

Communicative Misalignment in Autism Spectrum Disorder

Harshali Wadge^a, Rebecca Brewer^b, Geoffrey Bird^{c,d}, Ivan Toni^{e,*} & Arjen Stolk^{a,e,*}

^aHelen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA

^bDepartment of Psychology, Royal Holloway University of London, London, UK

^cDepartment of Experimental Psychology, University of Oxford, Oxford, UK

^dSocial, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

^eDonders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

*These authors contributed equally

Corresponding author: astolk@berkeley.edu

Abstract

Communication deficits are a defining feature of Autism Spectrum Disorder (ASD), manifest during social interactions. Previous studies investigating communicative deficits have largely focused on the perceptual biases, social motivation, cognitive flexibility, or mentalizing abilities of isolated individuals. By embedding autistic individuals in live non-verbal interactions, we characterized a novel cause for their communication deficits. Adults with ASD matched neurotypical individuals in their ability and propensity to generate and modify intelligible behaviors for a communicative partner. However, they struggled to align the meaning of those behaviors with their partner when meaning required referencing their recent communicative history. This communicative misalignment explains why autistic individuals are vulnerable in everyday interactions, which entail fleeting ambiguities, but succeed in social cognition tests involving stereotyped contextual cues. These findings illustrate the cognitive and clinical importance of considering social interaction as a communicative alignment challenge, and how ineffective human communication is without this key interactional ingredient.

Keywords:

Human communication, Social interaction, Conceptual alignment, Theory of mind, Mentalizing

1. Introduction

Autism Spectrum Disorder (ASD) is diagnosed on the basis of communicative deficits observed in everyday social interactions (American Psychiatric Association, 2013; Frith, 2003; Klin, McPartland, & Volkmar, 2013). The deficits are most evident in situations where the speaker's intention and a sentence's literal meaning strongly diverge, such as in the case of irony and sarcasm (Tesink et al., 2009; Zalla et al., 2014), and have been argued to be a product of a primary impairment in representing mental states (Baron-Cohen, Leslie, & Frith, 1985; Happé, 1993). However, empirical studies of this impairment have produced mixed results, including compelling observations of intact social perception and reasoning in individuals with ASD (Bowler, 1992; Cusack, Williams, & Neri, 2015; Pantelis & Kennedy, 2017; Sally & Hill, 2006). Other accounts suggest that the communication deficits arise from core difficulties with social motivation, social attention, or cognitive flexibility (Chambon et al., 2017; Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012; Geurts, Corbett, & Solomon, 2009). Yet other accounts emphasize biases in processing biological and multimodal linguistic cues used during face-to-face interactions (Constantino et al., 2017; J. Cook, Saygin, Swain, & Blakemore, 2009; Hobson, Ouston, & Lee, 1988; Hutchins & Brien, 2016; Nackaerts et al., 2012; Silverman, Bennetto, Campana, & Tanenhaus, 2010). These considerations highlight severe limitations in our understanding of communication in ASD, and consequently a lack of principled interventions for improving communication between autistic and neurotypical individuals.

Here we examined the possibility that individuals with ASD have difficulties in using the conceptual space defined by an ongoing interaction to resolve the pervasive ambiguity of human communicative signals (Goodman & Frank, 2016; Levinson, 1983; Stolk, Verhagen, & Toni, 2016). Human communication is often framed in terms of signal transmission, presupposing that communicators already share the same set of encoding-decoding rules, e.g., a common language (Akmajian, Farmer, Bickmore, Demers, & Harnish, 2017; Eco, 1976; Jakobson, 1971). Yet even commonly used words do not contain fixed meanings that are reliably shared across communicators (Grice, 1975; Rumelhart, 1979; Sperber & Wilson, 1996). Their meaning is flexibly coordinated through an online interpersonal alignment process by which people in dialogue seek and provide evidence that they understand one another (Susan E. Brennan, Galati, & Kuhlen, 2010; Fusaroli & Tylén, 2016; Garrod & Anderson, 1987; Misyak, Melkonyan, Zeitoun, & Chater, 2014; Stolk et al., 2016). This dynamic alignment process provides a conceptual frame of reference necessary for interpreting intrinsically ambiguous communicative signals (Stolk et al., 2016). The present study quantitatively tested whether

individuals with ASD have difficulties in dynamically aligning conceptualizations of their behaviors with a communicative partner.

This test was implemented in a novel communicative setting in which cognitively able adults with ASD interacted with another individual via a digital game board (Fig. 1A). This two-player computer game captures the open-ended and interpersonal nature of everyday communication by challenging players to generate, negotiate, and align the meaning of their non-verbal communicative behaviors (de Ruiter et al., 2010). The design of the computer game remediates several factors hypothesized to account for communicative impairments in ASD. For instance, the game prevents recourse to pre-existing shared communicative representations, such as those offered by some linguistic and gestural emblems (Groen, Zwiets, van der Gaag, & Buitelaar, 2008), and avoids verbal and face-to-face contact between players. This nullifies the effects of individual differences in processing biological and multimodal linguistic cues (Constantino et al., 2017; J. Cook et al., 2009; Hobson et al., 1988; Hutchins & Brien, 2016; Nackaerts et al., 2012; Silverman et al., 2010). Furthermore, the game manipulates the ambiguity of the communicative signals by introducing problems that are more easily solved in light of previous interactions. This feature mimics daily conversation, which consists of ambiguous words and behaviors that can only be discerned by individuals who know their context of use in an ongoing interaction. Under these controlled yet genuinely interactive experimental circumstances, it becomes feasible to quantitatively test whether communicative deficits in individuals with ASD arise specifically from difficulties establishing dynamic conceptual alignment of ambiguous signals with another person.

2. Materials and methods

2.1. Participants

Fifty-two adult participants were recruited to take part in this study ($M \pm SD$ 23.9 \pm 6.5 yrs of age, 20 females, 22 individuals with ASD). ASD participants were recruited from a database of research volunteers maintained by the authors. Typical participants were recruited from local participant pools populated by students of the University of London and members of the general public, and were selected to match the ASD sample on age, gender, and IQ (see below). Participants were assigned pairwise to either the ASD group (7 pairs, each containing two individuals with ASD), the Typical group (11 pairs, each containing two individuals with no clinical diagnosis), or the Mixed group (8 pairs, each including one individual with ASD and one individual with no clinical diagnosis). The Mixed group served as a control for behavioral

changes related to the clinical status of the communicative partners in the full ASD and Typical groups. All study procedures received ethical approval by the local institution's ethics committee and all participants provided informed consent in line with the declaration of Helsinki. Analysis took place after completion of data collection.

Individuals with and without ASD did not differ in terms of age (24.7 ± 6.5 vs. 23.2 ± 6.6 , $t(50) = 0.79$, $p = .44$, $d = 0.23$, 95% CI = [-0.33 0.78], $BF = 2.76$ in favor of the null hypothesis of no difference) or gender (6/22 vs. 14/30 females, $\chi^2(1,52) = 2.02$, $p = .16$, $d = 0.19$, 95% CI = [-0.06 0.45], $BF = 1.16$ in favor of the null hypothesis). They also did not differ in terms of IQ (102.1 ± 19.4 vs. 109.4 ± 14.1 , $t(50) = -1.59$, $p = .12$, $d = -0.46$, 95% CI = [-1.01 0.10], $BF = 1.28$ in favor of the null hypothesis), measured by the Wechsler Abbreviated Scale of Intelligence (WASI-II, (Wechsler, 2011)) in neurotypical individuals, and the Wechsler Adult Intelligence Scale (WAIS-IV, (Wechsler, 2008)) in those with ASD. Autistic traits were assessed in all individuals using the Autism-Spectrum Quotient (AQ, (Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006)), and were more prevalent in individuals with ASD than in neurotypical participants (32.6 ± 8.4 vs. 16.0 ± 8.0 , $t(50) = 7.23$, $p < .001$, $d = 2.07$, 95% CI = [1.39 2.75]). All individuals with ASD were diagnosed by an independent clinician, and completed the Autism Diagnostic Observation Schedule (ADOS-2, (Lord et al., 2012)). Six of the ASD participants met the ADOS criteria for Autism, while nine met the criteria for Autism Spectrum. Seven individuals with ASD did not meet ADOS criteria despite their clinical diagnosis, but were diagnosed by an independent clinician and reported a similar degree of autistic traits on the AQ as the individuals who did meet the ADOS criteria (31.5 ± 10.2 , 32.1 ± 9.7 , and 33.9 ± 5.7 for the Autism, Autism Spectrum, and the None group, respectively). Independent samples t-tests did not indicate differences in AQ scores between these three groups (all $p > .62$, all $BF > 1.88$ in favor of the null hypothesis). Supplemental Table S1 represents an overview of AQ scores and ADOS classifications for all individuals diagnosed with ASD.

