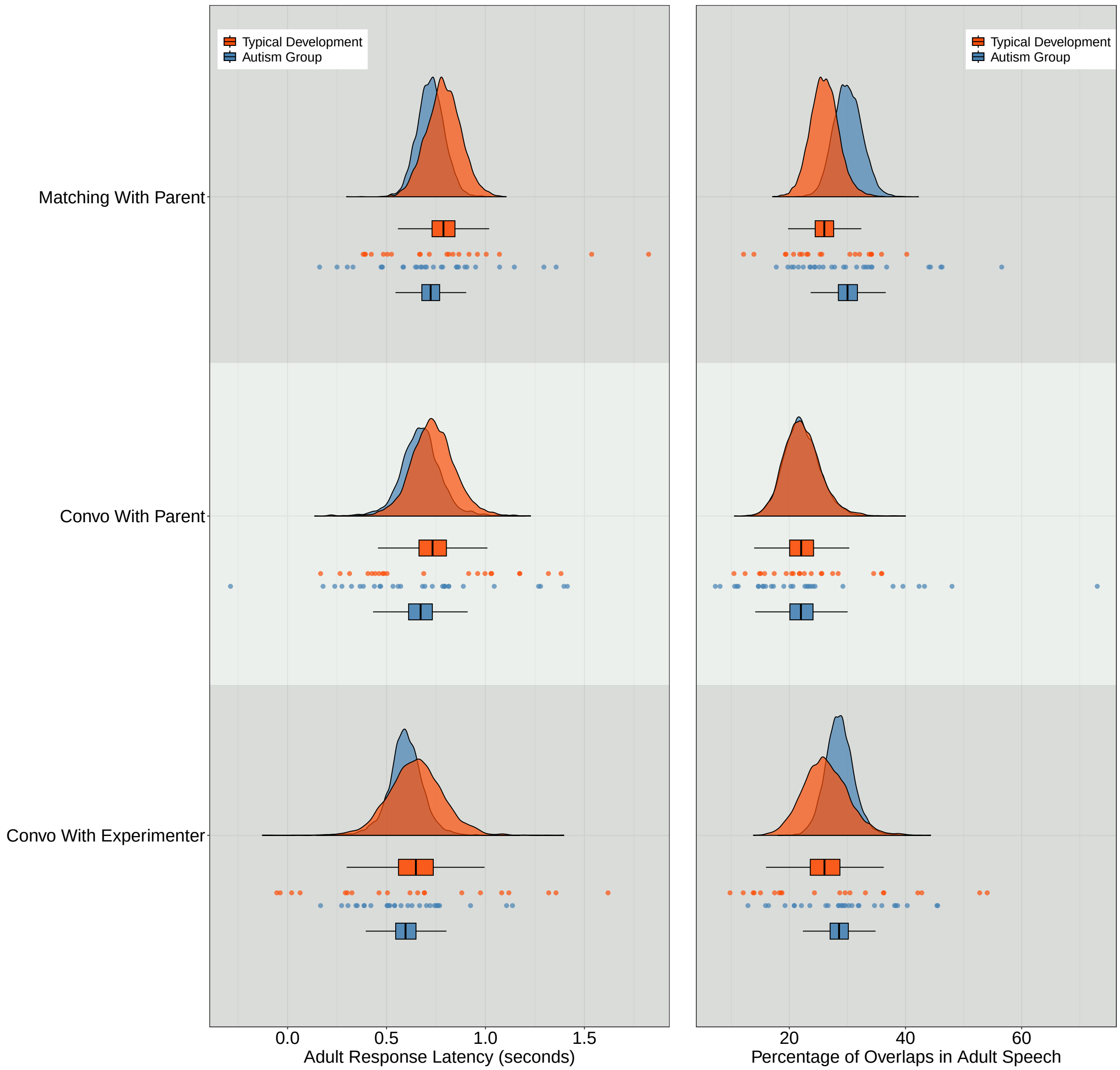

Figure 6: *Panel of plots to demonstrate model estimates for adult response latencies across different conversational contexts for the two diagnostic groups: Autism Group in blue and Typical Development Group in orange. The small points are posterior estimates for each individual participant. The density plots and boxplots show posterior predictions for the conversational activities.*

We observe a higher proportion of overlaps in adults' speech when interacting with autistic children when aggregated across tasks (*autism* > *typical development*: 0.02 [-0.02, 0.06], ER = 4), and when considering each condition separately (matching game, *autism* > *typical development*: 0.04 [-0.02, 0.1], ER = 7.5; parent conversations, *autism* > *typical development*: 0 [-0.07, 0.07], ER = 1; experimenter conversations (*autism* > *typical development*: 0.02 [-0.05, 0.1], ER = 2.5). There were strong differences in the proportions of overlaps produced by parents and experimenters in both the autism group and typical development group, with experimenters producing more overlaps (Autism Group, *Experimenter Conversation*

> *Parent Conversation*: 0.08 [0.02, 0.14], ER = 0; Typical Development, *Experimenter Conversation* > *Parent Conversation*: -0.04 [-0.12, 0.04], ER = 4.5); however, there were no changes in the strength of the difference between groups when comparing across parent and experimenter conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: 0.02 [-0.07, 0.12], ER = 0.5). There was a higher proportion of overlaps in the matching game than the parent conversations in both the autism group (*Matching Game* > *Parent Conversation*: 0.08 [0.02, 0.14], ER = 0) and the typical development group (*Matching Game* > *Parent Conversation*: 0.04 [-0.03, 0.1], ER = 5.6). The between-group difference in adults' overlap proportion was more pronounced in the matching game compared to parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: 0.04 [-0.04, 0.13], ER = 0.3).

Table 6: *Posterior estimates for child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency), exponential component (Beta), Residual Heterogeneity (Sigma), between-subjects standard deviation (Child SD), standard deviation across visits (Visit SD), Residual Heterogeneity across visits (Sigma Visit) and proportion of latencies below zero (Overlap Proportion). All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	724ms [613, 837]	788ms [642, 930]
	Convo With Parent	674ms [516, 837]	734ms [556, 919]
	Convo With Experimenter	597ms [450, 738]	648ms [430, 872]
	Aggregate Estimate	665ms [578, 754]	724ms [614, 833]
Overlap Proportion	Matching With Parent	0.3 [0.26, 0.34]	0.26 [0.22, 0.3]
	Convo With Parent	0.22 [0.18, 0.27]	0.22 [0.17, 0.28]
	Convo With Experimenter	0.29 [0.25, 0.33]	0.26 [0.2, 0.33]
	Aggregate Estimate	0.27 [0.24, 0.3]	0.25 [0.22, 0.28]
Sigma	Matching With Parent	506 [476, 539]	487 [459, 516]
	Convo With Parent	438 [389, 491]	526 [459, 604]
	Convo With Experimenter	861 [781, 945]	801 [725, 887]
	Aggregate Estimate	576ms [546, 608]	590ms [554, 627]
Beta	Matching With Parent	0.18 [0.1, 0.27]	0.15 [0.05, 0.26]
	Convo With Parent	-0.16 [-0.29, -0.04]	-0.14 [-0.32, 0.04]
	Convo With Experimenter	-0.44 [-0.59, -0.3]	-0.34 [-0.48, -0.2]
	Aggregate Estimate	-0.14 [-0.21, -0.07]	-0.11 [-0.2, -0.02]
Adult SD	Matching With Parent	280ms [220, 350]	347ms [264, 448]
	Convo With Parent	360ms [289, 447]	353ms [269, 458]
	Convo With Experimenter	251ms [194, 322]	475ms [372, 607]
	Aggregate Estimate	297ms [258, 343]	392ms [334, 455]
Visit SD	Matching With Parent	96ms [44, 178]	82ms [29, 159]
	Parent Conversation	98ms [9, 280]	90ms [7, 262]
	Convo With Experimenter	129ms [42, 285]	138ms [43, 308]
	Aggregate Estimate	108ms [54, 192]	103ms [49, 185]

There was evidence of a slightly higher degree of variance (i.e., sigma) for the parents in the typical development group compared to the autism group across conditions (*autism > typical development* on log scale: -0.02 [-0.11, 0.06], ER = 0.5). There was a higher degree of variability in the adults' response latencies for the autism group compared to in typical development group in the matching game (*autism > typical development* on log scale: 0.04 [-0.05, 0.13], ER = 3.4) and experimenter conversations (*autism > typical development* on log scale: 0.07 [-0.07, 0.21], ER = 4.2), but parents of autistic children exhibited less variability than those in typical development group in parent conversations (*autism > typical development* on log scale: -0.18 [-0.37, 0], ER = 19.2). Experimenters exhibited a substantially higher variance component than parents in both groups (Autism Group, *Experimenter > Parent*: on log scale: 0.68 [0.52, 0.83], ER = Inf; Typical Development, *Experimenter > Parent* on log scale: 0.42 [0.25, 0.59], ER = 2499). There was evidence for a more pronounced between-group difference in the extent of variance when comparing experimenter with parent conversations ($Experimenter_{asd-td} > Parent_{asd-td}$ on log scale: 0.25 [0.02, 0.48], ER = 26.3). Parents in the autism group showed a greater amount of variance than those in the matching game (*MatchingGame > Parent Conversation* on log scale: 0.15 [0.02, 0.28], ER = 27.7); however, adults in the typical development group showed the opposite pattern of higher variance in the parent conversations compared to in the matching game (*Matching Game > Parent Conversation* on log scale: -0.08 [-0.22, 0.07], ER = 4.1). There was evidence for a more pronounced between-group difference in the extent of variance in the matching game compared to in the parent conversations ($MatchingGame_{asd-td} < ParentConversation_{asd-td}$ on log scale: 0.22 [0.02, 0.42], ER = 25.3).

Adults' long pauses (i.e., the exponential tail) were slightly shorter in the autism group overall (*asd < td* on a log scale: -0.03 [-0.14, 0.08], ER = 2.1) and within each condition with the exception of the matching game (matching game, *asd < td*: 0.03 [-0.1, 0.16], ER = 0.6; parent conversations, *asd < td*: -0.03 [-0.25, 0.19], ER = 1.4; experimenter conversations, *asd < td*: -0.1 [-0.3, 0.1], ER = 3.9). The experimenter in both the autism group and typical development group showed shorter pauses (i.e., more left skew) compared to parents (Autism Group, *Experimenter > Parent*: -0.28 [-0.48, -0.08], ER = 0; Typical Development, *Experimenter > Parent*: -0.2 [-0.43, 0.02], ER = 0.1). The extent of between-group difference in long tails, however, did not differ in interactions with the experimenter compared to their parents ($Parent_{asd-td} < Experimenter_{asd-td}$: -0.07 [-0.38, 0.22], ER = 1.9). Parents in both the autism group and typical development showed longer pauses in the matching game compared to in parent conversations (Autism Group, *MatchingGame < Parent*: 0.34 [0.2, 0.49], ER = 999; Typical Development, *MatchingGame < Parent Conversations*: 0.29 [0.09, 0.5], ER = 127.2). The between-group difference was only slightly more pronounced in the parent conversations compared to the matching game ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: -0.06 [-0.31, 0.21], ER = 0.5).

The adults' standard deviation from group-level estimates was smaller in the autism group than in the typical development group across conditions (*asd < td*: -94 [-169, -22], ER = 52.2) and within each condition with the exception of parent conversations (matching game, *asd < td*: -67 [-187, 42], ER = 5, parent conversations: *asd < td*: 7 [-119, 130], ER = 0.9, experimenter conversation: *asd < td*: -223 [-119, -97], ER = 453.5). Parents in the autism group displayed greater degree of between-adult variability than experimenters, whereas experimenters in the typical development group displayed a greater degree of between-adult variability than parents (Autism, *Experimenter > Parent*: -108 [-214, -10], ER = 0; Typical Development, *Experimenter > Parent*: 122 [-24, 274], ER = 11). The groups showed more pronounced differences in between-adult variability when in conversations with the experimenter compared to with their parents ($Experimenter_{asd-td} < Parent_{asd-td}$: -230 [-410, -49], ER = 47.5). Parents in the autism group showed more between-adult variability in parent conversations than in the matching game (*Matching Game < Parent Conversation*: -80 [-183, 19], ER = 0.1), but parents in the typical development group showed no difference across the conversations and matching game (*Matching Game < Parent Conversation*: -6 [-144, 126], ER = 1.2). There was only a slightly greater difference in between-subjects variability across groups in the matching game compared to parent conversations ($MatchingGame_{asd-td} < ParentConversation_{asd-td}$: 74 [-95, 246], ER = 0.3).

6.2.2 Individual Differences

As shown in **Figure 7**, adults interacting with children with higher cognitive skills displayed faster latencies in both the autism group (*asd < 0*: -149ms [-254, -45], ER = 97) and the typical development group (*td < 0*: -358ms [-707, -13], ER = 21.6). There were robust between-group differences in children's cognitive skills, with a speedup of higher magnitude for adults interacting with children in the typical development group than in the autism group when aggregated across conditions (*autism > typical development*: 208ms [-155, 574], ER = 4.8). This difference in the effect of children's cognitive skills on adult response latencies applied in the matching game (*autism > typical development*: 177ms [-306, 666], ER = 2.8), experimenter conversations (*autism > typical development*: 233ms [-541, 983], ER = 2.4) and in parent conversations (*autism > typical development*: 215ms [-351, 806], ER = 2.7). Experimenters interacting with both autistic and typically developing children showed slightly weaker effects of children's cognitive skills than parents interacting with them (Autism Group, *Experimenter > Parent*: 209ms [-61, 479], ER = 9.5; Typical Development, *Experimenter > Parent*: 192ms [-712, 1120], ER = 0.6). The extent of between-group difference in the effect of cognitive skills on adult response latencies was similar in experimenter conversations and parent conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: 17ms [-932, 946], ER = 1.1). Parents interacting with both autistic and typically developing children with higher cognitive skills showed a speedup of higher magnitude in parent conversations than in the matching game (Autism Group, *Parent Conversations > Matching Game*: 58ms [-185, 294], ER = 0.5; Typical Development, *Parent Conversations > Matching Game*: 96ms [-602, 746], ER = 0.7). The extent of between-group difference similar across the matching game and parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: -38ms [-754, 694], ER = 0.8).