2.2. Task

We used the same two-player communication game employed in a previous experiment (Stolk, Verhagen, et al., 2013). The game involves pairs of participants interacting on a digital game board with a 3 x 3 grid layout, which was visually presented on each participant's computer screen (Fig. 1A). Each pair communicated in real-time over the course of 80 interactions, alternating between the roles of Communicator and Addressee across successive interactions. During each interaction, their goal was to reproduce a target configuration of two geometric

shapes on the game board. Each member of the pair controlled the movement of one of these shapes. The target configuration was shown to the Communicator only and thus a successful interaction required the Communicator to convey to the Addressee the target location and orientation of the Addressee's shape, while also ensuring that the final location and orientation of their own shape was as specified by the target configuration. In this game, the only means available to the Communicator for communicating with the Addressee is by moving the Communicator's own shape around the grid using horizontal and vertical translations, and 90° clockwise rotations controlled by button presses on a handheld game controller (four face buttons and one shoulder button, respectively). The only means available to the Addressee for completing the configuration is by inferring the target location and orientation of his or her own shape on the basis of the movements of the Communicator, and positioning it accordingly using a second handheld game controller.

At interaction onset, each player is assigned their role (Communicator or Addressee) and shape (event I in Fig. 1A), followed by presentation of the target configuration to the Communicator (event II). Both Communicator and Addressee know that the Communicator has unlimited time available for planning the movements of the Communicator's shape, but only 10 seconds to execute them (event III). As soon as the Communicator presses the start/stop button, the target configuration disappears from the Communicator's screen and the Communicator's shape appears in the center of the grid on both participants' screens, signaling readiness to move. All movements are then visible on both the Communicator's and Addressee's screens. After 10 seconds, or earlier if the Communicator presses the start/stop button for a second time, the Communicator's shape cannot move further and the Addressee's shape appears in the center of the grid, indicating control by the Addressee over the Addressee's own shape. Similarly, the Addressee has no time constraints on planning the movements of their shape, but only 10 seconds to position it in a location and orientation deemed correct on the basis of the movements of the Communicator (event IV). Finally, after 10 seconds, or earlier if the Addressee presses the start/stop button again, the same feedback is presented to both players in the form of a green check mark or red cross, indicating whether or not the participants had successfully reproduced the target configuration (event V). The 80 target configurations were presented in the same predetermined order to all 26 participant pairs.

There are no a priori correct communication strategies in this game nor can the Addressee solve the task by simply reproducing the movements of the Communicator's shape. Rather, the Addressee needs to disambiguate communicative and instrumental components of

the Communicator's movements, and identify the relationship between the Communicator's movements and the message they intend to convey. Several pieces of evidence indicate that the players jointly and dynamically establish an agreement, also known as a 'conceptual pact' (S. E. Brennan & Clark, 1996), concerning the meaning of their behaviors. For instance, the same communicative behavior can be used by different pairs to convey different meanings. The same behavior can even have different meanings in different interactions of the same pair when the communicative agreement is jointly revised, and across various pairs (for examples see movies in (Stolk, Noordzij, Verhagen, et al., 2014; Stolk, Verhagen, et al., 2013)). To drive participants to continuously (re)negotiate the meaning of their communicative behaviors rather than to exploit already-established communicative conventions, we increased task difficulty across successive interactions. This was achieved by introducing deliberate mismatches between the geometrical characteristics of the pairs' shapes, and by introducing target orientations incompatible with the Communicator's shape. For instance, if a Communicator was able to successfully communicate the target orientation of the Addressee's shape by rotating their own shape, this strategy would be negated when the Communicator's shape changed to a circle and the Addressee's to a triangle. The Communicator would then have to find a new way to indicate the target orientation of the Addressee's shape, because rotations of the circle shape would not be visible. A further level of difficulty could be introduced by maintaining the same circle-triangle shape combination, but having the Addressee's triangle point outside the grid. If the Communicator had previously used a signal that involved repeatedly stepping in and out of the Addressee's target location in the direction of the target orientation (as seen in Fig. 1A), an interaction where the Addressee's triangle pointed outside of the grid would be impossible to solve using that signal. Therefore, the Communicator would have to devise a new behavior to solve this new problem type.

Prior to task performance, participants completed an individual training session in order to familiarize themselves with the handheld game controller, ensuring that each could manipulate their shape quickly and accurately. Given that atypical movements kinematics have been reported in ASD (J. L. Cook, Blakemore, & Press, 2013), we extended the allotted movement time for Communicator and Addressee from 5 to 10 seconds in this study, which was found to be sufficient for participants with ASD to execute their planned movements. Participants also jointly completed 20 practice interactions to familiarize themselves with the communicative setting and the sequence of events occurring during an interaction. To minimize any effect stemming from potential individual differences in working memory capacity, we

allowed Communicators to refer to a hard copy of the target configuration throughout the interaction. The order of trials was consistent across participant pairs. The experiment was programmed using Presentation (Neurobehavioral Systems, Albany, CA, USA) software on a Windows XP personal computer, and lasted approximately 40 minutes.

2.3. Analysis

Participants' behaviors, communicative success, and alignment of communicative signal use were analyzed offline using custom MATLAB (Mathworks, Natick, MA, USA), Python (Python Software Foundation, DE, USA) and Java code (Oracle, CA, USA), the SPSS (IBM, Armonk, NY, USA) and JASP (JASP Team, jasp-stats.org) statistical software packages, and the Gephi (Gephi Consortium, gephi.org) network visualization software package. Communicative success was calculated as the percentage of spatial configurations successfully reproduced by each pair over the course of their interactions.

Pairwise alignment was calculated as the percentage of interaction-by-interaction overlap in signals used by pair members in the Communicator role for communicating the target location and orientation of the Addressee's shape. The pair members were considered aligned at a given interaction if the type of communicative signal used in that interaction was conceptually identical (irrespective of movement speed and trial-specific trajectories) to the signal in the interactions preceding or following it, when the other member of the pair was the Communicator (Fig. 1C). The bidirectional sensitivity of the measure captures conceptual alignment even in instances where new communicative signals are introduced, and avoids ascribing transient task difficulty-related to signal changes to misalignment. Furthermore, unlike the inherently joint measure of communicative success, alignment can be calculated between members of different pairs. As outlined below, these pseudo-pair combinations provided a quantitative index of a problem's ambiguity, for testing whether autistic and neurotypical pairs differed in their ability to achieve alignment as a function of problem ambiguity.

An experimenter familiar with the communicative task (author H.W.) performed an interaction-by-interaction classification of the communicative signals using a custom-made replayer tool, while remaining blind to the clinical status of each participant. The signals varied from common pauses for distinguishing a target location from other visited locations, to idiosyncratic drawing, rotation, or 'wiggling' behaviors (stepping out from and back into a location, e.g., the double-arrowed number 2 action in Fig. 1A) for conveying the target location and orientation of the Addressee's shape. A list of signals and their descriptions was

constructed on the basis of previously identified behaviors from using this game (Blokpoel et al., 2012; Stolk, Noordzij, Verhagen, et al., 2014; Volman, Noordzij, & Toni, 2012). In cases where no pre-existing description was considered sufficiently accurate for an observed behavior, a new description was created for that communicative behavior. A total of 18 unique signals were observed, see Supplemental Table S2. It should be emphasized that in this game it cannot be determined on the basis of a single interaction alone whether, for instance, a wiggling behavior is intended to emphatically indicate the direction in which the Addressee's shape needs to point, or whether the number of wiggles corresponds to the number of clockwise rotations required for the Addressee's shape to reach the target orientation. Moreover, the communicative signals not only need to be disambiguated from other signals but also from instrumental elements of a Communicator's behavior which are necessary to achieve the Communicator's own target location and orientation. We controlled for subjectivity in the interpretation of the inherently ambiguous communicative behaviors in two ways. First, we asked a second rater to perform an interaction-by-interaction classification of the same dataset using the descriptions of behaviors observed by the first rater. We based our analyses on the classifications of the second rater given that the second rater was unaware of the details of this study. Second, we report statistical inferences based on the measure of pairwise alignment. Pairwise alignment is less prone to rater-specific interpretation and labeling of communicative signals, given that it considers the parity of signals across each pair's interactions as classified by the same rater. To validate the final signal classifications and resulting alignment values, we compared these measures between the two raters. The raters' individual signal classifications overlapped by 82% (intraclass correlation = 0.97, Kappa = 0.78). The subsequent alignment values for each pair overlapped by 84% (intraclass correlation = 0.81, Kappa = 0.68).

The main analysis tested for between-group differences in joint communicative success and pairwise alignment of communicative signals, over and above individual differences in general cognitive function. The fixed effect of Group (ASD, Mixed, Typical) was assessed with two univariate ANCOVAs using pairs' communicative success and pairwise alignment as dependent variables, respectively (Fisher, 1925). The covariate in these analyses considered the inter-pair variance in general cognitive function accounted for by the IQ in each pair (Reber, Walkenfeld, & Hernstadt, 1991). We selected between the mean and minimum IQ in a pair to describe this relationship, with the latter providing a better fit to communicative success ($F(1,24) = 16.17, p = .001, R^2_{adj} = 0.38$) than the former ($F(1,24) = 4.72, p = .04, R^2_{adj} = 0.16$). We report values adjusted for IQ where appropriate to resolve the variability in the pairwise

measures. The sources of the between pair differences in communicative success and pairwise alignment were further qualified with post-hoc comparisons using Fisher's least significant difference (LSD). The predictive strength of pairwise alignment on communicative success was determined using linear regression analysis. We report effect size estimates (partial η^2 and Cohen's d) and 95% confidence intervals for each ANOVA and post-hoc comparison to facilitate cumulative science (Lakens, 2013). We report Bayes Factors (BF) for statistical tests evaluating evidence in favor of the null hypothesis. Bayes Factors express the relative likelihood of the data under the models at hand.

A follow-up analysis tested for between-group differences in the dynamic relationship between pairwise alignment and problem ambiguity. The fixed effect of Group (ASD, Mixed, Typical) was assessed with a univariate ANOVA of Pearson's correlations between the two dependent variables and further qualified using post-hoc comparisons as in the main analysis. Problem ambiguity was measured as 1 minus the average pairwise alignment in pseudo-pairs, which is calculated in the same way as in the original pairs (Fig. 1C), and indicates consistent signal use across at least two consecutive trials by two members who are originally from different pairs, e.g., between individual 1 of pair A and individual 2 of pair B. This pairing produces a total of 650 possible pseudo-pair combinations between members of all 26 pairs. Averaging the alignment across all pseudo-pairs for each trial results in the trial-by-trial chance-distribution seen in Figure 1F. In this distribution, a trial has low ambiguity because pseudo-pairs were frequently aligned in their signal use for that problem, indicating consistent signal use across individuals regardless of communicative history. For instance, target configurations that did not require Addressees to rotate their shapes to reach a target orientation (e.g., circle-circle combinations presented in early interactions, as seen in Fig. 1F, trial 3) tended to evoke consistent solutions across pairs. On these trials, Communicators needed to distinguish Addressees' target locations from other locations visited on the game board in their movement, which several Communicators achieved by using a brief movement pause on the Addressees' target location (signal A). Conversely, high ambiguity trials were solved more easily with signals that had been negotiated and coordinated through pairs' unique communicative histories. For instance, problems involving mismatches between the shapes' geometrical characteristics and that require Addressees to rotate (e.g., circle-triangle combinations, in later interactions such as trial 39, shown in Fig. 1F) were often solved with different signals by different pairs (cf. signals E through S in Supplemental Table S2). Therefore, pseudo-pairs had lower alignment during these trials.

A network analysis provided further insight into how Typical and ASD pairs navigated their solution space over the course of 80 interactions. To visualize the distribution of individually and jointly visited solutions in each pair, networks consisting of two sets of edges (indicated by color, one for each member) followed the two pair members' signals at each trial. The blue edges represented player 1, and connected the nodes for odd trials (1, 3, 5, 7, and so on). The orange edges represented player 2, and connected the nodes for even trials (2, 4, 6, 8, and so on). All 80 trials, represented as nodes and connected by their Communicators' corresponding edges, were then clustered by signal type to visualize each player's solution space in respect to the other. A Fruchterman Reingold algorithm (Fruchterman & Reingold, 1991), implemented in Gephi, minimized the energy of the system by moving the nodes and changing the forces between them. This provided a two-dimensional layout of the pairwise trajectories through the pairs' solution spaces. Hamming distance, a metric used to measure differences in networks (Hamming, 1950), was calculated between pair members' solution spaces to show how much the members differed in their solution sets.

2.4. Additional analyses

Five additional analyses were used to assess the specificity of communicative success and pairwise alignment, beginning with two control analyses. First, we tested whether ASD-related differences in these variables were driven by generic motor-related differences between individuals in each group, using a multivariate ANOVA with a between-participants factor of Group (ASD, Typical) and the dependent variables of planning time, movement time, number of moves and time spent on target and other locations of the game board. Second, we tested whether the reduced pairwise alignment in ASD could be a consequence of this sample population generating a reduced number of intelligible solutions to the novel communicative problems, owing, for example, to perseveration being common in ASD (Geurts et al., 2009; Hill, 2004). To test this possibility, we calculated the number of distinguishable signals that were used in the interactions by each participant in the Communicator role, and tested for between-group differences using an independent samples t-test. We additionally examined whether individuals with ASD differed in the type of communicative signals they generated, using a Chi-squared test to assess the between-group overlap in signals used as a fraction of the total number of signals observed across both groups. An additional Kullback-Leibler divergence test was used to measure the difference between the two groups' signal frequency distributions. Third, we assessed whether the reduced conceptual alignment in ASD could be a

consequence of this sample population not being able or motivated to change their communicative signal following a misunderstanding during the last time they were the Communicator, by testing for between-group differences in signal changes (in interaction i) following an error (interaction i - 2) using an independent samples t-test. Fourth, we assessed whether between-group differences in communicative success and pairwise alignment could be attributed to differences in a motivation to communicate. To this end, we used a repeated-measures ANOVA to test for between-group differences (ASD, Typical) in the communicative emphasis participants in the Communicator role spontaneously placed on the Addressee's target location relative to other visited locations (Target, Non-target). These two analyses were adopted from other similar studies (Stolk, D'Imperio, di Pellegrino, & Toni, 2015; Stolk, Hunnius, Bekkering, & Toni, 2013; Stolk, Verhagen, et al., 2013). Fifth, we assessed whether communication impairment in ASD resulted selectively from an inability to align production and not interpretation of a behavior to a partner. For this exploratory analysis, we used a repeated-measures ANOVA to test for between-group differences (ASD, Typical) in communicative success in the role of Communicator and Addressee.

3. Results

3.1. Communication impairment

To quantify communicative abilities, we asked participant pairs to communicate over the course of 80 interactions, alternating between the roles of Communicator and Addressee. In any given interaction, the Communicator and Addressee must recreate a spatial configuration of two assigned shapes on the digital game board. The target configuration is only shown to the Communicator (event II in Fig. 1A), who must use his or her own assigned shape to relay to the Addressee his or her target location (event III, hereafter referred to as a 'communicative signal'). The Communicator must then reach his or her own target position. The Addressee, who has not seen the target configuration, must then infer his or her target position based on the Communicator's communicative signal and move accordingly (event IV). After the Addressee has moved, the same feedback is presented to both players to indicate communicative success, i.e., whether or not they jointly reproduced the target configuration (event V). Fifty-two adult participants were recruited to take part in this study: 22 individuals with ASD and 30 neurotypical individuals, matched on gender, age, and IQ. Participants were assigned pairwise to either the ASD group (7 pairs), the Typical group (11 pairs), or the Mixed group (8 pairs, each including one individual with ASD and one individual with no clinical diagnosis). Confirming the

study's predictions, the three groups differed in their overall communicative success ($F(2,22) = 11.40, p < .001, \text{partial } \eta^2 = 0.51$). As seen in Figure 1B, this effect was driven by pairs containing one or more individuals with ASD (ASD and Mixed pairs) successfully solving fewer communicative problems than Typical pairs (Typical versus ASD, $p < .001, d = 2.41, 95\% \text{ CI} = [1.18 \ 3.65]$; Typical versus Mixed, $p = .048, d = 1.06, 95\% \text{ CI} = [0.09 \ 2.04]$; Mixed versus ASD, $p = .047, d = 1.27, 95\% \text{ CI} = [0.16 \ 2.38]$), providing quantitative evidence for communication impairment as a core diagnostic feature of ASD (American Psychiatric Association, 2013).