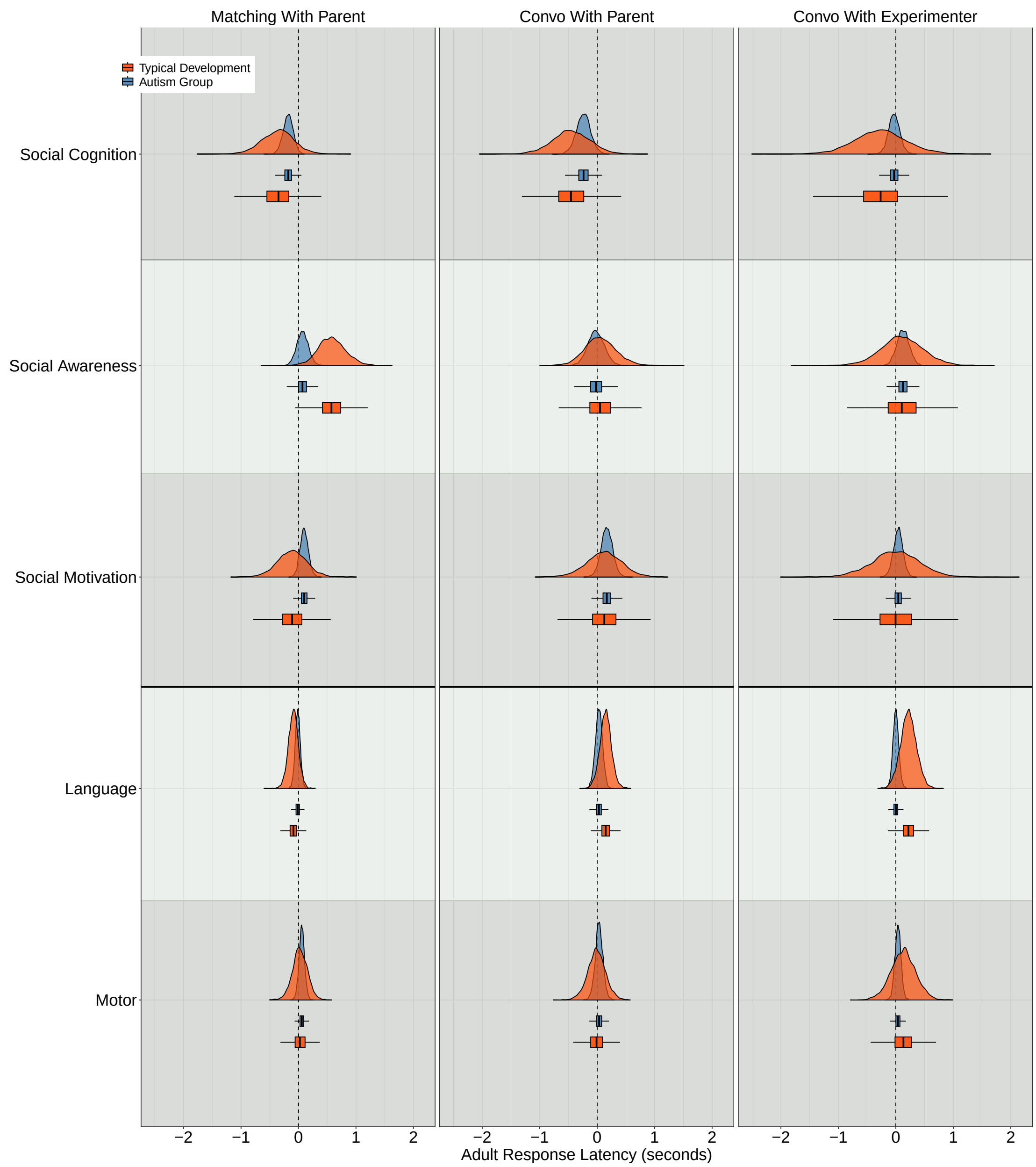

Figure 7: *Panel of plots to demonstrate model estimates for adult response latencies across individual differences and conversational contexts.*

There was evidence that adults interacting with children with higher motor skills had slower response latencies in the autism group (*asd* < *0*: 43ms [-14, 100], ER = 0.1) and weaker evidence for the same pattern in the typical development group (*td* < *0*: 48ms [-114, 212], ER = 0.4). There was no evidence of between-group differences in the effect of motor skills on adult response latencies when aggregated across conditions (*asd* < *tdc*: -6ms [-177, 166], ER = 0.9) and within each condition separately (Matching game, *asd* < *tdc*: 32ms [-199, 264], ER = 1.5; Parent Conversations: *asd* < *tdc*: 44ms [-235, 335], ER = 0.7; Experimenter Conversations: *asd* < *tdc*: -93ms [-467, 278], ER = 0.5). There was no evidence that adults interacting with children with higher motor skills slowed down their latencies less when interacting with the experimenter compared to their parent in both the autism and typical development group (*Autism Group, Parent < Experimenter*: 2.25ms [-145, 148], ER = 1; *Typical Development, Parent < Experimenter*: 36ms [-287, 356], ER = 1.4). The extent of between-group difference in the effect of motor skills on adult response latencies did not change across conversations with the experimenter and parents ($Parent_{asd-td}$ < $Experimenter_{asd-td}$: -137ms [-603, 327], ER = 0.4). There was also no evidence that adults interacting children with higher motor skills slowed down their latencies more when interacting their parents in the matching game than in the conversation in the autism group (*Autism Group, Parent Conversations < Matching Game*: 24ms [-105, 155], ER = 1.7) and the typical development group (*Typical Development, Parent Conversations < Experimenter Conversations*: 139ms [-297, 579], ER = 2.4). The extent of between-group difference in the effect on adult response latencies remained similar in the matching game compared to in parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$:-12ms [-363, 345], ER = 0.9).

Parents interacting with children with higher linguistic skills showed no clear differences in the autism group (*asd > 0*: 6ms [-48, 60], ER = 1.4) anda slowdown in response times in the typical development group (*td > 0*: 92ms [-17, 201], ER = 11). There was evidence for an effect of between-group differences for the effect of linguistic skills on adult response latencies when aggregated across conditions (*asd < td*: -87ms [-209, 34], ER = 0.1) and within each condition separately (matching game, *asd < tdc*: 77ms [-85, 239], ER = 3.7; parent conversations: *asd < tdc*: -117ms [-315, 82], ER = 0.2; experimenter conversations: *asd < tdc*: -220ms [-465, 25], ER = 0.1). There was weak evidence that experimenters interacting with children with higher linguistic skills slowed down their latencies less compared to parents in the autism group, but the opposite pattern was true for the typical development group (*Autism Group, Parent < Experimenter*: -31ms [-164, 107], ER = 0.5; *Typical Development, Parent Conversations < Experimenter Conversations*: 73ms [-201, 351], ER = 2.1). The extent of between-group difference in the effect of linguistic skills on adult response latencies did not change across the experimenter conversations and parent conversations ($Parent_{asd-td}$ < $Experimenter_{asd-td}$: -103ms [-410, 208], ER = 0.4). There was no evidence that parents interacting with children with higher linguistic skills slowed down their latencies more in the matching game than in the conversation in the autism group, but there was evidence for this greater degree of slowdown in the conversations in the typical development group (*Autism Group, Parent Conversations < Matching Game*: -43ms [-160, 77], ER = 0.4; *Typical Development, Parent Conversations < Experimenter Conversations*: -237ms [-449, -30], ER = 30.6). The extent of between-group difference was higher in the matching game compared to parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: 194ms [-39, 437], ER = 11).

Table 7: *Posterior estimates for how adult response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of each of the types of skills listed below (Motor, Cognitive Skills, Language Skills, Social Awareness and Social Motivation). The estimates denote the change in adult response latency as a function of one standard deviation increase in the skills of the children.*

	Skills	**Autism Group**	**Typical Development**
Motor	Matching With Parent	58ms [-24, 140]	26ms [-194, 240]
	Parent Conversations	34ms [-80, 145]	-10ms [-274, 250]
	Convo With Experimenter	36ms [-52, 122]	129ms [-229, 490]
	Aggregate Estimate	43ms [-14, 100]	48ms [-114, 212]
Cognition	Matching With Parent	-181ms [-329, -40]	-358ms [-834, 107]
	Convo With Parent	-239ms [-445, -36]	-454ms [-994, 85]
	Convo With Experimenter	-29ms [-187, 134]	-262ms [-1001, 500]
	Aggregate Estimate	-149ms [-254, -45]	-358ms [-707, -13]
Language	Matching With Parent	-12ms [-87, 64]	-90ms [-236, 54]
	Convo With Parent	30ms [-75, 133]	147ms [-22, 316]
	Convo With Experimenter	-0.11ms [-83, 85]	220ms [-12, 452]
	Aggregate Estimate	6ms [-48, 60]	92ms [-17, 201]
Awareness	Matching With Parent	71ms [-95, 241]	579ms [196, 993]
	Convo With Parent	-18ms [-252, 222]	58ms [-379, 532]
	Convo With Experimenter	124ms [-59, 307]	107ms [-501, 733]
	Aggregate Estimate	59ms [-55, 174]	248ms [-33, 554]
Motivation	Matching With Parent	99ms [-15, 221]	-113ms [-538, 299]
	Convo With Parent	169ms [2.16, 335.28]	118ms [-395, 619]
	Convo With Experimenter	40ms [-97, 174]	-8ms [-715, 664]
	Aggregate Estimate	103ms [20, 187]	-0.89ms [-319, 299]

Adults interacting with children with higher degrees of social awareness produced slower response latencies in the autism group (*asd > 0*: 59ms [-55, 174], ER = 4.1) and in the typical development group (*asd > 0*: 248ms [-33, 554], ER = 13.1). There was evidence of a bigger effect of social awareness scores on adult response latencies in the typical development group compared to the autism group when aggregated across conditions (*asd < td*: -189ms [-514, 122], ER = 5.6) and especially within the matching game but not parent or experimenter conversations (matching game, *asd < tdc*: -508ms [-957, -79], ER = 37.2; parent conversations: *asd < tdc*: -76ms [-601, 424], ER = 1.4; experimenter conversations: *asd < tdc*: 17ms [-629, 666], ER = 0.9). There was some evidence that experimenters interacting with autistic children with higher social awareness skills slowed down their latencies more than parents, but there was no evidence for this pattern for adults interacting with the typically developing children (*Autism Group, Parent < Experimenter*: 142ms [-162, 444], ER = 0.3; *Typical Development, Parent < Experimenter*: 49ms [-692, 764], ER = 1.2). The extent of between-group difference in the effect of social awareness skills on adult response latencies was not any stronger in the experimenter conversations compared to parent conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: 93ms [-693, 892], ER = 0.7). There was no evidence that parents interacting with autistic children with higher social awareness skills slow down their latencies more in the matching game than in the conversations, but there was evidence for this pattern in parents interacting with typically developing children (*Autism Group, Parent Conversations < Matching Game*: 90ms [-184, 370], ER = 0.4; *Typical Development, Parent Conversations < Experimenter Conversations*: 521ms [-61, 1098], ER = 14.1). The extent of between-group difference in the effect of awareness on adult response latencies was higher in the matching game compared to parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$:-432ms [-1074, 211], ER = 6.7).

Adults interacting with children who displayed higher social motivation scores produced slightly slower response times in the autism group (*asd > 0*: 103ms [20, 187], ER = 53.9) but similar response times in the typical development group (*td > 0*: -0.89ms [-319, 299], ER = 1). When comparing the estimates aggregated across conditions between the autism and typical development groups, the model showed weak evidence of robust between-group differences in the effect of social motivation on adult response latencies (*autism > typical development*: 104ms [-215, 431], ER = 2.3). Within each condition separately, all between-group comparisons showed no robust evidence of differences with the exception of the matching game where adults interacting with children with higher social motivation scores displayed a slowdown in the autism group compared to a speedup in the typical development group (matching game, *autism > typical development*: 212ms [-224, 653], ER = 3.8; parent conversations: *asd < tdc*: 50ms [-468, 594], ER = 0.8; experimenter conversations: *asd < tdc*: 48ms [-642, 759], ER = 0.8). There was evidence that experimenters interacting with children with higher social motivation skills in both the autism and typical development group slowed down their latencies less compared to parents (*Autism Group, Parent < Experimenter*: -128ms [-349, 91], ER = 0.2; *Typical Development, Parent < Experimenter*: -126ms [-978, 706], ER = 0.7). The extent of between-group difference in the effect of social motivation skills on parent response latencies was not stronger in the experimenter conversations compared to parent conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: -2.44ms [-886, 879], ER = 1). There was weak evidence that parents interacting with autistic children with higher social motivation skills slow down their latencies less in conversations than in the matching game (*Autism Group, Parent Conversations < Matching Game*: -69ms [-274, 130], ER = 0.4). Parents interacting with the typically developing children with higher social motivation skills, on the other hand, showed evidence of a speedup in the matching game compared to a slowdown in parent conversations (*Typical Development, Parent Conversations > Matching Game*: -231ms [-846, 420], ER = 2.8). There was evidence that the extent of between-group difference was slightly higher in the matching game compared to parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: 162ms [-514, 810], ER = 2).