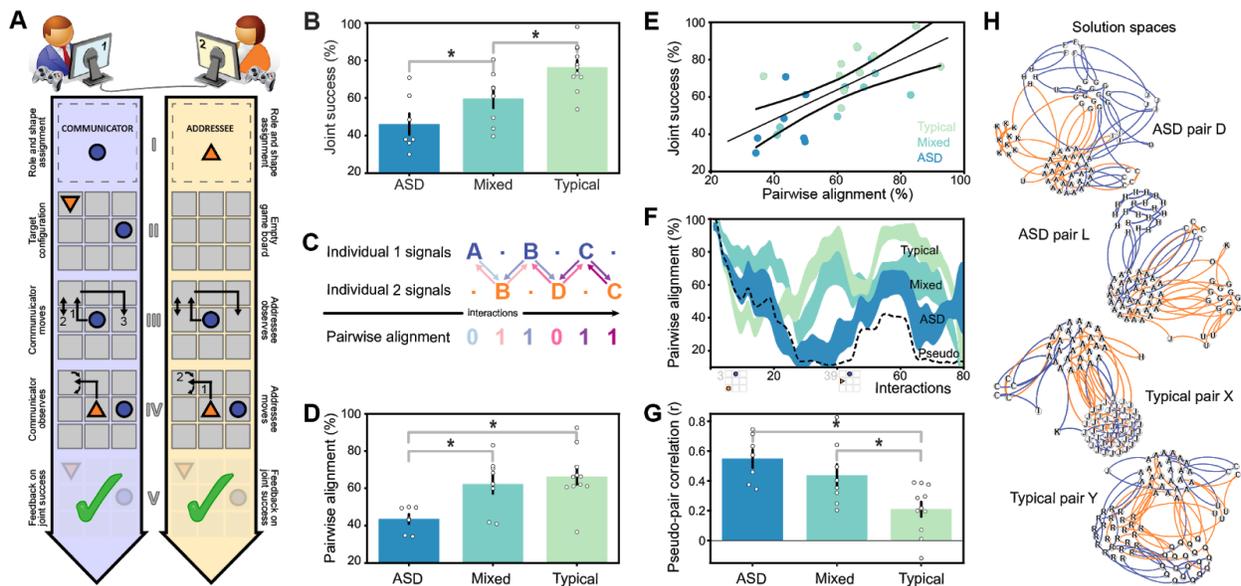


Fig. 1 - Communication impairment and misalignment in ASD. (A) The joint goal of Communicator and Addressee is to reproduce a target configuration of their two given shapes on a digital game board, shown only to the Communicator (event II). Given that the Addressee cannot see the target configuration, a successful interaction requires the Communicator and Addressee to construct and comprehend how the behavior of the Communicator's shape indicates the target location and orientation of the Addressee's shape (a 'communicative signal', event III). (B) Pairs containing one or more individuals diagnosed with ASD (ASD and Mixed pairs) jointly solved fewer communicative problems than pairs containing two individuals with no clinical diagnosis (Typical pairs). (C and D) Pairwise alignment measured the percentage of interaction-by-interaction overlap in signals produced by pair members, and revealed less frequent dynamic alignment of communicative signal use in ASD pairs than in the other pair types. Pairs were considered aligned at a given interaction if the type of communicative signal used in that interaction was identical to the signal in the interactions directly preceding or following it, when the other member of the pair was the Communicator. Letter indices correspond to a descriptive list of observed signals

(Supplemental Table S2). **(E)** Pairwise alignment predicted communicative success across pairs. Lines show linear fit and mean 95% confidence intervals. **(F and G)** Pairwise alignment was also calculated for pseudo-pairs, constituted of members from different pairs. Given that pseudo-pair members lack any interpersonal coordination, their distribution of trial-by-trial alignment (dashed black line) provided a quantitative index of problem ambiguity, where low ambiguity indicated high pseudo-pair alignment because those configurations evoked consistent solutions across different pairs, cf. interactions 3 and 39. The bar graph indicates that ASD and Mixed pairs' alignment tended to follow the distribution of pairwise alignment found in the pseudo-pairs, whereas Typical pairs achieved alignment even when the problem space afforded multiple solutions. Lines' spread indicates ± 1 SEM. A version of the line graph with trials ordered from lowest to highest ambiguity can be seen in Supplemental Figure S1. **(H)** Network visualizations of four pairs' trajectories through their idiosyncratic solution spaces. The nodes represent the signals constructed by the Communicator of each interaction and are clustered to show signals that were used repeatedly. The colored edges connect each individual's consecutively produced signals as Communicators (blue for individual 1, orange for individual 2). ASD pairs showed more individual exploration than Typical pairs of their solution space, as indicated by relatively large clusters of individually-visited solutions (clusters of nodes that were connected only by edges of a single color) and small clusters of jointly-visited solutions.

3.2. Communicative misalignment

The task is primarily designed to quantify and manipulate interpersonal alignment in communication. Similar to how idiosyncratic shared constructs emerge from everyday dialogue (S. E. Brennan & Clark, 1996), this task allows pairs to converge on unique meanings for the same behavior. For instance, some pairs solve the communicative problem illustrated in Figure 1A by stepping in and out multiple times from the Addressee's target location, to indicate the direction in which the Addressee's shape needs to point. Other pairs use the exact same communicative signal to indicate the number of rotations the Addressee needs to apply to his/her shape to obtain the desired orientation (Supplemental Table S2 provides a list of identified signals). These examples also illustrate how the meaning of the ambiguous signals produced in this game cannot be determined from a single interaction. A pair has converged on a shared meaning for a signal only if both individuals manage to comprehend and reproduce that signal successfully. Therefore, pairs were considered aligned during a given interaction if its Communicator used a communicative signal that was conceptually identical to the signal used in the interactions directly preceding or following it, when the other member of the pair was the Communicator (Fig. 1C). This trial-by-trial alignment is analogous to two communicators using

an ambiguous word or gesture that only they would know the exact meaning of. Their ability to produce and understand this behavior is proof of their pair-specific conceptual agreement.

To manipulate alignment over the course of the task, target configurations with multiple possible solutions were introduced in a consistent and deliberate order. This experimental manipulation makes it possible to calculate a trial-by-trial chance-distribution of pseudo-pairs' communicative alignment, containing for each trial the average pairwise alignment of communicative signals between members of different pairs (dashed black line in Fig. 1F). Given that pseudo-pair members lack any interpersonal coordination, their distribution of trial-by-trial alignment allows us to quantitatively differentiate between configurations which evoke consistent solutions across pairs and configurations solved with different signals by different pairs (cf. interactions 3 and 39 in Fig. 1F). We took these differences in the solution space to indicate problem ambiguity, where high ambiguity equated to a larger set of empirically observed signals used to solve that problem. We predicted problem ambiguity to strongly modulate alignment in pairs containing one or more individuals with ASD, consistent with their inability to establish conceptual alignment.

The main finding of this study quantitatively illustrates that reduced conceptual alignment explains communicative impairment in ASD. Comparison of signals used across contiguous interactions showed differing alignment across the three groups ($F(2,22) = 9.07, p = .001$, partial $\eta^2 = 0.45$). As shown in Figure 1D and Movies [S1](#) and [S2](#), full ASD pairs aligned their communicative signals across contiguous interactions less frequently than the other pair types (Typical versus ASD, $p < .001, d = 2.11, 95\% \text{ CI} = [0.94 \text{ } 3.29]$; Mixed versus ASD, $p = .003, d = 2.03, 95\% \text{ CI} = [0.78 \text{ } 3.28]$; Typical versus Mixed, $p = .99, d = 0.01, 95\% \text{ CI} = [-0.90 \text{ } 0.92]$). This analysis leads to two independent observations about conceptual alignment. First, a correlation analysis showed a strong relationship between alignment and success across all pairs, confirming conceptual misalignment as a predictor of communicative impairment ($F(1,24) = 32.79, p < .001, R^2_{adj} = 0.56, r = 0.76, 95\% \text{ CI} = [0.53 \text{ } 0.89]$), see Figure 1E. Second, full- and mixed-ASD pairs largely followed the distribution of communicative alignment found in pseudo-pairs, during both unambiguous and ambiguous problems (Fig. 1F and G). This observation indicates that the interaction dynamics of full- and mixed-ASD pairs were comparable to those of pairs of individuals with no previous communication. In contrast, Typical pairs achieved higher alignment than pseudo-pairs even during ambiguous problems, indicating that individuals in Typical pairs capitalized on their communicative history. These between-group differences are supported by a statistical analysis testing for the effect of

problem ambiguity on pairwise alignment ($F(2,23) = 8.51, p = .002, \text{partial } \eta^2 = 0.43$; Typical versus ASD, $p = .001, d = 2.02, 95\% \text{ CI} = [0.86 \ 3.17]$; Typical versus Mixed, $p = .012, d = 1.34, 95\% \text{ CI} = [0.33 \ 2.34]$; Mixed versus ASD, $p = .24, d = 0.68, 95\% \text{ CI} = [-0.36 \ 1.73]$).