6.2.3 Utterance Predictability

As shown in **Table 8**, the model indicates a influence of the predictability of the child's previous utterance on response latencies in the autism group but not the typical development group (Autism Group, *asd < 0*: -15ms [-35, 6], ER = 7.3; Typical Development, *td < 0*: 4.66ms [-19, 29], ER = 0.6). There is evidence of between-group differences in the influence of the predictability of previous child utterances on adult response latencies when aggregated across all three conditions (*autism > typical development*: -19ms [-51, 13], ER = 0.2). Within conditions, parents in the autism group exhibited a steeper negative slope in the matching game than the typical development group (*asd < td*: -34ms [-78, 9], ER = 9.5), in the parentt conversations (*autism > typical development*: -21ms [-77, 36], ER = 0.4) but not in experimenter conversations (*autism > typical development*: -2.46ms [-65, 61], ER = 0.9). Experimenters in both the autism and typical development group exhibited stronger effects of predictability than the parents in conversations (Autism Group, *Experimenter > Parent*: 26ms [-26, 80], ER = 0.3; Typical Development, *Experimenter > Parent*: 45ms [-21, 111], ER = 0.1). The between-group differences were slightly more pronounced in parent conversations than in experimenter conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: -19ms [-103, 68], ER = 0.6). Parents showed slightly smaller effects in parent conversations than in the matching game (Autism Group, *Parent < Matching Game*: 13ms [-35, 62], ER = 2; Typical Development, *Parent < Matching Game*: 26ms [-26, 78], ER = 3.9). The difference between the slopes in the two groups was more pronounced in the matching game than in parent conversations ($ParentConversation_{asd-td} > MatchingGame_{asd-td}$: -13ms [-84, 59], ER = 1.6).

Table 8: *Posterior estimates for how adult response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous child utterance and adult utterance. The estimates denote the change in child response latency as a function of one standard deviation increase in the predictability of the utterance*

.

	Tasks	Autism Group	Typical Development
Predictability of Previous Child Utterance	Matching With Parent	2.46ms [-26, 32]	37ms [6, 71]
	Convo With Parent	-10ms [-48, 28]	11ms [-29, 52]
	Convo With Experimenter	-36ms [-73, 1]	-34ms [-85, 16]
	Aggregate Estimate	-15ms [-35, 6]	4.66ms [-19, 29]
Predictability of Own Utterance	Matching With Parent	-8ms [-29, 14]	-4.16ms [-28, 19]
	Convo With Parent	-14ms [-38, 10]	4.98ms [-25, 35]
	Convo With Experimenter	47ms [14, 80]	55ms [16, 96]
	Aggregate Estimate	8ms [-8, 24]	19ms [-0.35, 37.7]

The model showed a small but robust positive influence of the predictability of the adult's utterance on adult response latencies for both groups (Autism Group, *asd > 0*: 8ms [-8, 24], ER = 4.6; Typical Development, *td > 0*: 19ms [-0.35, 37.7], ER = 17.9). There is moderate evidence that the predictability of adults' utterances exerts a stronger positive influence on adult response latencies in the typical development group than in the autism group when aggregated across all three conditions (*autism > typical development*: -10ms [-35, 15], ER = 0.3). Parents in the autism group showed a slightly more negative slope in the matching game than those in the typical development group (*autism > typical development*: 5ms [-27, 38], ER = 0.6). In the parent conversations, parents in the autism group showed a slightly negative effect of predictability on their response times, while those in the typical development group showed a slightly positive effect (*autism > typical development*: -19ms [-57, 20], ER = 0.3). The experimenter in the autism group showed similar positive effects of predictability compared to in the typical development group (*autism > typical development*: -8ms [-59, 43], ER = 0.6). Compared to parents, the experimenter exhibited stronger positive effects of predictability on their own response latencies in both the autism and typical development group (Autism Group, *Experimenter > Parent*: -61ms [-102, -20], ER = 137.9; Typical Development, *Experimenter > Parent*: -50ms [-101, 0], ER = 19.3). There was no evidence that the between-group differences were more pronounced in parent conversations than in experimenter conversations ($Parent_{asd-td} < Experimenter_{asd-td}$: -10ms [-74, 52], ER = 1.5). Parents in the autism group showed a slightly less steep negative slope in the matching game compared to in conversations, whereas parents in the typical development group exhibited a slightly negative slope in the matching game versus a slightly positive slope in parent conversations (Autism Group, *Matching Game < Parent Conversation*: 5ms [-27, 38], ER = 0.6; Typical Development, *td < 0*: -9ms [-47, 30], ER = 0.5). The difference between the slopes in the two groups was slightly more pronounced in the matching game compared to in parent conversations ($ParentConversation_{asd-td} > MatchingGame_{asd-td}$: 15ms [-37, 66], ER = 0.5).

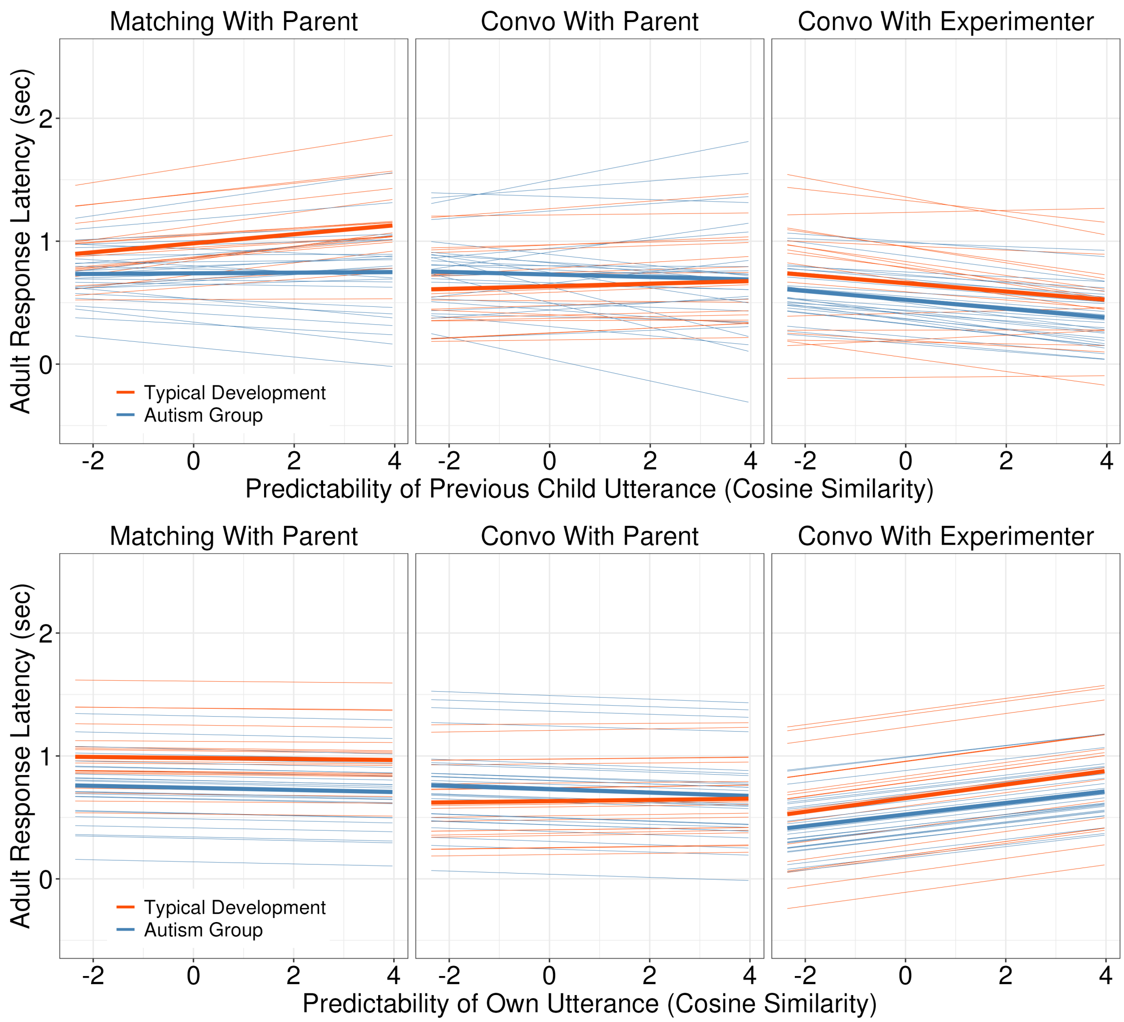

Figure 8: *Panel of plots to demonstrate how model estimates for adult response latencies change with the predictability of the previous child response latency (top) and semantic predictability of own utterance (bottom).*

6.2.4 Turn-By-Turn Adjustment

As shown in **Table 9** and **Figure 9**, the multivariate outcome model including interpersonal adjustment between child and adult showed that adults in the autism group and the typical development group produced longer response latencies the longer the response latency of the previous child's utterance (Autism Group, *asd > 0*: 33ms [14, 52], ER = 374; Typical Development, *td > 0*: 32ms [6, 58], ER = 43.1). When aggregated across conditions, there were no robust differences between adults in the autism and typical development group (*autism > typical development*: 1.58ms [-31, 34], ER = 1.1).

Parents in the autism group displayed a greater degree of interpersonal adjustment than those in the typical development group in the matching game, parents exhibited similar degrees of interpersonal adjustment across the groups in the parent conversations, and the experimenter exhibited a higher degree of interpersonal adjustment in the typical development group than in the autism group (matching game, *autism > typical development*: 22ms [-12, 56], ER = 6.7; parent conversations, *autism > typical development*: 1.08ms [-65, 67], ER = 1, experimenter conversations, *autism > typical development*: -19ms [-79, 43], ER = 0.4). There was some evidence that the experimenter adapted less to the response latencies of the children compared to the parents in the autism group, but this pattern was not present in the typical development group (Autism Group, *Experimenter < Parent*: 18ms [-36, 74], ER = 2.4; Typical Development, *Experimenter Conversations < Parent Conversations*: -1.31ms [-73, 70], ER = 0.9). There was no clear evidence for a greater between-group difference in interpersonal adjustment when children were engaging with their parents than with the experimenter ($Parent_{asd-td} > Experimenter_{asd-td}$: 20ms [-71, 111], ER = 1.8). There was also no clear evidence that parents adapted more to the response latencies of their children in the matching game compared to in parent conversations in neither the autism group or typical development group (Autism Group, $Parent_{asd-td} < Experimenter_{asd-td}$: 4.5ms [-36, 44], ER = 0.7; Typical Development, *Experimenter Conversations < Parent Conversations*: -17ms [-78, 43], ER = 0.5). There was some evidence for a slightly smaller between-group difference in parent conversations compared to the matching game ($ParentConversation_{asd-td} > MatchingGame_{asd-td}$: 21ms [-50, 93], ER = 0.5).

Table 9: *Posterior estimates for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the latency of the previous adult utterance (Interpersonal Adjustment) and previous child response latency (Self-Adjustment). The estimates denote the change in child response latency as a function of a one-second increase in the response latency of the utterance. The parameter estimates with an evidence ratio of above 10 for a test of difference to null are marked with a *.*

	Tasks	**Autism Group**	**Typical Development**
Interpersonal Adjustment	Matching With Parent	42ms [25, 60]	20ms [-8, 48]
	Convo With Parent	38ms [2.11, 75.53]	37ms [-18, 92]
	Convo With Experimenter	20ms [-21, 61]	38ms [-9, 83]
	Aggregate Estimate	33ms [14, 52]	32ms [6, 58]
Self Adjustment	Matching With Parent	41ms [23, 58]	34ms [9, 60]
	Convo With Parent	73ms [43, 103]	34ms [-2.33, 70.35]
	Convo With Experimenter	106ms [56, 155]	80ms [48, 111]
	Aggregate Estimate	73ms [53, 93]	49ms [31, 68]

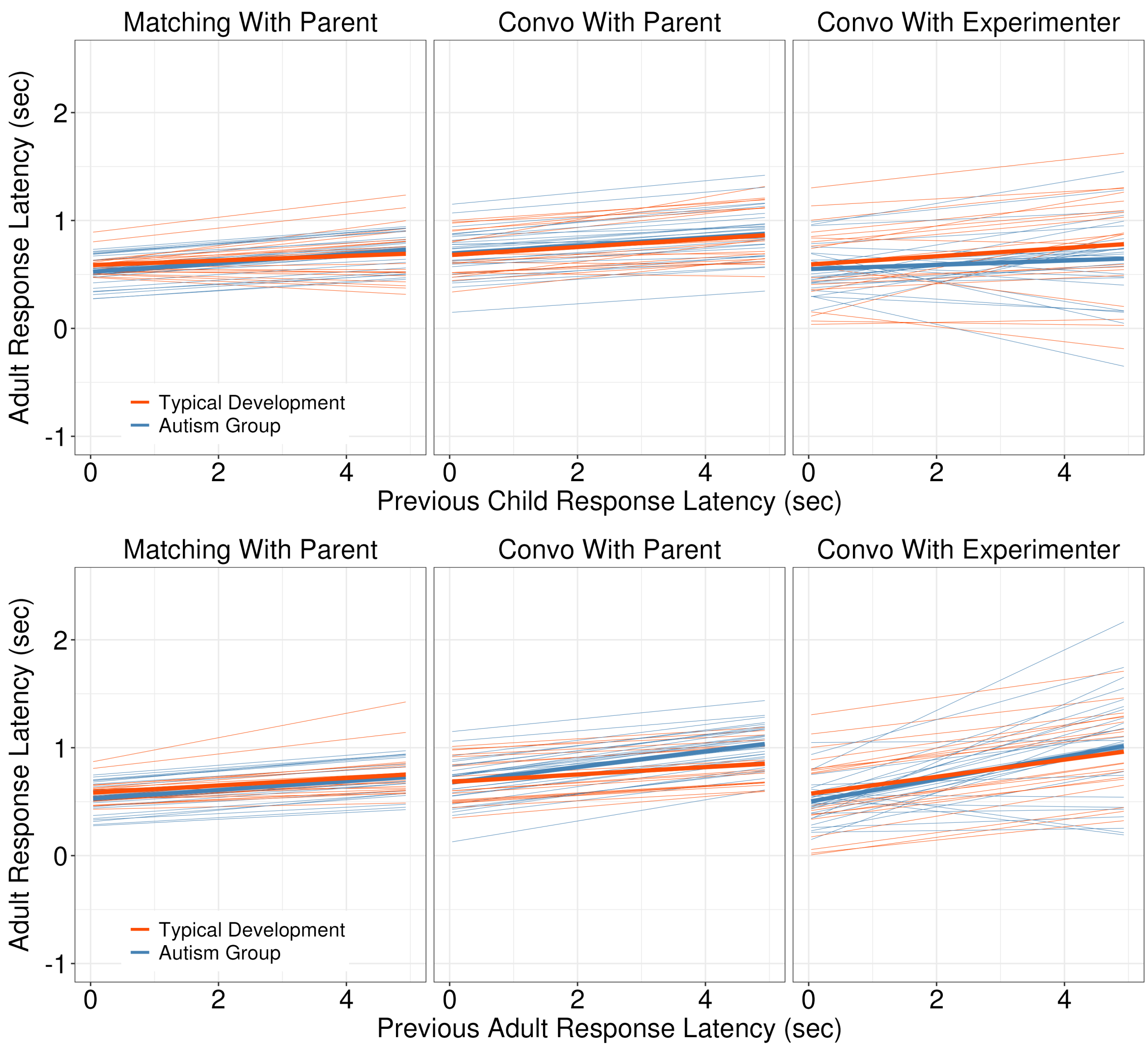

Figure 9: *Panel of plots to demonstrate how model estimates for adult response latencies change with the latency of the previous child response latency (top) and the previous adult response latency (bottom).*