The group-based observations are also supported by descriptive analyses of pairwise trajectories through their idiosyncratic solution spaces, where pair members' signals at every interaction are represented as nodes in a network (Fig. 1H). These networks allowed us to follow individual pair members' behaviors in relation to one another, and visualize patterns of exploration and alignment through the solution space. It can be seen that in Typical pairs, both individuals frequently used the same signal (e.g., signals J and C in pair X; signals R and Q in pair Y), while occasionally exploring other solutions (e.g., signals K, H, and I in pair X; signal J in pair Y). However, those options were not pursued further if the other pair member did not adopt those signals. The large size of shared clusters, and the point-like structure of the individually-visited solutions, confirm that Typical pairs develop strong conceptual agreements across the 80 interactions. In contrast, ASD pairs show more individually-visited solutions and small clusters of jointly-visited solutions, the exception being large clusters of signal A that was used predominantly during problems requiring no rotation by the Addressee (e.g., circle-circle problem configurations). As seen through the spread of behaviors by each member, these pairs were able to explore signals, but unlike Typical pairs, struggled to converge on shared signals. This is supported by an analysis comparing the dissimilarity between signal sets produced by members of ASD and neurotypical pairs (Hamming distance: 0.67 ± 0.04 vs. $0.50 \pm 0.04, M \pm \text{SEM}, t(16) = 2.84, p = .012, d = 1.46, 95\% \text{ CI} = [0.40 \ 2.52]$). Thus, members of ASD pairs produced signals that were largely independent from those of their partners, and failed to reach conceptual alignment.

3.3. Preserved cognitive abilities and communicative propensities

The experimental setting and additional empirical observations exclude several proposed causes for the communicative impairment observed in individuals with ASD. First, planning time, movement time, number of moves, time spent on target and other locations of the game board were consistently matched between neurotypical and individuals with ASD (Fig. 2A, $F(5,46) = 2.00, p = .10, \text{partial } \eta^2 = 0.18$). This finding indicates that the communicative impairment is not a consequence of the sensorimotor demands of the task, nor of misapprehension of its communicative demands. Second, individuals with and without ASD exhibited a comparable heterogeneity of communicative behaviors throughout their interactions (Fig. 2A, 5.41 ± 0.35 vs.

5.07 ± 0.24 distinguishable signals in each participant, $M \pm SEM$, $t(50) = 0.84$, $p = .40$, $d = 0.29$, 95% CI = [-0.27 0.84], $BF = 2.66$ in favor of the null hypothesis; 13 out of 18 identical signals across both groups, $\chi^2(1,18) = 7.1$, $p = .008$, $d = 0.44$, 95% CI = [0.15 0.74]; Kullback-Leibler divergence of signal frequency distributions = 0.016), indicating that the communicative misalignment was not due to perseveration commonly associated with ASD. Moreover, individuals in both groups were similarly inclined to re-use signals that were previously understood by their partner (Fig. 2A, Kullback-Leibler divergence of success rate distributions = 0.029), showing that misalignment persisted despite an intact ability to remember and learn from successful interactions. Furthermore, individuals with ASD were also similarly inclined to change their signal following a communicative failure (Fig. 2B, $t(50) = 1.28$, $p = .21$, $d = 0.37$, 95% CI = [-0.19 0.92], $BF = 1.83$ in favor of the null hypothesis), excluding the possibility that reduced success or misalignment was due to a failure to detect or react to a misunderstanding with their partner. Third, similar to neurotypical individuals, individuals with ASD spontaneously spent more time on communicatively relevant locations of the game board (Fig. 2C, main effect of location $F(1,50) = 188.88$, $p < .001$, partial $\eta^2 = 0.79$; no main effect of group or interaction involving ASD and location, both $p > .15$), indicating that they were willing to invest resources in switching between communicative and instrumental portions of their actions to mark a communicatively relevant location for the benefit of their partner. Finally, communicative success was lowest when individuals with ASD were Addressees (interaction effect of group and role, $F(1,50) = 11.31$, $p = .001$, partial $\eta^2 = 0.18$; no main effect of role, $p = .61$; Typical vs. ASD Communicator success, $t(50) = 1.80$, $p = .08$, $d = 0.52$, 95% CI = [-0.04 1.07]; Typical vs. ASD Addressee success, $t(50) = 3.26$, $p = .002$, $d = 0.94$, 95% CI = [0.36 1.51]). Taken together with their misalignment, this finding suggest that individuals with ASD struggled to both produce and comprehend behaviors in light of the context of interaction.

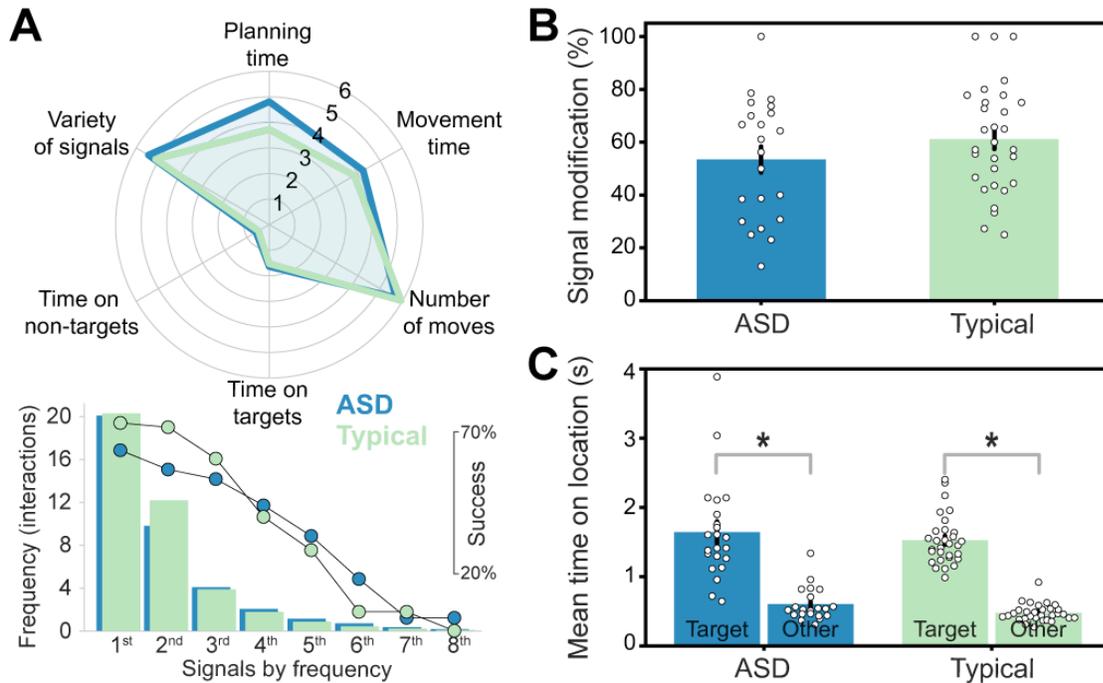


Fig. 2 - Preserved cognitive abilities and communicative propensities in ASD. (A) Planning time, movement time, the number of moves, time spent on target and other locations of the game board were consistently matched between neurotypical and autistic individuals. Individuals with and without ASD also exhibited a comparable heterogeneity of communicative signals throughout the task, and were similarly inclined to re-use signals that were previously understood by their partner as indicated by the comparable slopes of the frequency and success distributions. Times are measured in seconds. (B) Individuals with and without ASD showed a similar propensity for modifying their behavior as Communicators if not understood previously. (C) Similar to neurotypical individuals, individuals with ASD spontaneously spent more time on communicatively relevant locations of the game board as Communicators, discriminating them from other visited locations for the benefit of their communicative partner.

4. Discussion

The findings reported in this study demarcate a key cognitive challenge intrinsic to human interpersonal communication that individuals with ASD struggle to overcome, despite having otherwise indistinguishable performance from neurotypical adults across several task metrics. This study shows that communicative impairments in ASD are not simply a consequence of neglect of communicative demands during interaction nor of altered sensory processing, motor performance, interaction memory, social motivation or attention, or cognitive perseveration. As the communicative demands of the current task prohibit recourse to known biological and

linguistic cues, ASD communicative impairment was also unaffected by altered processing of those stimuli (Constantino et al., 2017; J. Cook et al., 2009; Hobson et al., 1988; Hutchins & Brien, 2016; Nackaerts et al., 2012; Silverman et al., 2010). Furthermore, individuals with ASD showed comparable ability and motivation to neurotypical individuals in producing intelligible communicative behaviors. They even modulated the use of these behaviors based on their partners' responses, questioning suggestions of universally diminished social motivation or impaired cognitive flexibility in ASD (Chevallier et al., 2012; Geurts et al., 2009).

Yet, individuals with ASD struggled to align the conceptualizations of their communicative signals with those of their interaction partners when the problem space afforded multiple solutions. This impairment could be isolated because the novel communicative setting prevented access to pre-existing contextual cues that cognitively-able individuals with ASD can capitalize on to resolve ambiguity (Au-Yeung, Kaakinen, Liversedge, & Benson, 2015; Birmingham, Stanley, Nair, & Adolphs, 2015; Branigan, Tosi, & Gillespie-Smith, 2016; Brewer, Biotti, Bird, & Cook, 2017; Hahn, Snedeker, & Rabagliati, 2015; Nadig, Seth, & Sasson, 2015; Pijnacker, Hagoort, Buitelaar, Teunisse, & Geurts, 2009). Under these experimentally-generated conditions, built to recreate the fleeting ambiguities of everyday interaction, communication requires more than pruning a decision tree of possible signals or iteratively optimizing behavioral outcomes (Botvinick & Weinstein, 2014; Donoso, Collins, & Koechlin, 2014; Keysers & Perrett, 2004). Despite their ability to consistently produce, modify, and remember interpretable communicative behaviors, i.e. behaviors that were also part of the neurotypical solution space, autistic individuals were less likely to select signals in light of their partner's behaviors. This resulted in greater misalignment, especially when full- and mixed-ASD pairs were presented with ambiguous problems. In contrast, neurotypical individuals navigated through epochs of communicative ambiguity by considering and aligning to their partner's recent signals. The observed differences in task performance between neurotypical and full- and mixed-ASD pairs arose specifically from this reduced ability to produce and comprehend communicative behaviors informed by their recent communicative history. This observation explains why autistic individuals are vulnerable to the transient and interaction-specific ambiguities in everyday social situations, and provides novel boundaries to the general notion that ASD is linked to altered mentalizing abilities (Baron-Cohen et al., 1985; Happé, 1993; Schaafsma, Pfaff, Spunt, & Adolphs, 2015). Although adults with ASD can communicate on the basis of presumed knowledge about a generic partner, they fail to dynamically update that conceptual knowledge according to the ongoing interaction with a specific individual. Precise

characterization of dynamic conceptual updating might provide a new window into understanding autistic communication, and how signals derive their meaning from the communicative context in which they are embedded.

Failures in using the conceptual space implied by the ongoing interaction are likely to have two important consequences. First, those failures might affect the recognition of the relevance of a communicative signal for jointly coordinating the shared conceptual space. Second, those failures might affect the resolution of the ambiguity intrinsic in those signals. It remains to be seen whether other social deficits observed in ASD (and controlled for in this study), such as abnormalities in eye contact, facial expressions, speech, and turn-taking (Madipakkam, Rothkirch, Dziobek, & Sterzer, 2017; Shriberg et al., 2001; Tager-Flusberg & Anderson, 1991), could in fact be downstream consequences of difficulties in predicting and monitoring mutual understanding (Stolk et al., 2016). It will also be of interest to know whether and how conceptual alignment deficits interact with cognitive traits and environmental factors to give rise to the considerable behavioral and developmental variability observed in ASD, opening the way for principled interventions to improve communication between autistic and neurotypical individuals (Edey et al., 2016; Fusaroli, Weed, Fein, & Naigles, 2018; Greenberg, Warrier, Allison, & Baron-Cohen, 2018; Perry, Levy-Gigi, Richter-Levin, & Shamay-Tsoory, 2015; Stolk, Hunnius, et al., 2013; Stolk, Noordzij, Volman, et al., 2014). This work illustrates that to answer these key questions, it is both feasible and pertinent to study autistic individuals engaged in social interactions with others. This is the natural context in which communication is learned, where it is used, and where individuals with ASD experience difficulties.

5. Conclusion

This study provides a novel and precise characterization of communicative deficits in ASD, one of its core diagnostic features (American Psychiatric Association, 2013). The results suggest that individuals with ASD and neurotypical individuals would be equally able and motivated to communicate if human communication could be reduced to an information transfer problem in a signal encoding–decoding framework (Shannon, 1948). Yet, the ASD communicative deficits observed here indicate that human communication is best characterized as a solution to a conceptual alignment challenge, organized to predict and monitor mutual understanding (Stolk et al., 2016). This study illustrates how the efficacy of the evolutionarily anomalous human communicative system is severely limited without this key interactional ingredient.

Author Contributions. R.B., G.B., I.T., and A.S. conceived and designed the study. R.B. and G.B. conducted the study and collected the data. H.W., I.T., and A.S. conducted the data analyses. H.W., I.T., and A.S. drafted the manuscript, and R.B. and G.B. provided critical revisions. All authors approved the final version of the manuscript.

Acknowledgments. This work was supported by a Rubicon grant from the Dutch Organization for Scientific Research NWO (446-14-007) and a Marie Skłodowska-Curie Global Fellowship from the European Union (658868) awarded to A.S. We are grateful to the study participants and Loes Bijl for helping with the classification of behaviors.

Data and Code Availability. The data and analysis code are available from the corresponding author upon request.

References

- Akmajian, A., Farmer, A. K., Bickmore, L., Demers, R. A., & Harnish, R. M. (2017). *Linguistics: An Introduction to Language and Communication*. MIT Press.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Au-Yeung, S. K., Kaakinen, J. K., Liversedge, S. P., & Benson, V. (2015). Processing of Written Irony in Autism Spectrum Disorder: An Eye-Movement Study. *Autism Research: Official Journal of the International Society for Autism Research*, 8(6), 749–760.
- Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism-Spectrum Quotient (AQ)--adolescent version. *Journal of Autism and Developmental Disorders*, 36(3), 343–350.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), 37–46.
- Birmingham, E., Stanley, D., Nair, R., & Adolphs, R. (2015). Implicit Social Biases in People With Autism. *Psychological Science*, 26(11), 1693–1705.
- Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & van Rooij, I. (2012). Recipient design in human communication: simple heuristics or perspective taking? *Frontiers in Human Neuroscience*, 6. <https://doi.org/10.3389/fnhum.2012.00253>
- Botvinick, M., & Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1655). <https://doi.org/10.1098/rstb.2013.0480>
- Bowler, D. M. (1992). “Theory of Mind” in Asperger’s Syndrome Dermot M. Bowler. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 33(5), 877–893.
- Branigan, H. P., Tosi, A., & Gillespie-Smith, K. (2016). Spontaneous lexical alignment in

- children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 42(11), 1821–1831.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 22(6), 1482–1493.
- Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two Minds, One Dialog. In *Psychology of Learning and Motivation* (pp. 301–344).
- Brewer, R., Biotti, F., Bird, G., & Cook, R. (2017). Typical integration of emotion cues from bodies and faces in Autism Spectrum Disorder. *Cognition*, 165, 82–87.
- Chambon, V., Farrer, C., Pacherie, E., Jacquet, P. O., Leboyer, M., & Zalla, T. (2017). Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. *Cognition*, 160, 17–26.
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4), 231–239.
- Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., ... Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*, 547(7663), 340–344.
- Cook, J. L., Blakemore, S.-J., & Press, C. (2013). Atypical basic movement kinematics in autism spectrum conditions. *Brain: A Journal of Neurology*, 136(Pt 9), 2816–2824.
- Cook, J., Saygin, A. P., Swain, R., & Blakemore, S.-J. (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia*, 47(14), 3275–3278.
- Cusack, J. P., Williams, J. H. G., & Neri, P. (2015). Action Perception Is Intact in Autism Spectrum Disorder. *Journal of Neuroscience*, 35(5), 1849–1857.
- de Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Newman-Norlund, R., Hagoort, P.,

- Levinson, ... Toni, I. (2010). Exploring the cognitive infrastructure of communication. *Interaction Studies*, 11(1), 51–77.
- Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191), 1481–1486.
- Eco, U. (1976). *A Theory of Semiotics*. Bloomington: Indiana University Press.
- Edey, R., Cook, J., Brewer, R., Johnson, M. H., Bird, G., & Press, C. (2016). Interaction takes two: Typical adults exhibit mind-blindness towards those with autism spectrum disorder. *Journal of Abnormal Psychology*, 125(7), 879–885.
- Fisher, R. A. (1925). *Statistical Methods For Research Workers*. Genesis Publishing Pvt Ltd.
- Frith, U. (2003). *Autism: Explaining the Enigma*. Wiley-Blackwell.
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice & Experience*, 21(11), 1129–1164.
- 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. (2018). Hearing me hearing you: Reciprocal effects between child and parent language in autism and typical development. *Cognition*, 183, 1–18.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181–218.
- Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences*, 13(2), 74–82.
- Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, 20(11), 818–829.