The results for self-adjustment showed that adults in both the autism and typical development group adjusted their response latencies according to their previous response latency (Autism Group, *asd* $>$ *0*: 73ms [53, 93], ER = Inf; Typical Development, *td* $>$ *0*: 49ms [31, 68], ER = Inf). When aggregated across conditions, adults in the autism group displayed a greater degree of self-adjustment than the typical development group (*autism* $>$ *typical development*: 24ms [-2.91, 52.01], ER = 13.7). Within conditions, parents in the autism group displayed a greater degree of self-adjustment than the typical development group in the matching game (*autism* $>$ *typical development*: 7ms [-24, 37], ER = 0.5), parent conversations (*autism* $>$ *typical development*: 39ms [-9, 86], ER = 11.1) and experimenter conversations (*autism* $>$ *typical development*: 26ms [-34, 87], ER = 3.3). The experimenter exhibited a higher degree of self-adjustment than parents in both the autism and typical development group (Autism Group, *Experimenter Conversations* $<$ *Parent Conversations*: -33ms [-91, 25], ER = 0.2; Typical Development, *Experimenter Conversations* $<$ *Parent Conversations*: -46ms [-93, 2], ER = 0.1). There was no evidence for a greater between-group difference in adults' degree of self-adjustment in interactions with experimenters compared to parents ($Parent_{asd-td} < Experimenter_{asd-td}$: 13ms [-62, 88], ER = 0.6). There was evidence that parents in the autism group exhibited less self-adjustment in the matching game compared to parent conversations, whereas parents in the typical development group exhibited similar degrees of self-adjustment in the matching game and in conversations (Autism Group, *Matching Game* $<$ *Parent Conversations*: -32ms [-67, 2], ER = 14.7; Typical Development, *Matching Game* $<$ *Parent Conversations*: -0.04ms [-43, 43], ER = 1). There was thus a smaller between-group difference for the parents' self-adjustment in the parent conversations compared to in the matching game ($ParentConversation_{asd-td} > MatchingGame_{asd-td}$: -32ms [-88, 23], ER = 5).

6.3. Test-Re-Test Reliability

To investigate the test-retest reliability of our measures of turn-taking, we extracted estimates of mu, sigma and beta (mean and variance of the Gaussian component, exponential tail) by visit and diagnostic group. We then calculated a correlation matrix of the estimates within each diagnostic group across the 7 weeks and assessed both average correlation and the impact of distance on the correlation estimate. As shown in **Figure 10**, there was evidence of a high degree of test re-test reliability in both the autism group and typical development group in parent conversations (Autism Group: 0.83 [0.81, 0.86]; Typical Development: 0.81 [0.8, 0.82]) and experimenter conversations (Autism Group: 0.7 [0.66, 0.78]; Typical Development: 0.61 [0.48, 0.68]), but less so in the matching game (Autism Group: 0.42 [0.24, 0.48]; 0.28 [-0.08, 0.48]).

Figure 10: *Panel of plots showing strength of correlations among visits for child latencies across different conditions*

6.4. Control Analyses

6.4.1 Child Latency With No Overlaps

To check the extent to which the inclusion of overlaps influenced our estimates in the main manuscript, we ran control model without overlaps, the estimates for which are in **Table 10**. Child response latencies for the autism group remained faster than those for the typical development group both when aggregating across conditions ($asd < td$: -97ms [-182, -8], ER = 26.6) and considering each condition separately with the exception of parent conversations (matching game, $asd < td$: -90 [-193, 11], ER = 13.1; parent conversations, $asd < td$: -12 [-135, 110], ER = 1.3; experimenter conversations, $asd < td$: -189 [-347, -24], ER = 30.6). Both the autism and typical development group exhibited slower response latencies when engaging with the experimenter compared to their parent (Autism, $Parent < Experimenter$: 393 [275, 508], ER = Inf; Typical Development, $Parent < Experimenter$: 569 [402, 731], ER = Inf), and there were stronger differences between the groups when the children were conversing with the experimenter compared to with their respective parents ($Parent_{asd-td} < Experimenter_{asd-td}$: -177 [-379, 27], ER = 12). Both groups also showed slower response latencies when interacting with their parents in the matching game compared to in conversations (Autism, $ParentConversation < MatchingGame$: 365 [283, 450], ER = 4999; Typical Development, $ParentConversation < MatchingGame$: 443 [328, 566], ER = 4999), and the extent of group difference was stronger in parent conversations compared to experimenter conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: -78 [-226, 70], ER = 0.2).

Table 10: *Posterior estimates for child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency), exponential component (Beta), Residiual Heterogeneity (Heterogeneity), between-subjects heterogeneity (Child SD), heterogeneity across visits (Visit SD), and correlations among visit estimates (Correlations). All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	1238ms [1181, 1298]	1328ms [1246, 1414]
	Convo With Parent	873ms [808, 938]	885ms [788, 986]
	Convo With Experimenter	1265ms [1165, 1360]	1454ms [1327, 1581]
	Aggregate Estimate	1125ms [1077, 1174]	1222ms [1152, 1291]
Beta	Matching With Parent	0.2 [0.15, 0.25]	0.27 [0.2, 0.33]
	Convo With Parent	-0.17 [-0.24, -0.09]	-0.17 [-0.27, -0.06]
	Convo With Experimenter	-0.09 [-0.17, -0.01]	-0.07 [-0.19, 0.05]
	Aggregate Estimate	-0.02 [-0.06, 0.03]	0.01 [-0.05, 0.07]
Sigma	Matching With Parent	7 [3.76, 11.67]	6 [3.33, 10.76]
	Convo With Parent	10 [4.82, 20.03]	23 [10, 45]
	Convo With Experimenter	201 [163, 245]	326 [286, 372]
	Aggregate Estimate	24ms [18, 33]	36ms [25, 48]

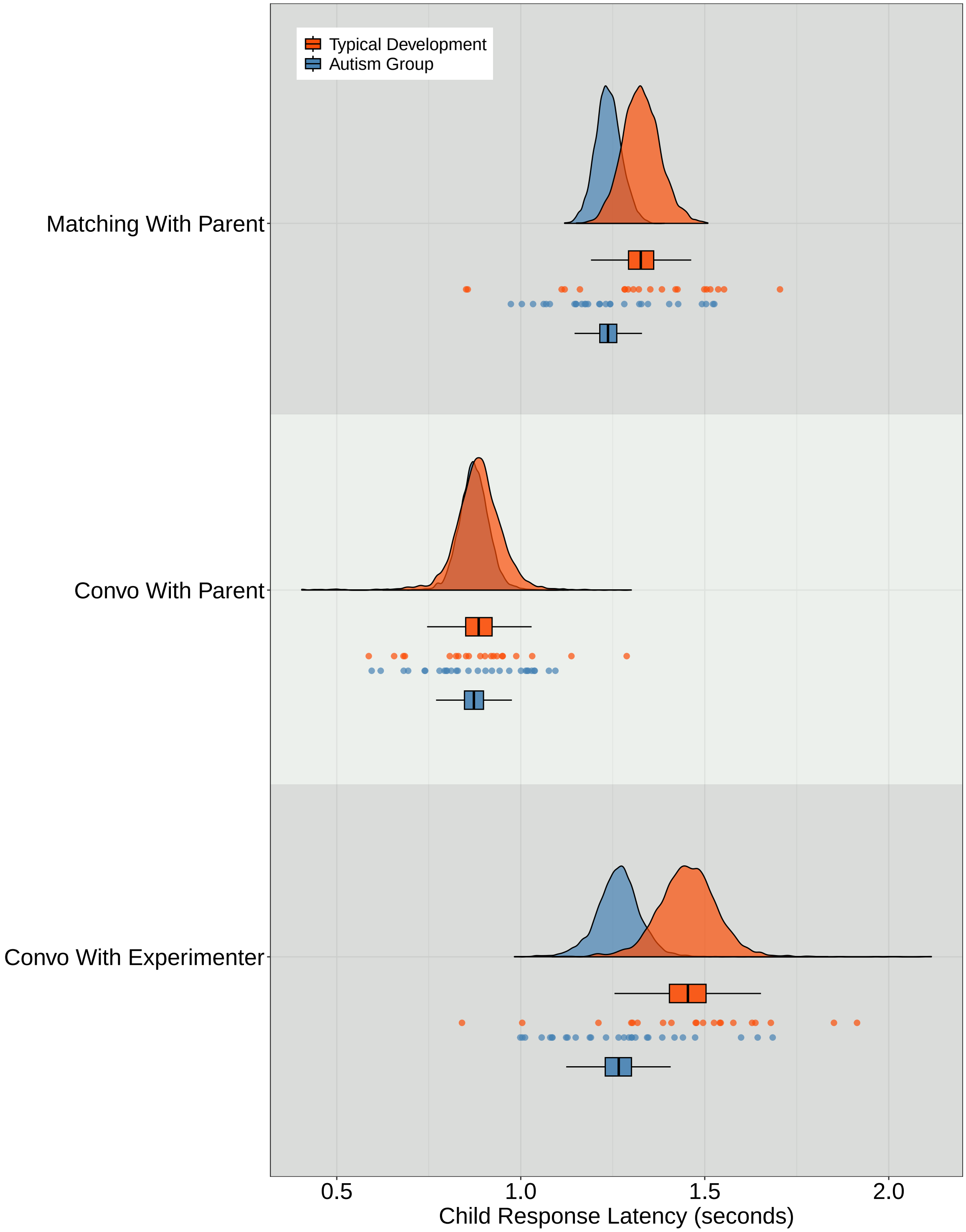

Figure 11: *Panel of plots showing child response latency estimates from a model where overlaps (i.e., negative latencies) were excluded.*

6.4.2 Adult Latency With No Overlaps

As shown in **Table 11**, adult response latencies for the autism group were only slightly faster than those for the typical development group when aggregating across conditions ($asd < td$: -29ms [-122, 62], ER = 2.2) and no clear patterns across each condition separately (matching game, $asd < td$: 2.68 [-121, 124], ER = 1.1; parent conversations, $asd < td$: -80 [-260, 77], ER = 3.9; experimenter conversations, $asd < td$: -9 [-185, 167], ER = 1.1). The experimenters exhibited slower response times in both the autism and typical development group compared to parents (Autism, $Parent < Experimenter$: 393 [275, 508], ER = Inf; Typical Development, $Parent < Experimenter$: 569 [402, 731], ER = Inf), but there were only slightly stronger differences between the groups when the children were conversing with the experimenter compared to with their respective parents ($Parent_{asd-td} < Experimenter_{asd-td}$: -177 [-379, 27], ER = 12). Both adult interactants also showed slower response latencies in the matching game compared to in conversations (Autism, $ParentConversation < MatchingGame$: 365 [283, 450], ER = 4999; Typical Development, $ParentConversation < MatchingGame$: 443 [328, 566], ER = 4999), but again the extent of group difference was only slightly stronger in the matching game compared to parent conversations ($ParentConversation_{asd-td} < MatchingGame_{asd-td}$: -78 [-226, 70], ER = 0.2).

Table 11: *Posterior estimates for child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency), exponential component (Beta), Residual Heterogeneity (Heterogeneity), between-subjects heterogeneity (Child SD), heterogeneity across visits (Visit SD), and correlations among visit estimates (Correlations). All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	Autism Group	Typical Development
Latency	Matching With Parent	1195ms [1135, 1257]	1192ms [1092, 1298]
	Convo With Parent	942ms [853, 1027]	1021ms [898, 1157]
	Convo With Experimenter	1135ms [1032, 1240]	1144ms [1009, 1279]
	Aggregate Estimate	1090ms [1036, 1142]	1119ms [1046, 1196]
Beta	Matching With Parent	0.17 [0.12, 0.22]	0.14 [0.05, 0.22]
	Convo With Parent	-0.09 [-0.16, -0.01]	-0.06 [-0.18, 0.07]
	Convo With Experimenter	-0.34 [-0.47, -0.21]	-0.31 [-0.43, -0.19]
	Aggregate Estimate	-0.08 [-0.14, -0.03]	-0.07 [-0.14, -0.01]
Sigma	Matching With Parent	6 [3.1, 9.5]	7 [3.73, 11.57]
	Convo With Parent	9 [4.62, 14.61]	15 [9, 26]
	Convo With Experimenter	249 [217, 284]	234 [197, 276]
	Aggregate Estimate	23ms [17, 30]	29ms [22, 37]

Figure 12: *Panel of plots showing child response latency estimates from a model where overlaps (i.e., negative latencies) were excluded.*

6.4.3 Child Latency Estimates Across Gender

As shown in **Table 12**, Child response latencies exhibited a difference across gender for the autism group when aggregated across conditions (*male > female*: 72 [-85, 228], ER = 3.5; *male > non-binary*: 152 [-129, 421], ER = 4.5; *female > non-binary*: 80 [-205, 363], ER = 2.1). There was evidence for similar differences within the matching game (*male > female*: 61 [-143, 272], ER = 2.2, *male > non-binary*: 152 [-188, 487], ER = 3.4, *female > non-binary*: 90 [-252, 437], ER = 1.9), parent conversations (*male > female*: 100 [-105, 306], ER = 3.9, *male > non-binary*: 209 [-160, 573], ER = 4.7, *female > non-binary*: 109 [-287, 485], ER = 2.2) and experimenter conversations (*male > female*: 55 [-184, 288], ER = 1.9, *male > non-binary*: 42 [-426, 495], ER = 1.3, *female > non-binary*: 42 [-426, 495], ER = 1.3). When aggregated across conditions in the typical development group, on the other hand, there were no clear differences in response latencies between genders (*male > female*: 37 [-156, 234], ER = 1.6). However, female children produced faster response latencies in the matching game (*male > female*: 138 [-100, 372], ER = 5.4) and parent conversations (*male > female*: 93 [-190, 372], ER = 2.5) but slower response latencies in experimenter conversations (*male < female*: -120 [-468, 225], ER = 2.6).