- Greenberg, D. M., Warrier, V., Allison, C., & Baron-Cohen, S. (2018). Testing the Empathizing-Systemizing theory of sex differences and the Extreme Male Brain theory of autism in half a million people. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), 12152–12157.
- Grice, H. P. (1975). Logic and conversation. In C. P. Morgan JL (Ed.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). Academic Press.
- Groen, W. B., Zwiers, M. P., van der Gaag, R.-J., & Buitelaar, J. K. (2008). The phenotype and neural correlates of language in autism: an integrative review. *Neuroscience and Biobehavioral Reviews*, *32*(8), 1416–1425.
- Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid Linguistic Ambiguity Resolution in Young Children with Autism Spectrum Disorder: Eye Tracking Evidence for the Limits of Weak Central Coherence. *Autism Research: Official Journal of the International Society for Autism Research*, *8*(6), 717–726.
- Hamming, R. W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, *29*(2), 147–160.
- Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*(2), 101–119.
- Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*(1), 26–32.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology*, *79*(4), 441–453.
- Hutchins, T. L., & Brien, A. (2016). Conversational topic moderates social attention in autism spectrum disorder: Talking about emotions is like driving in a snowstorm. *Research in Autism Spectrum Disorders*, *26*, 99–110.
- Jakobson, R. (1971). *Selected Writings: Word and Language*. Walter de Gruyter.

- Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, 8(11), 501–507.
- Klin, A., McPartland, J., & Volkmar, F. R. (2013). Asperger Syndrome. In *Handbook of Autism and Pervasive Developmental Disorders* (pp. 88–125).
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Levinson, S. C. (1983). *Pragmatics*. (M. Hattaway, Ed.). Cambridge University Press.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)*.
- Madipakkam, A. R., Rothkirch, M., Dziobek, I., & Sterzer, P. (2017). Unconscious avoidance of eye contact in autism spectrum disorder. *Scientific Reports*, 7(1), 13378.
- Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, 18(10), 512–519.
- Nackaerts, E., Wagemans, J., Helsen, W., Swinnen, S. P., Wenderoth, N., & Alaerts, K. (2012). Recognizing biological motion and emotions from point-light displays in autism spectrum disorders. *PloS One*, 7(9), e44473.
- Nadig, A., Seth, S., & Sasson, M. (2015). Global Similarities and Multifaceted Differences in the Production of Partner-Specific Referential Acts by Adults with Autism Spectrum Disorders. *Frontiers in Psychology*, 6, 1888.
- Pantelis, P. C., & Kennedy, D. P. (2017). Autism does not limit strategic thinking in the “beauty contest” game. *Cognition*, 160, 91–97.
- Perry, A., Levy-Gigi, E., Richter-Levin, G., & Shamay-Tsoory, S. G. (2015). Interpersonal distance and social anxiety in autistic spectrum disorders: A behavioral and ERP study.

Social Neuroscience, 1–12.

- Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J.-P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 39(4), 607–618.
- Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: individual differences and IQ. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 17(5), 888–896.
- Rumelhart, D. E. (1979). Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 71–82). Cambridge University Press.
- Sally, D., & Hill, E. (2006). The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *Journal Of Economic Psychology*, 27(1), 73–97.
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, 19(2), 65–72.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(4), 623–656.
- Shriberg, L. D., Paul, R., McSweeney, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research: JSLHR*, 44(5), 1097–1115.
- Silverman, L. B., Bennetto, L., Campana, E., & Tanenhaus, M. K. (2010). Speech-and-gesture integration in high functioning autism. *Cognition*, 115(3), 380–393.
- Sperber, D., & Wilson, D. (1996). *Relevance: Communication and Cognition*. Wiley-Blackwell.
- Stolk, A., D'Imperio, D., di Pellegrino, G., & Toni, I. (2015). Altered communicative decisions following ventromedial prefrontal lesions. *Current Biology*, 25(11), 1469–1474.

- Stolk, A., Hunnius, S., Bekkering, H., & Toni, I. (2013). Early social experience predicts referential communicative adjustments in five-year-old children. *PLoS One*, *8*(8), e72667.
- Stolk, A., Noordzij, M. L., Verhagen, L., Volman, I., Schoffelen, J.-M., Oostenveld, R., ... Toni, I. (2014). Cerebral coherence between communicators marks the emergence of meaning. *Proceedings of the National Academy of Sciences*, *111*(51), 18183–18188.
- Stolk, A., Noordzij, M. L., Volman, I., Verhagen, L., Overeem, S., van Elswijk, G., ... Toni, I. (2014). Understanding communicative actions: a repetitive TMS study. *Cortex*, *51*, 25–34.
- Stolk, A., Verhagen, L., Schoffelen, J.-M., Oostenveld, R., Blokpoel, M., Hagoort, P., ... Toni, I. (2013). Neural mechanisms of communicative innovation. *Proceedings of the National Academy of Sciences*, *110*(36), 14574–14579.
- Stolk, A., Verhagen, L., & Toni, I. (2016). Conceptual Alignment: How Brains Achieve Mutual Understanding. *Trends in Cognitive Sciences*, *20*(3), 180–191.
- Tager-Flusberg, H., & Anderson, M. (1991). The Development of Contingent Discourse Ability in Autistic Children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *32*(7), 1123–1134.
- Tesink, C. M. J. Y., Buitelaar, J. K., Petersson, K. M., van der Gaag, R. J., Kan, C. C., Tendolkar, I., & Hagoort, P. (2009). Neural correlates of pragmatic language comprehension in autism spectrum disorders. *Brain: A Journal of Neurology*, *132*(Pt 7), 1941–1952.
- Volman, I., Noordzij, M. L., & Toni, I. (2012). Sources of variability in human communicative skills. *Frontiers in Human Neuroscience*, *6*. <https://doi.org/10.3389/fnhum.2012.00310>
- Wechsler, D. (2008). Wechsler Adult Intelligence Scale--Fourth Edition. *PsycTESTS Dataset*. <https://doi.org/10.1037/t15169-000>
- Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence--Second Edition. *PsycTESTS*

Dataset. <https://doi.org/10.1037/t15171-000>

Zalla, T., Amsellem, F., Chaste, P., Ervas, F., Leboyer, M., & Champagne-Lavau, M. (2014).

Individuals with autism spectrum disorders do not use social stereotypes in irony comprehension. *PloS One*, 9(4), e95568.

Table S1 - Autism Quotient (AQ) score and Autism Diagnostic Observational Schedule (ADOS) classification for twenty-two participants diagnosed with ASD. Scores on the Social Interaction and Communication modules of the ADOS were totalled in their ADOS score, and all three scores were then used to classify the participants. Participants were classified as follows; *Autism*: a Communication score of 3+, Social Interaction score of 6+, and a total score of 10+; *Autism Spectrum*: a Communication score of 2+, Social Interaction score of 4+, and a total score of 7+; *None*: if any of the criteria for Autism Spectrum were not met. These individuals were included based on an independent diagnosis by a clinician and reported a similar degree of autistic traits on the AQ as the individuals who did meet the ADOS criteria.

Pair	AQ	ADOS	Social Interaction	Communication	Classification
A	39	9	7	2	Autism Spectrum
A	41	6	4	2	None
B	42	10	6	4	Autism
B	35	10	5	5	Autism Spectrum
C	36	8	6	2	Autism Spectrum
C	19	11	7	4	Autism
D	29	2	1	1	None
D	37	11	6	5	Autism
E	31	9	6	3	Autism Spectrum
E	36	6	5	1	None
F	35	9	6	3	Autism Spectrum
F	28	7	3	4	None
I	19	10	5	5	Autism Spectrum
J	22	9	6	3	Autism Spectrum
K	46	8	6	2	Autism Spectrum
L	39	1	1	0	None
L	46	7	5	2	Autism Spectrum
M	23	9	7	2	Autism Spectrum
N	28	11	8	3	Autism
O	27	5	5	0	None
P	21	8	6	2	Autism Spectrum
Q	37	4	3	1	None

Table S2 - Communicative signals devised and used by Communicators for communicating the Addressee's target location and orientation. The descriptions do not include completion of the Communicator's own target location and orientation, nor their effectiveness for the various problem types. Signal L was not observed in the current sample population, whereas signals B, Q and R had not been reported before. This table is not a comprehensive list of all possible signals for this task, but only a compilation of observed signals to date. Signal occurrences are calculated on the entire sample population.