When comparing within genders and across diagnoses, the model showed evidence in line with the model in the main manuscript; that is, autistic male children exhibited faster response latencies compared to typically developing male children aggregated across conditions (*ASDmale < TDmale*: -129 [-308, 51], ER = 7.7) and within individual conditions (matching game, *ASDmale < TDmale*: -141 [-370, 93], ER = 5.7, parent conversations: *ASDmale < TDmale*: -98 [-335, 147], ER = 2.9, experimenter conversations, *ASDmale < TDmale*: -148 [-453, 154], ER = 3.8). Similarly, when comparing female children with and without autism, there was evidence of the autism group being faster, both when aggregated across conditions (*ASDfemale < TDfemale*: -164 [-337, 7], ER = 16.2) and within individual conditions (matching game, *ASDmale < TDmale*: -64 [-273, 149], ER = 2.3, parent conversations: *ASDmale < TDmale*: -105 [-343, 132], ER = 3.3, experimenter conversations, *ASDmale < TDmale*: -322 [-615, -30], ER = 28.1).

Autistic male children still demonstrated faster response times than typically developing female children when aggregated across conditions (*ASDmale < TDfemale*: -92 [-250, 66], ER = 4.9). However, because typically developing female children had faster response latencies than typically developing male children in the parent and experimenter conversations – and because autisic male children exhibited slower response latencies compared to autistic female children in the parent and experimenter conversations – there was no evidence of any difference between autistic male response latencies and typically developing female response latencies in these two conditions (matching game: *ASDmale < TDfemale*: -3.06 [-206, 199], ER = 1; parent conversations: -4.44 [-221, 210], ER = 1). When comparing autistic male children with typically developing female children in the experimenter conversations, however, there were clear differences between the groups (experimenter conversations, *ASDmale < TDfemale*: -268 [-535, -1], ER = 19.3).

Autistic female children also demonstrated faster response times than typically developing male children when aggregated across conditions (*ASDmale < TDfemale*: -201 [-398, -1], ER = 19.6) and within conditions (matching game: *ASDfemale < TDmale*: -203 [-441, 38], ER = 10.5; parent conversations: -198 [-472, 76], ER = 7.9; experimenter conversations, *ASDfemale*

$<$ *TDmale*: -202 [-536, 131], ER = 5.3).

Autistic non-binary children also demonstrated faster response times than both typically developing male and female children when aggregated across conditions (male, *ASDnonbinary* $<$ *TDmale*: -152 [-421, 129], ER = 4.5; female, *ASDnonbinary* $<$ *TDfemale*: -80 [-363, 205], ER = 2.1) and within conditions (matching game, *ASDnonbinary* $<$ *TDmale*: -293 [-657, 65], ER = 10.3; *ASDnonbinary* $<$ *TDfemale*: -155 [-504, 180], ER = 3.3, parent conversations: *ASDnonbinary* $<$ *TDmale*: -307 [-712, 112], ER = 8.2; *ASDnonbinary* $<$ *TDfemale*: -214 [-607, 179], ER = 4.5, experimenter conversations: *ASDnonbinary* $<$ *TDmale*: -244 [-734, 249], ER = 3.7; *ASDnonbinary* $<$ *TDfemale*: -364 [-833, 110], ER = 9).

Table 12: *Posterior estimates for child response latencies across genders and conditions for both diagnostic groups. All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Male	**Female**	**NonBinary**
ASD			
Matching With Parent	892ms [755, 1029]	830ms [679, 981]	740ms [428, 1041]
Convo With Parent	621ms [496, 748]	521ms [356, 690]	412ms [70, 773]
Convo With Experimenter	793ms [646, 938]	738ms [552, 926]	696ms [286, 1114]
Aggregate	769ms [672, 864]	697ms [577, 817]	616ms [357, 883]
TD			
Matching With Parent	1033ms [853, 1213]	895ms [748, 1040]	
Convo With Parent	719ms [505, 934]	626ms [454, 799]	
Convo With Experimenter	941ms [670, 1208]	1061ms [829, 1282]	
Aggregate	898ms [743, 1051]	860ms [737, 982]	

6.4.4 Model Estimates With No Sigma and No Beta

To explore how the estimates change with a vaying beta and varying sigma component, we ran a control model without the varying beta and varying sigma component. The posterior estimates from this model can be viewed in **Table 13** below.

Table 13: *Posterior estimates for child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters in the model without a varying beta and sigma: the Gaussian component (Latency), between-subjects standard deviation (Child SD), and standard deviation across visits (Visit SD). All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	785ms [704, 868]	813ms [729, 899]
	Convo With Parent	726ms [585, 867]	832ms [662, 1002]
	Convo With Experimenter	766ms [629, 909]	970ms [761, 1177]
	Aggregate Estimate	759ms [672, 844]	872ms [764, 981]
Child SD	Matching With Parent	204ms [157, 261]	195ms [137, 269]
	Convo With Parent	210ms [157, 277]	340ms [253, 447]
	Convo With Experimenter	352ms [277, 442]	460ms [354, 589]
	Aggregate Estimate	256ms [215, 303]	332ms [276, 395]
Visit SD	Matching With Parent	69ms [25, 134]	42ms [3.62, 100.02]
	Convo With Parent	115ms [16, 308]	78ms [4.7, 254.14]
	Convo With Experimenter	81ms [8, 214]	114ms [26, 273]
	Aggregate Estimate	88ms [37, 172]	78ms [30, 158]

6.4.5 Model Estimates With Sigma But No Beta

To explore how the estimates change with a vaying beta component, we ran a control model without the varying beta component but retained the varying sigma component. The posterior estimates from this model can be viewed in **Table 14** below.

Table 14: *Posterior estimates for child response latencies across individual conditions (Matching Game, Convo With Parent and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters in the model without a varying beta and sigma: the Gaussian component (Latency), between-subjects standard deviation (Child SD), and standard deviation across visits (Visit SD). All of the parameter estimates listed in this table have an evidence ratio above 40 for a test of difference to zero.*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	758ms [679, 839]	786ms [711, 868]
	Convo With Parent	699ms [584, 809]	793ms [642, 947]
	Convo With Experimenter	830ms [699, 957]	1057ms [855, 1256]
	Aggregate Estimate	762ms [683, 837]	879ms [779, 982]
Sigma	Matching With Parent	645 [599, 694]	569 [520, 622]
	Convo With Parent	495 [440, 557]	571 [496, 660]
	Convo With Experimenter	897 [834, 963]	879 [808, 956]
	Aggregate Estimate	659ms [626, 695]	659ms [617, 701]
Child SD	Matching With Parent	189ms [145, 242]	183ms [130, 251]
	Convo With Parent	204ms [151, 267]	308ms [233, 404]
	Convo With Experimenter	328ms [259, 410]	453ms [340, 592]
	Aggregate Estimate	240ms [202, 282]	315ms [260, 377]
Visit SD	Matching With Parent	72ms [31, 134]	38ms [2.98, 93.19]
	Convo With Parent	82ms [6, 248]	80ms [4.95, 248.07]
	Convo With Experimenter	84ms [11, 213]	94ms [9, 240]
	Aggregate Estimate	79ms [34, 153]	71ms [23, 147]
Sigma Visit	Matching With Parent	-0.03 [-0.05, -0.01]	0 [-0.02, 0.03]
	Convo With Parent	0.02 [-0.01, 0.06]	-0.04 [-0.09, 0]
	Convo With Experimenter	-0.02 [-0.04, 0]	0.01 [-0.01, 0.03]
	Aggregate Estimate	-0.01 [-0.02, 0.01]	-0.01 [-0.03, 0.01]

6.4.6 Overlaps according to Individual Differences

To explore the extent to which potential increases or decreases in response latencies were a function of a greater proportion of overlaps, we modelled the proportion of overlaps (i.e., negative latencies) according to individual differences among children. As shown in **Figure 13**,

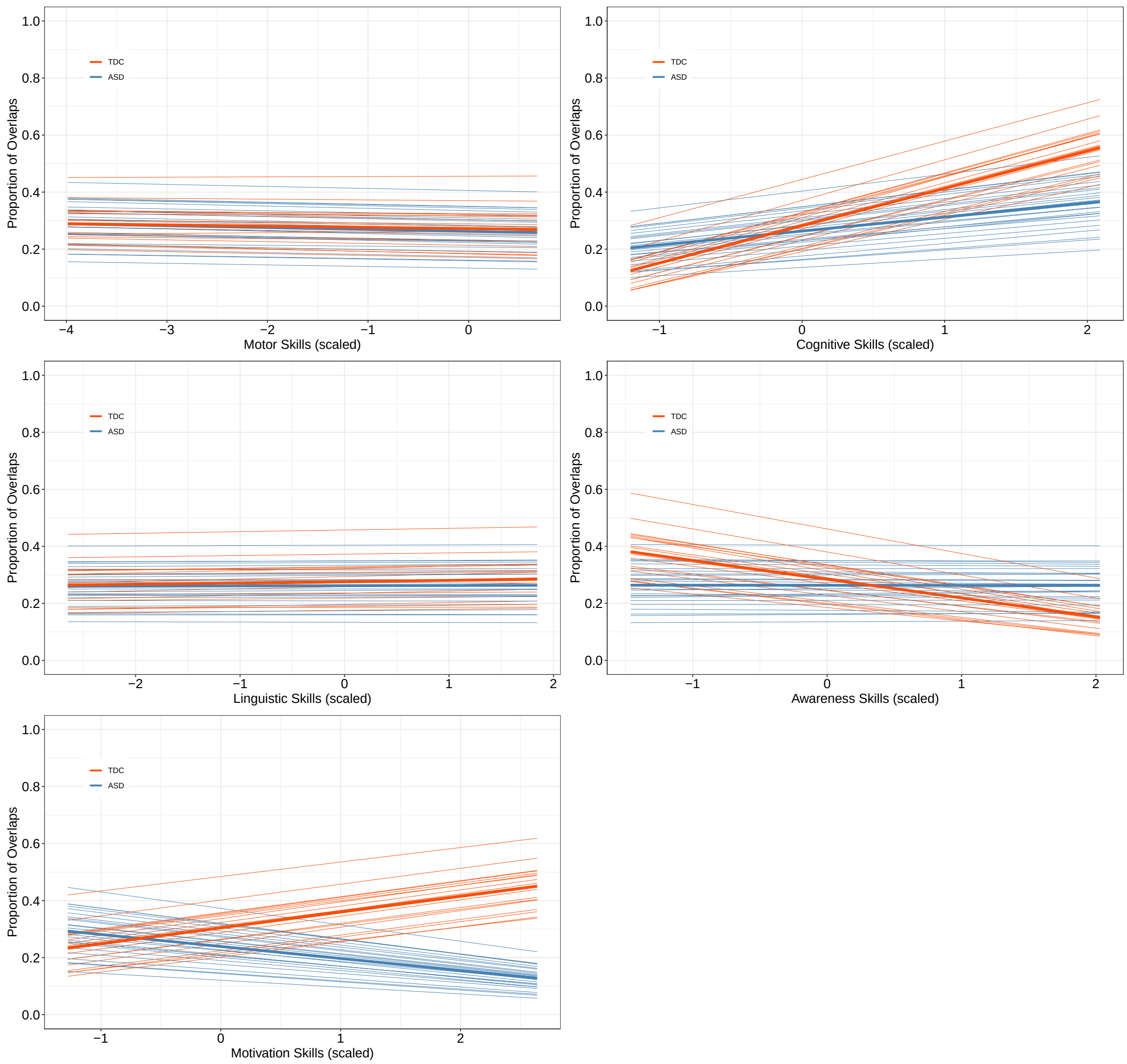

Figure 13: *This panel of plots displays the proportion of overlaps according to individual differences of the children. The faded lines are posterior predictions from model for individual subjects and how the proportion of overlaps changes for each of the continuous predictors.*

6.4.7 Individual Differences With No Overlaps

To check the extent to which overlaps were driving the estimates of change in response latencies according to individual differences, we ran the same individual differences model on data where overlaps (i.e., negative latencies) were excluded. The posterior estimates from this model can be viewed in **Table 15** below.