Communicative signal	Description	Occurrence
A - pause	The Communicator spends more time at the Addressee's target location than at any other visited location on the game board. Used by all pairs during the first four interactions, which involved target configurations consisting of identical geometric shapes across the players (e.g., circle-circle combinations) that did not require any rotation.	37.4%
B - prolonged pause	The Communicator pauses on the Addressee's target location for periods of time proportionate to the number of rotations the Addressee needs to make.	2.9%
C - match	The Communicator matches the Addressee's target orientation by rotating in place at the Addressee's target location (the players' shapes are identical).	5.5%
D - rotate	The Communicator goes to the Addressee's target location and rotates in place however many times required for the Addressee to reach the target orientation.	0.7%
E - entry	The Communicator enters the Addressee's target location from the direction the Addressee's shape should be pointing.	0.1%
F - exit	The Communicator exits the Addressee's target location along the direction the Addressee's shape should be pointing.	1.7%
G - line	The Communicator steps out multiple steps from and returns to the Addressee's target location to indicate the direction the Addressee's shape should be pointing.	4.5%
H - single wiggle from target	The Communicator steps in and out of the Addressee's target location in the direction of the Addressee's target orientation.	10.0%
I - multiple wiggles from target	The Communicator steps in and out of the Addressee's target location multiple times in the direction of the Addressee's target orientation (more emphatic version of H).	3.8%
J - wiggle count from target	The Communicator goes to the Addressee's target location and steps in and out of that location however many times the Addressee needs to rotate to reach the target orientation.	16.0%

K - wiggle count from target neighbor	The Communicator goes to a square adjacent to the Addressee's target location and steps into and out of the Addressee's target location however many times the Addressee needs to rotate to reach the target orientation. This signal will result in one less visit to the target location than J.	1.2%
L - wiggle from center	The Communicator visits the Addressee's target location followed by the game board's center, and steps in and out of the central location however many times the Addressee needs to rotate to reach the target orientation.	0.0%
M - exit from center	The Communicator exits the central start location along the direction the Addressee's shape should be pointing, before going to the Addressee's target location.	2.0%
N - wiggle elsewhere	The Communicator steps out in the direction of the Addressee's target orientation at a game board location other than the Addressee's target location or the central location.	0.2%
O - circle target count	The Communicator goes to the Addressee's target location and circles around it however many times the Addressee needs to rotate.	5.3%
P - circle target direction	The Communicator goes to the Addressee's target location and circles around it in the direction of the Addressee's target orientation.	0.9%
Q - circle board direction	The Communicator goes to the Addressee's target location and circles along the border of the entire board in the direction of the Addressee's target orientation.	1.0%
R - circle board count	The Communicator goes to the Addressee's target location and circles along the border of the entire board however many times the Addressee needs to rotate.	1.3%
S - draw	The Communicator uses a large section of the game board to sketch the Addressee's overall target configuration.	2.1%
U - unclassified	Signals that could not be classified or involved procedural errors by the Communicator.	3.6%

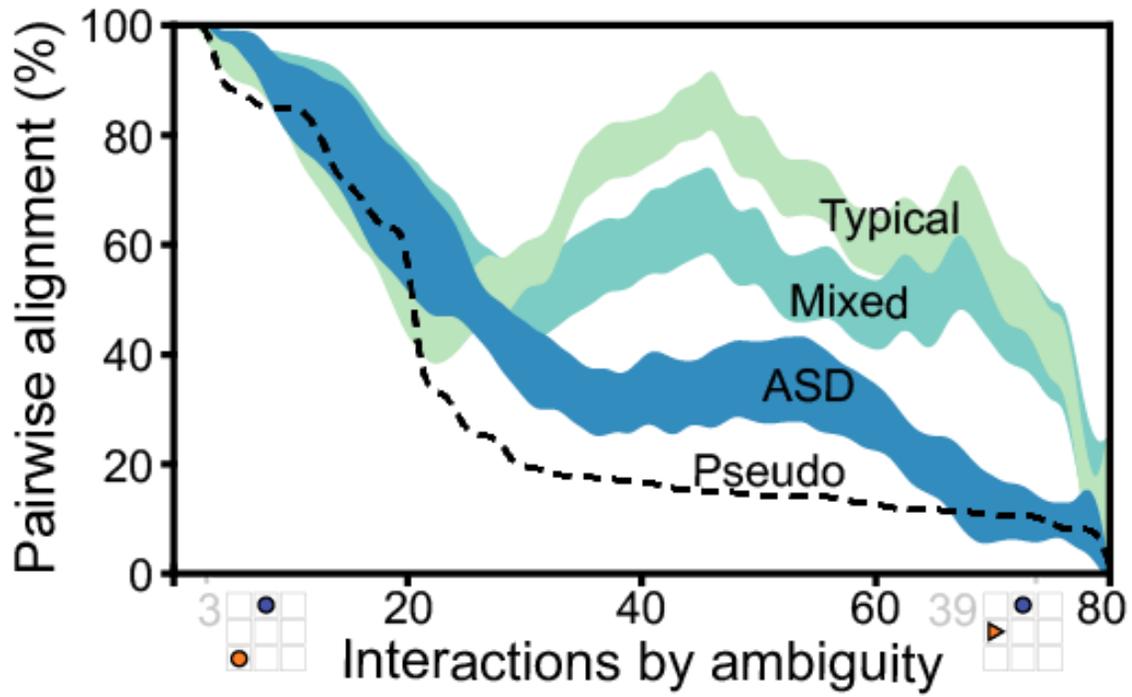


Fig. S1 - Replication of Figure 1F with trials ordered from lowest to highest ambiguity to better illustrate the relationship between pairwise alignment and problem ambiguity in each group.

Movie S1 - Communicative alignment in a neurotypical pair. This movie features 3 interactions illustrating how neurotypical individuals converge on a shared meaning of a communicative behavior in this game. In trial 27, the blue player, who is the Communicator of that interaction, steps out from the orange player's target location repeatedly in the direction the orange player's triangle should be pointing (signal I). The orange player, however, interprets this behavior differently from how it was intended. She rotates her triangular token twice, corresponding to the number of times the blue player stepped out from the presumed target location (signal J), thereby ending up facing in a direction different from the target orientation. During her own turn as a Communicator in trial 28, the orange player expands on her previous (incorrect) interpretation by visiting, with ostensible emphasis, the blue player's target location three times, which is however many times the blue player needs to rotate to reach the target orientation with her triangle (signal J). Arguably informed by both her partner's previous interpretation and current emphatic behavior, this proposition is rapidly understood and agreed upon by the blue player who correctly interprets the signal (J) in trial 28 and decides to use it at her turn as a Communicator in trial 29. That is, in trial 29, she steps out from the orange player's target location once to indicate to the orange player that she needs to rotate her triangle once (signal J). The pair continues to successfully apply this signal until trial 80 wherever possible (not shown in the video).

Movie S2 - Communicative misalignment in an ASD pair. This movie illustrates how cognitively able individuals diagnosed with ASD (IQs of 135 and 118) struggle to align their communication with one another over the course of 5 interactions, both when selecting their own and interpreting the other's behavior. In trial 34, the orange player, who is the Communicator of that interaction, visits the blue player's target location however many times the blue player needs to rotate to reach the target orientation (signal K). In trial 35, the blue player, who is now the Communicator, draws a line on the game board to indicate the direction the orange player's triangle should be pointing (signal G). The orange player, however, interprets this behavior according to his own signal (signal K) and rotates his triangle three times, which is how many times the blue player visited the presumed target location. Also during his own turn as a Communicator in trial 36, the orange player uses the number of visits to indicate orientation (signal K). In a similar vein, the blue player interprets that behavior not according to his partner's but to his own signaling (signal G) and points his triangle in the direction the orange player's circle stepped out from the target location. In trial 37, both players reach a successful outcome. This surprising success is presumably due to the blue player incidentally visiting the orange player's presumed target location three times (in line with orange player's signal K) while executing his line drawing signal (G). Trial 38 reinforces the hypothetical incidental nature of the signal overlap in trial 37, with the orange player visiting the target location twice to indicate two rotations (signal K). Yet the blue player points his triangle in the direction the orange player's circle stepped out (equaling 3 rotations), consistent with his signal G. The same

misalignment continues to trial 44 when both players switch to different strategies each and the misalignment repeats (not shown in the video).