Table 15: *Posterior estimates when excluding overlaps for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of each of the types of skills listed below (Motor, Cognitive Skills, Language Skills, Social Awareness and Social Motivation). The estimates denote the change in child response latency as a function of one standard deviation increase in the skills of the children. The parameter estimates with an evidence ratio of above 10 for a test of difference to null are marked with a *.*

	Skills	**Autism Group**	**Typical Development**
Social Cognition	Matching With Parent	-13ms [-42, 15]	-95ms [-382, 147]
	Convo With Parent	12ms [-30, 55]	50ms [-171, 291]
	Convo With Experimenter	-76ms [-176, 27]	-490ms [-883, -83]
	Aggregate Estimate	-26ms [-67, 16]	-178ms [-367, -4]
Social Awareness	Matching With Parent	4.24ms [-29, 37]	69ms [-81, 252]
	Convo With Parent	27ms [-21, 77]	-59ms [-205, 81]
	Convo With Experimenter	-4.11ms [-117, 108]	190ms [-65, 449]
	Aggregate Estimate	9ms [-39, 56]	67ms [-44, 189]
Social Motivation	Matching With Parent	4.48ms [-21, 30]	-46ms [-219, 116]
	Convo With Parent	-26ms [-64, 11]	78ms [-77, 245]
	Convo With Experimenter	84ms [-3.38, 169.97]	90ms [-194, 377]
	Aggregate Estimate	21ms [-15, 57]	40ms [-84, 165]
Language	Matching With Parent	1.94ms [-14, 18]	20ms [-33, 84]
	Convo With Parent	16ms [-8, 43]	-18ms [-73, 35]
	Convo With Experimenter	0.01ms [-57, 58]	35ms [-59, 131]
	Aggregate Estimate	6ms [-18, 31]	12ms [-29, 55]
Motor	Matching With Parent	0.18ms [-18, 18]	11ms [-69, 94]
	Convo With Parent	-6ms [-35, 21]	-9ms [-89, 67]
	Convo With Experimenter	41ms [-21, 101]	-15ms [-155, 130]
	Aggregate Estimate	12ms [-14, 36]	-4.13ms [-67, 57]

6.4.8 Predictability Model With No Backchannels

As shown in **Table 16**, the overall degree of utterance predictability differed slightly across interactional contexts, with both the child and adult utterances agreeing with each other in terms of their overall degree of predictability.

Table 16: *Overview of Utterance Predictability Across Different Contexts*

	Condition	Diagnosis	OtherPredictability	OwnPredictability
1	Matching With Parent	Autism Group	0 (0.98)	-0.02 (0.96)
2	Matching With Parent	Typical Development	-0.02 (1.02)	-0.07 (0.97)
3	Convo With Parent	Autism Group	0.08 (1.05)	0.19 (1.02)
4	Convo With Parent	Typical Development	0.08 (1.1)	0.23 (1.04)
5	Convo With Experimenter	Autism Group	-0.08 (1)	0.01 (0.98)
6	Convo With Experimenter	Typical Development	-0.15 (0.97)	-0.03 (0.98)

To check the potential influence of backchannels on the posterior estimates, we screened utterances that occurred most frequently ($n > 10$) in the corpus, removed backchannels from the data, and ran the same predictability model. **Table 17** provides an overview of the 15 most frequently occurring backchannels in the corpus. The estimates from this control model are shown in **Table 18**.

Table 17: *Table with examples of backchannels that were excluded from this control model.*

Backchannel	n
mhm	1089
yeah	951
okay	938
laugh	303
yes	249
no	195
yep	138
um	132
oh	123
hm	108
uhhuh	86
uh	72
mm	70
mhm okay	66
oh okay	61

Table 18: *Posterior estimates for how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous adult utterance and child utterance. The estimates denote the change in child response latency as a function of one standard deviation increase in the predictability of the utterance (i.e., cosine similarity).*

	Tasks	**Autism Group**	**Typical Development**
Predictability of Previous Adult Utterance	Matching With Parent	-38ms [-65, -12]	-16ms [-51, 23]
	Convo With Parent	-39ms [-73, -3]	-69ms [-127, -11]
	Convo With Experimenter	-13ms [-51, 26]	-12ms [-69, 43]
	Aggregate Estimate	-30ms [-50, -11]	-32ms [-62, -2]
Predictability of Own Utterance	Matching With Parent	-21ms [-44, 3]	13ms [-17, 41]
	Convo With Parent	47ms [19, 76]	6ms [-28, 40]
	Convo With Experimenter	164ms [127, 200]	119ms [71, 166]
	Aggregate Estimate	63ms [46, 81]	46ms [24, 67]

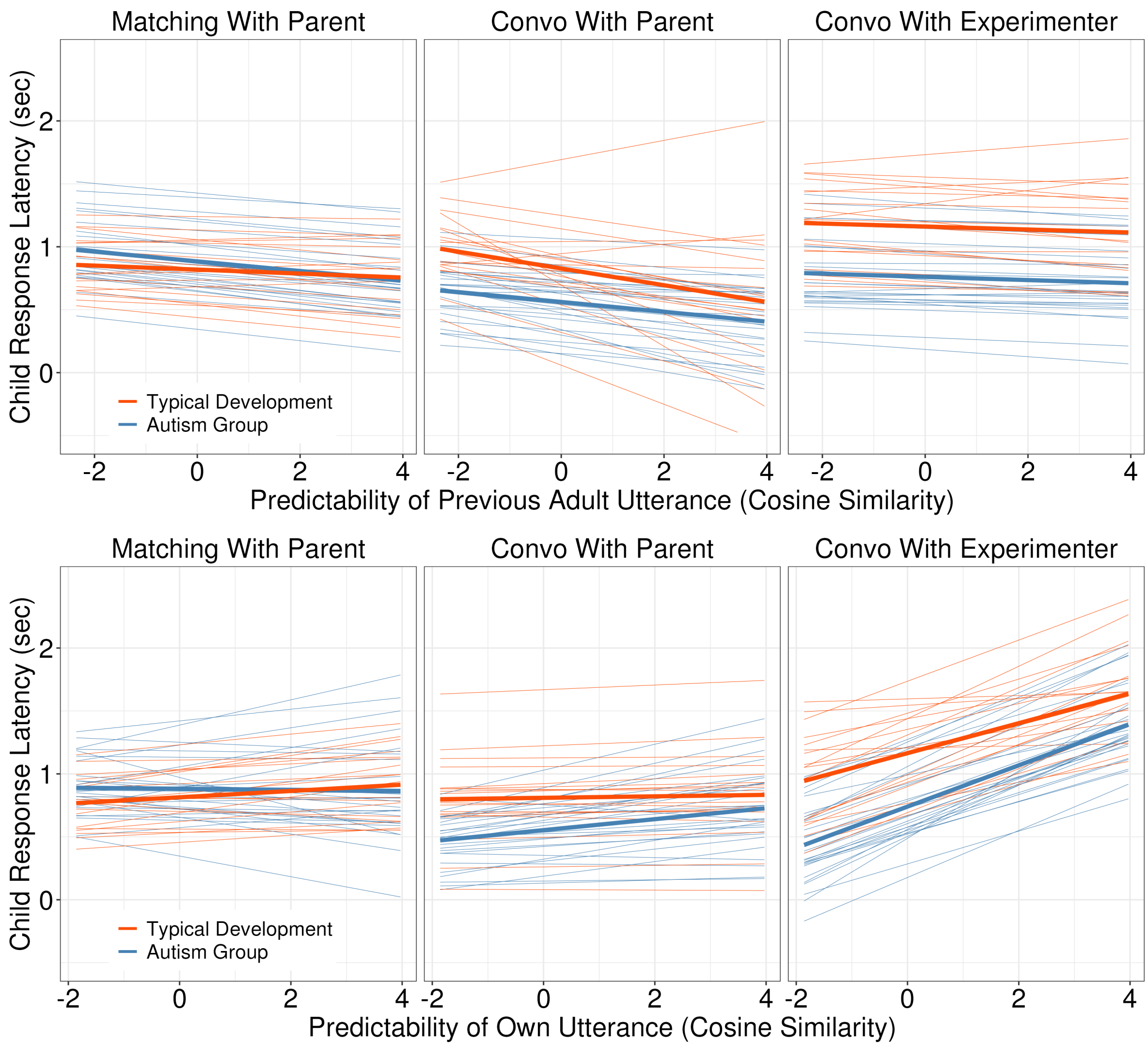

Figure 14: *Posterior estimates in a model excluding backchannels. The plot shows how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous adult utterance and current child utterance.*

6.4.9 Predictability Model With Short Utterances Removed

To check the potential influence of shorter utterances on the posterior estimates, we removed all utterances under three words in the corpus and ran the same predictability model as in the main manuscript. The estimates from this control model are shown in **Table 19**.

Table 19: *Posterior estimates for how child response latencies change across individual conditions (Matching With Parent, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous adult utterance and child utterance. The estimates denote the change in child response latency as a function of one standard deviation increase in the predictability of the utterance (i.e., cosine similarity).*

	Tasks	**Autism Group**	**Typical Development**
Predictability of Previous Adult Utterance	Matching With Parent	-54ms [-93, -14]	33ms [-18, 86]
	Convo With Parent	-25ms [-66, 17]	-72ms [-136, -6]
	Convo With Experimenter	-36ms [-93, 21]	-11ms [-83, 61]
	Aggregate Estimate	-38ms [-66, -11]	-17ms [-53, 20]
Predictability of Own Utterance	Matching With Parent	-31ms [-60, -3]	-7ms [-47, 32]
	Convo With Parent	10ms [-26, 47]	-12ms [-58, 34]
	Convo With Experimenter	188ms [145, 230]	108ms [50, 166]
	Aggregate Estimate	56ms [34, 77]	29ms [1.13, 56.88]

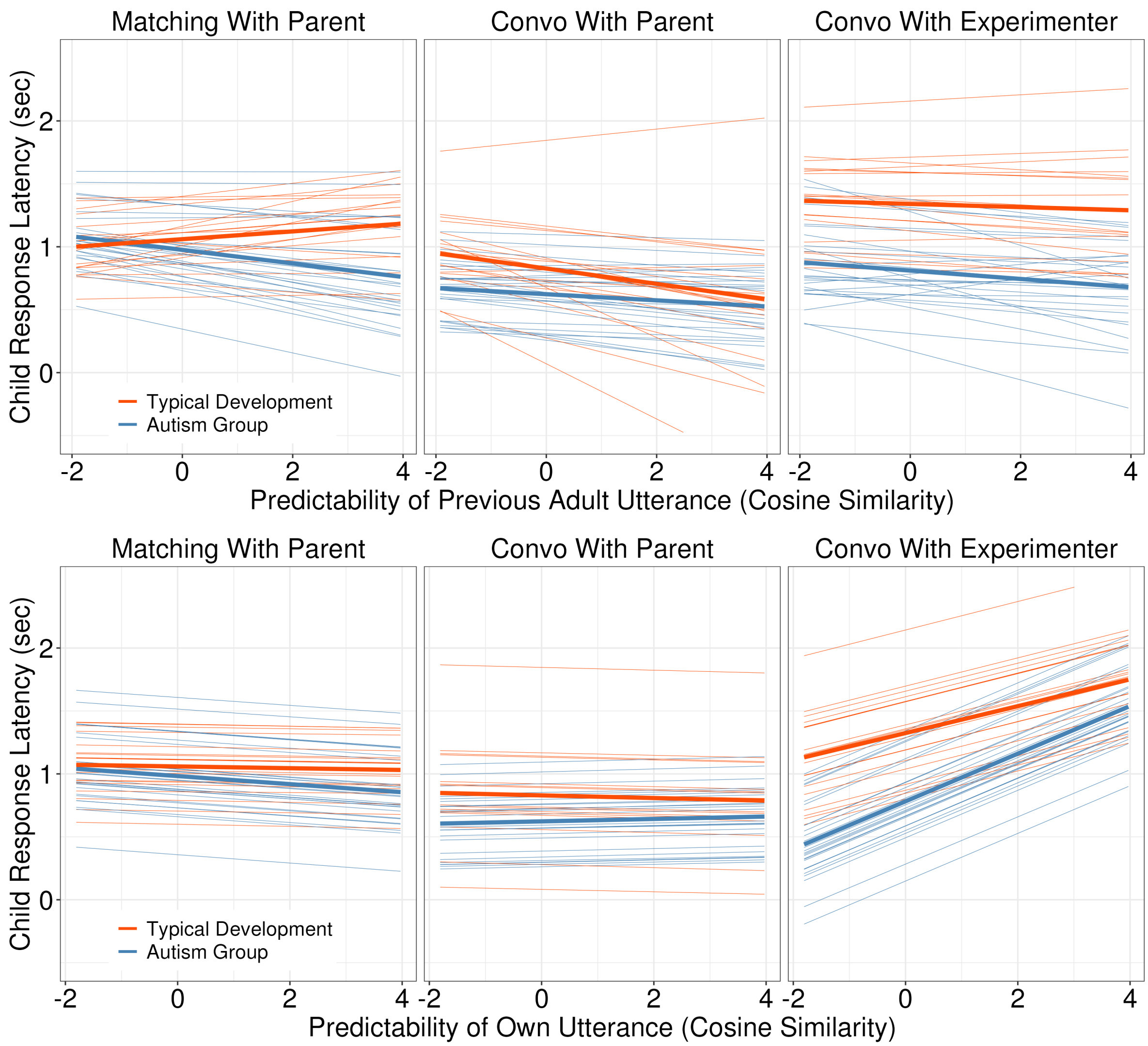

Figure 15: *Posterior estimates in a model excluding utterances below three words. The plot shows how child response latencies change across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and all conditions (Aggregate Estimate) as a function of the predictability of the previous adult utterance and current child utterance.*

6.4.10 Model With Surrogate Pairs

As a control analysis, we ran the first model after rotating the dataset; that is, pairing the utterances of one child with the interactant of another child and recalculating the response latencies between turns. The estimates from this control model are shown in **Table 20**.

Table 20: *Posterior estimates for child response latencies across individual conditions (Matching Game, Convo With Parent and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency), exponential component (Beta), Residual Heterogeneity (Sigma), between-subjects standard deviation (Child SD), standard deviation across visits (Visit SD), Residual Heterogeneity across visits (Sigma Visit) and proportion of latencies below zero (Overlap Proportion).*

	Tasks	**Autism Group**	**Typical Development**
Latency	Matching With Parent	27ms [-1632, 1806]	-202ms [-1896, 1544]
	Convo With Parent	-1823ms [-2576, -1087]	-2070ms [-3003, -1079]
	Convo With Experimenter	-2927ms [-3751, -2156]	-2009ms [-2671, -1330]
	Aggregate Estimate	-1574ms [-2293, -831]	-1427ms [-2183, -632]
Sigma	Matching With Parent	17735 [13366, 23464]	16909 [12333, 23217]
	Convo With Parent	4737 [3299, 6966]	4662 [2355, 8587]
	Convo With Experimenter	8091 [5725, 11433]	5827 [3685, 9164]
	Aggregate Estimate	8793ms [7070, 10988]	7716ms [5553, 10557]
Child SD	Matching With Parent	3015ms [1756, 4627]	3100ms [1917, 4587]
	Convo With Parent	1102ms [744, 1526]	823ms [458, 1261]
	Convo With Experimenter	1579ms [1070, 2196]	1096ms [687, 1611]
	Aggregate Estimate	1899ms [1405, 2496]	1673ms [1224, 2220]

Figure 16: *Posterior estimates for surrogate child response latencies across individual conditions (Matching Game, Parent Conversations and Experimenter Conversations) and aggregated across conditions (Aggregate Estimate) for different model parameters: the Gaussian component (Latency), exponential component (Beta), Residual Heterogeneity (Sigma), between-subjects standard deviation (Child SD), standard deviation across visits (Visit SD), Residual Heterogeneity across visits (Sigma Visit) and proportion of latencies below zero (Overlap Proportion).*

6.5. Model Information and Quality Checks

6.5.1 Choice of Priors

We chose weakly informative priors in order to ensure that their influence on the estimates was small and to discount extreme effect sizes as unlikely (cf. Lemoine, 2019; Gelman, Simpson & Betancourt, 2017). For the overall distribution, we chose an ex-Gaussian distribution. Our prior for the Gaussian component was specified as having a mean of 1 and standard deviation of 1 based on our prior expectations for child response latencies in a turn-taking context (Nguyen et al., 2022). This prior implies that we expect approximately 95% of the child latencies to fall between -1sec and 3sec. For the between-subject variability across different conversational contexts, we encoded our expectations with a truncated (i.e., standard deviation must have positive values) Gaussian prior with a mean of 0 and a standard deviation of 0.3, which implies that we expect the vast majority of values for the between-subject variability within diagnostic group to be between 0 and 0.6sec. For the rate parameter of the exponential component of the distribution, we specified a Gaussian prior with a mean of 0 and standard deviation of 1 and allowed this to vary for each task and familiarity for each participant.

The models were fitted with Hamiltonian Monte Carlo samplers with 2 parallel chains with 5,000 iterations each, an adapt delta of 0.99 and a maximum tree depth of 20 in order to ensure no divergence in the estimation process. The quality of the models was assessed by i) ensuring Rhat statistics to be lower than 1.1, ii) carrying out prior and posterior predictive checks, iii) plotting prior against posterior estimates and assessing whether the posteriors had lower variance than the priors, iv) ensuring no divergences in the process of estimation, v) checking that the number of effective bulk and tail samples was above 200, vi) conducting prior sensitivity analyses.

As an example of what the basic *brms* formula looked like, here is the formula for the first model (cf., code on OSF for the full code and *brms* formulae for other models):

Latency ~ 0 + Diagnosis:Task:Familiarity + (0 + Task:Familiarity | p | gr(ID, by = Diagnosis)) + (0 + Diagnosis:Task:Familiarity | r | Visit),

Sigma ~ 0 + Diagnosis:Task:Familiarity + Diagnosis:Task:Familiarity:Visit + (0 + Task:Familiarity:Visit | p | gr(ID, by = Diagnosis)),

Beta ~ 0 + Diagnosis:Task:Familiarity + (0 + Task:Familiarity | p | gr(ID, by = Diagnosis)) + (0 + Diagnosis:Task:Familiarity | r | Visit)

6.5.2 Prior and Posterior Predictive Checks

We performed quality checks of the models by carrying out prior and posterior predictive checks. The below prior predictive checks (on the left) indicate that our priors predict values within the order of magnitude of the distribution. The posterior predictive checks (on the right) indicate that the models have captured the distributions of data. These plots provide reassurance that our models capture relevant aspects of the overall distributions of dependent variables.

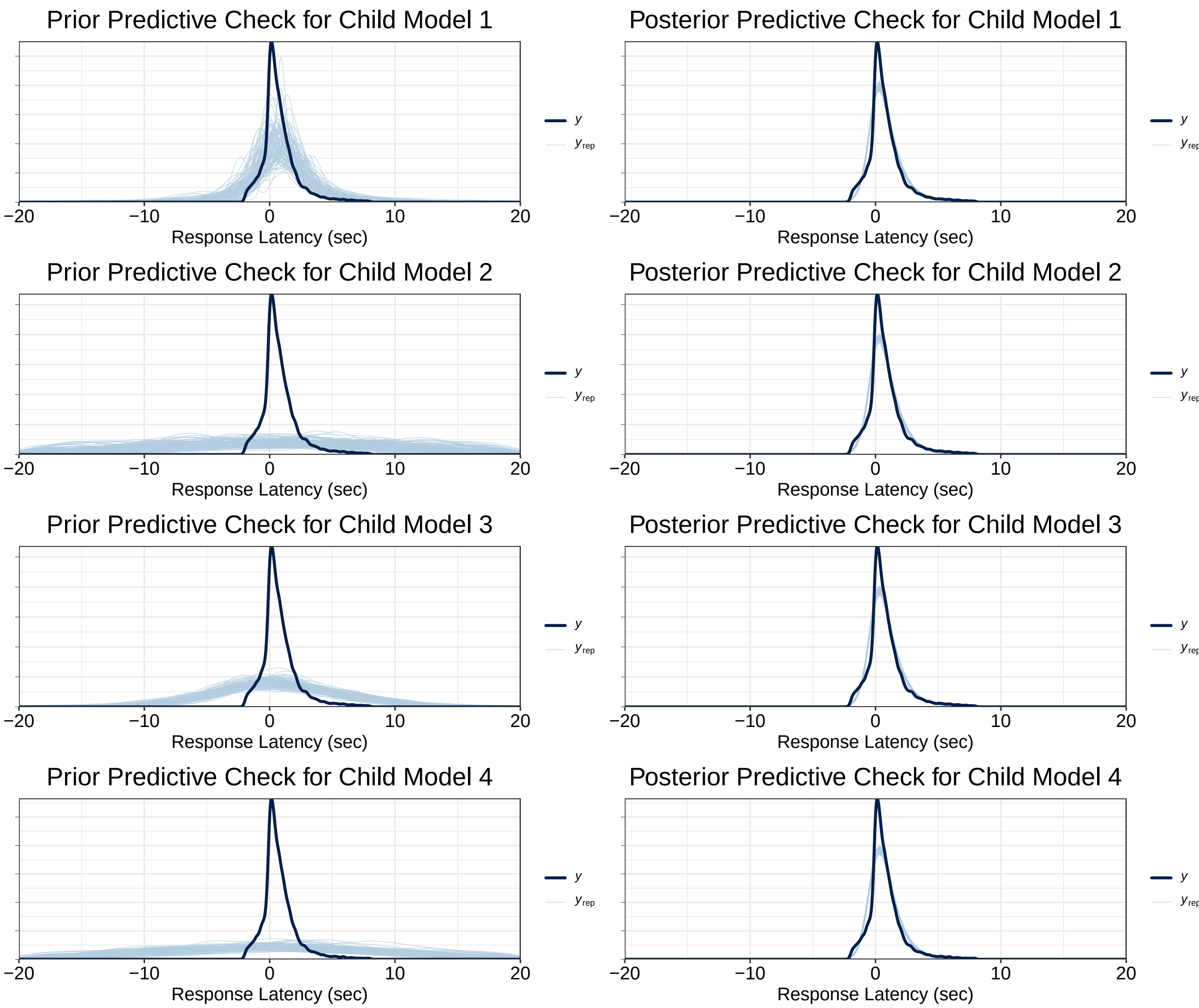

Figure 17: *Prior predictive checks (left column) and posterior predictive checks (right column) for child latency in models 1-4.*

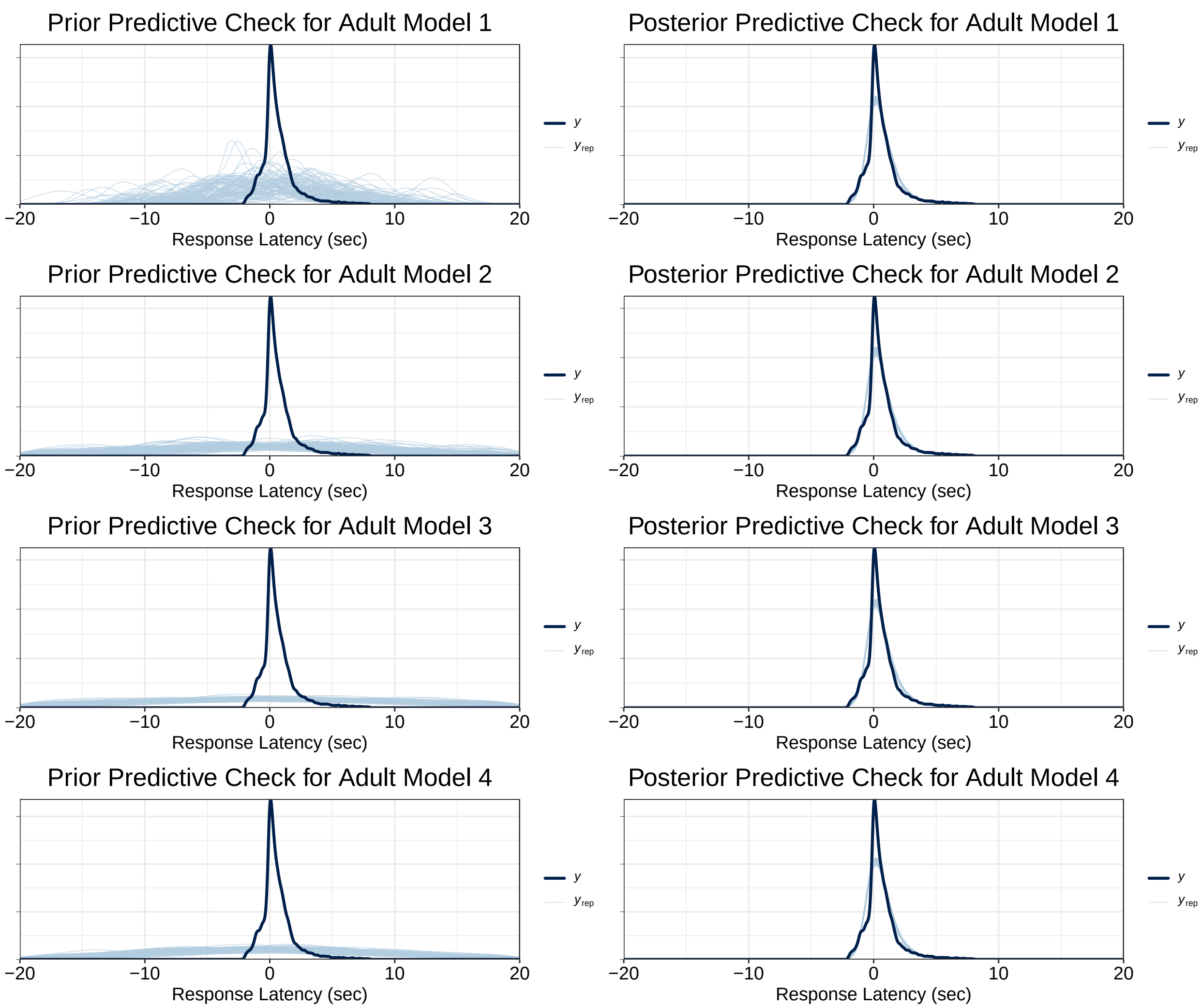

Figure 18: *Prior predictive checks (left column) and posterior predictive checks (right column) for adult latency in models 1-4.*

6.5.3 Prior-Posterior Update Plots

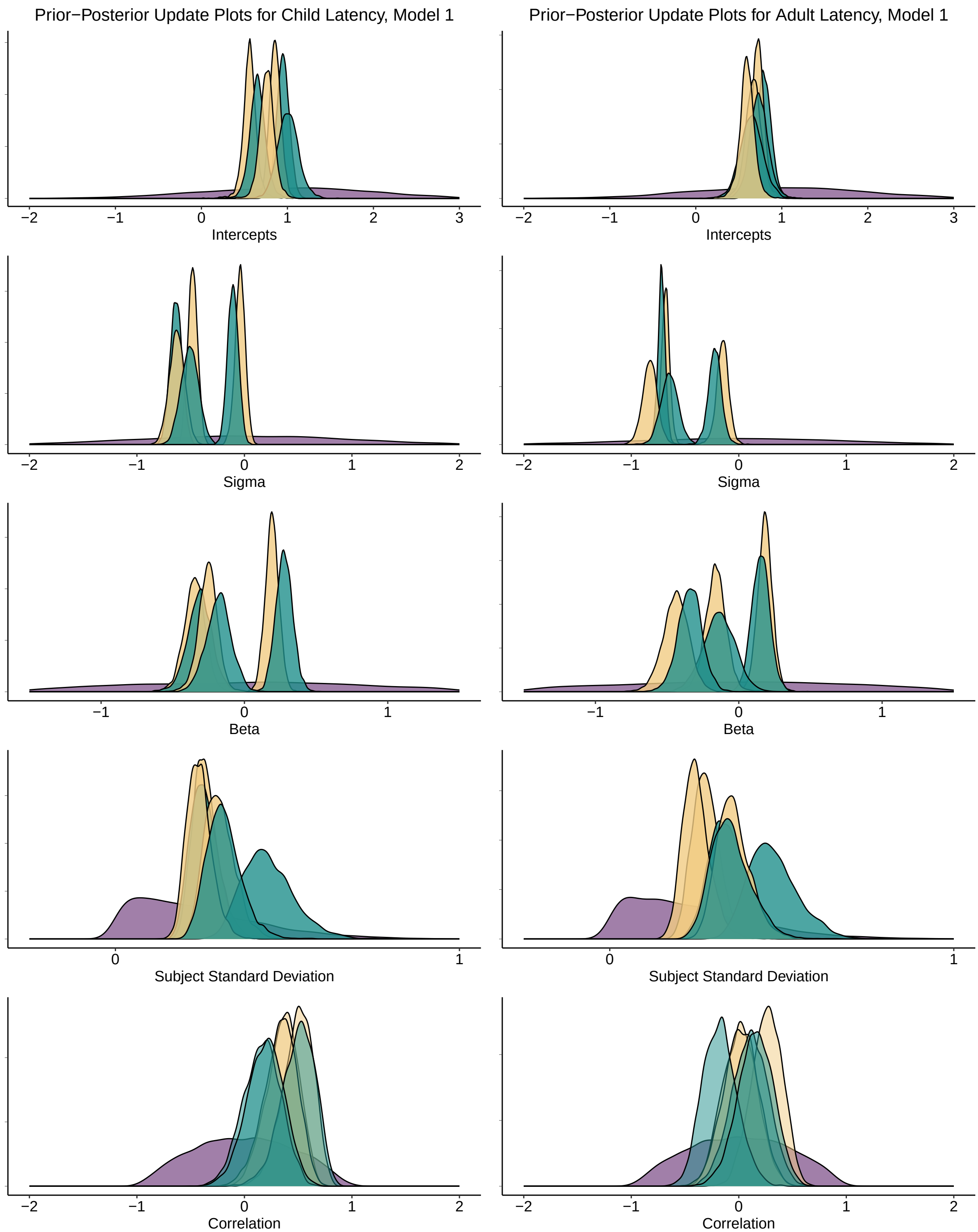

Figure 19: *Prior-posterior update checks for child and adult latency in Model 1. Purple density plots show the prior predictive density plot, turquoise density plots indicate posterior predictions for the typical development group, and yellow density plots denote the predicted estimates for the autism group.*

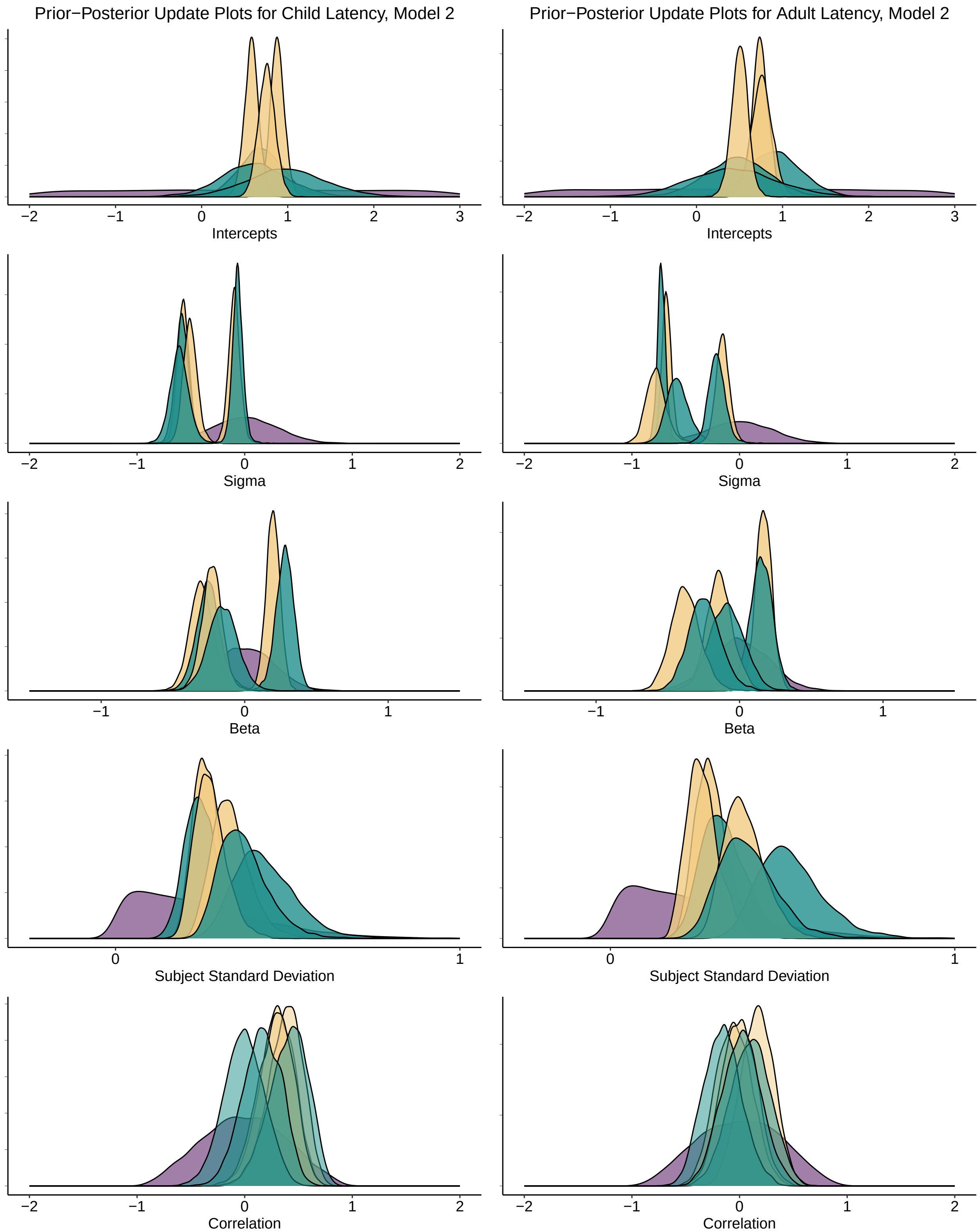

Figure 20: *Prior-posterior update checks for child and adult latency in Model 2. Purple density plots show the prior predictive density plot, turquoise density plots indicate posterior predictions for the typical development group, and yellow density plots denote the predicted estimates for the autism group.*

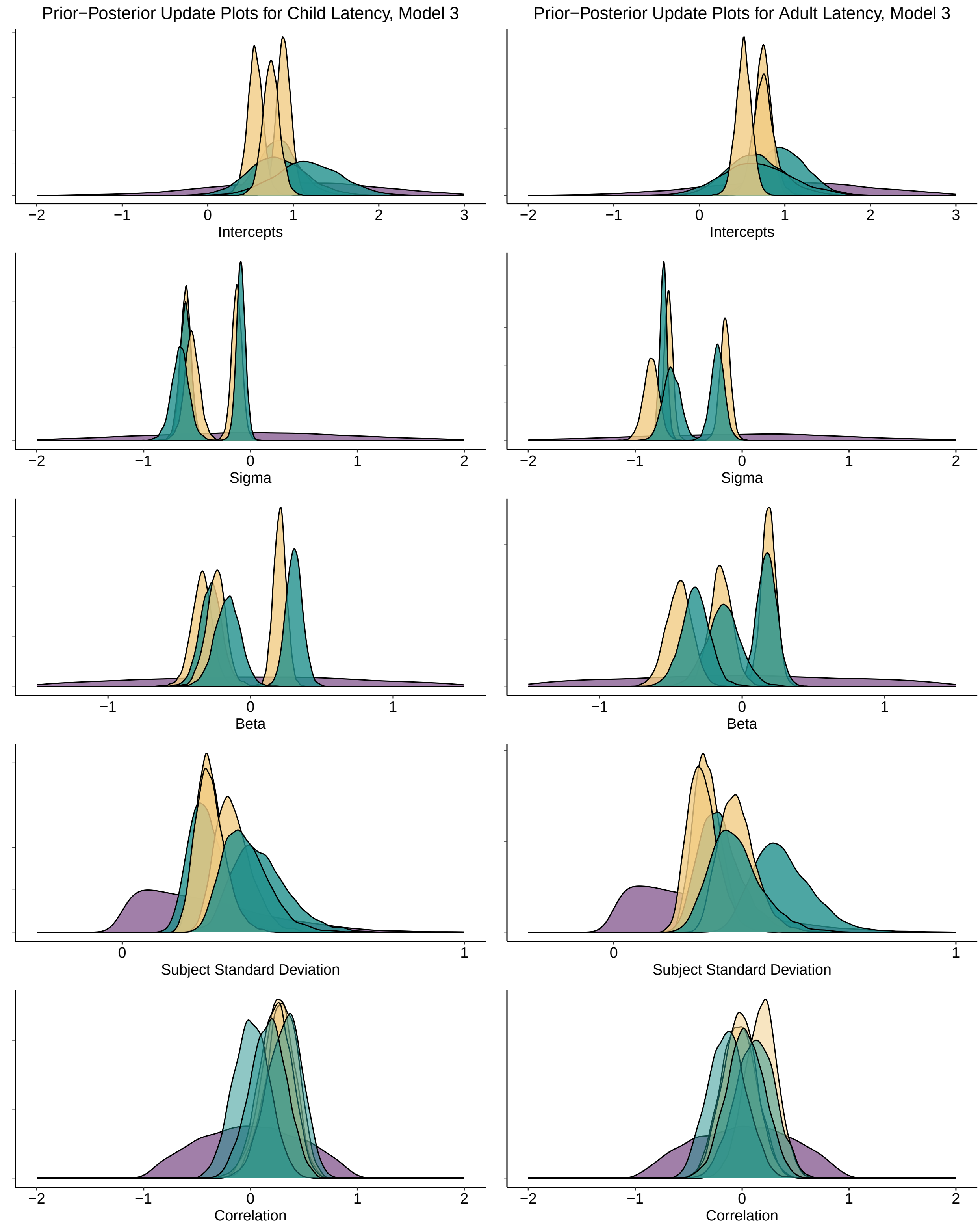

Figure 21: *Prior-posterior update checks for child and adult latency in Model 3. Purple density plots show the prior predictive density plot, turquoise density plots indicate posterior predictions for the typical development group, and yellow density plots denote the predicted estimates for the autism group.*

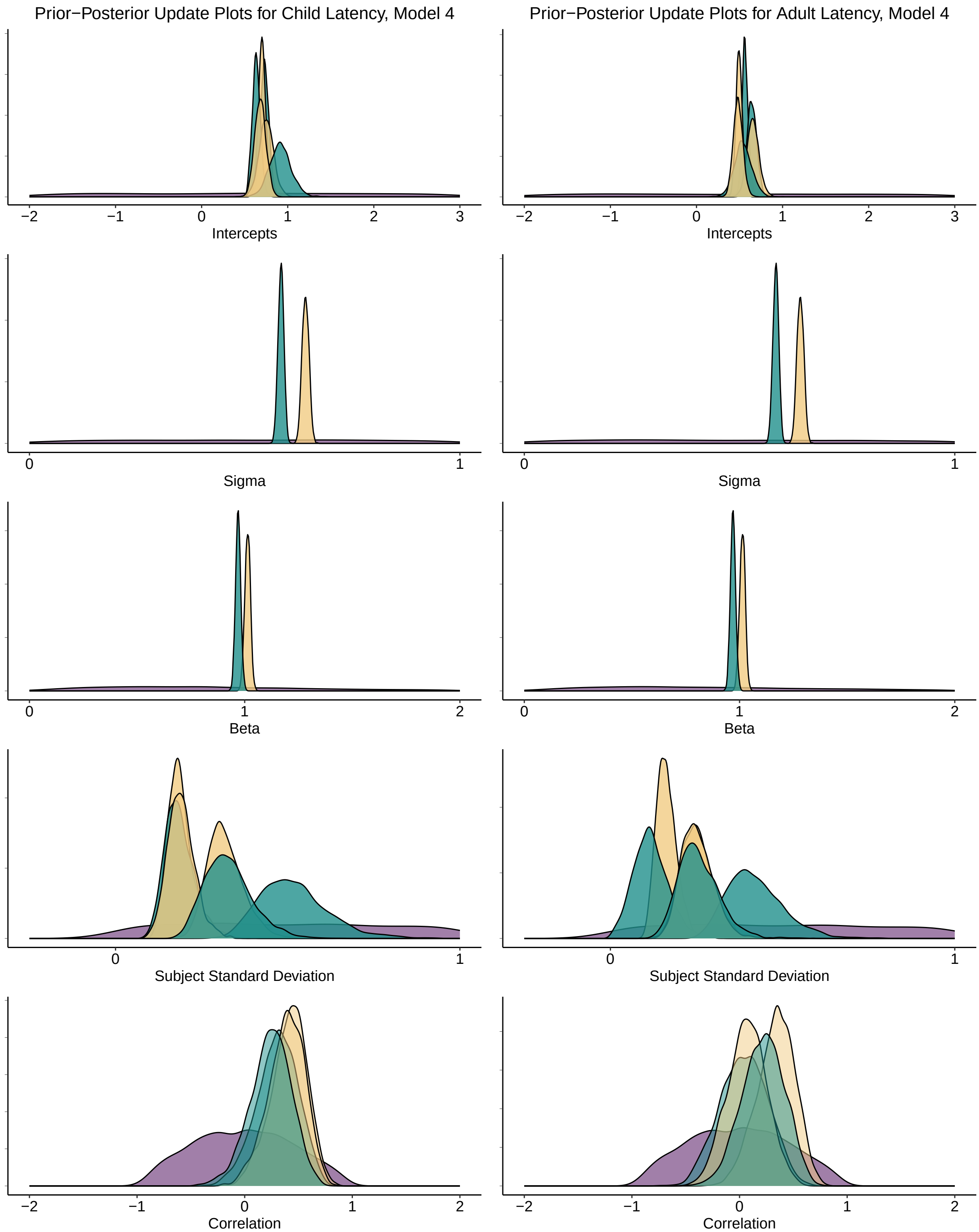

Figure 22: *Prior-posterior update checks for child and adult latency in the multivariate Model 4. Purple density plots show the prior predictive density plot, turquoise density plots indicate posterior predictions for the typical development group, and yellow density plots denote the predicted estimates for the autism group.*