f

EMS: Biomedical Sciences

University of Edinburgh

Honours Project

Neuroscience

2023-24

Investigating differences in verbal engagement and rates of backchannel response between autistic and non-autistic adults

Table of Contents

Abstract

The Double Empathy Problem suggests that the interaction difficulties associated with autism are the result of differences in social behaviours between autistic and non-autistic individuals rather than the result of cognitive deficits. In my research project I investigated the differences of verbal engagement and rates of verbal backchannel response between autistic, non-autistic and mixed dyads and their relationship to rapport. I found strong evidence that both verbal engagement and the rate of verbal backchannel response were higher in non-autistic dyads compared to mixed dyads. Conversely, there was insufficient evidence of differences between autistic and mixed dyads on these variables. Nonetheless, there was moderate evidence indicating a relationship between the rate of backchannel response and rapport within autistic dyads. This finding suggests a nuanced dynamic, possibly reflecting the autistic individuals' prioritisation of task completion over social interaction.

Lay Abstract

Autistic and non-autistic people often experience breakdowns in communication when interacting with each other. It is important to understand what contributes to these breakdowns to ensure effective communication. In my study, I looked at how autistic and non-autistic individuals communicate in different pairings. I found that non-autistic individuals talked more and responded more when interacting with other non-autistic individuals compared to mixed pairs, where one person was autistic and the other was non-autistic. Interestingly, there weren't clear differences between pairs where both were autistic and mixed pairs. This might be because autistic individuals focus more on completing tasks than socialising, but more research is needed to understand this better.

Introduction

Double Empathy Problem of Autism

Currently, autism is defined as a group of highly heterogenous neurodevelopmental disorders defined by difficulties in two areas: (1) “deficits in reciprocal social interaction and communication”, and (2) “restricted, repetitive and inflexible patterns of behaviour” (WHO, 2018). In these areas, autism can present with a wide range of symptoms (see Figure 1).

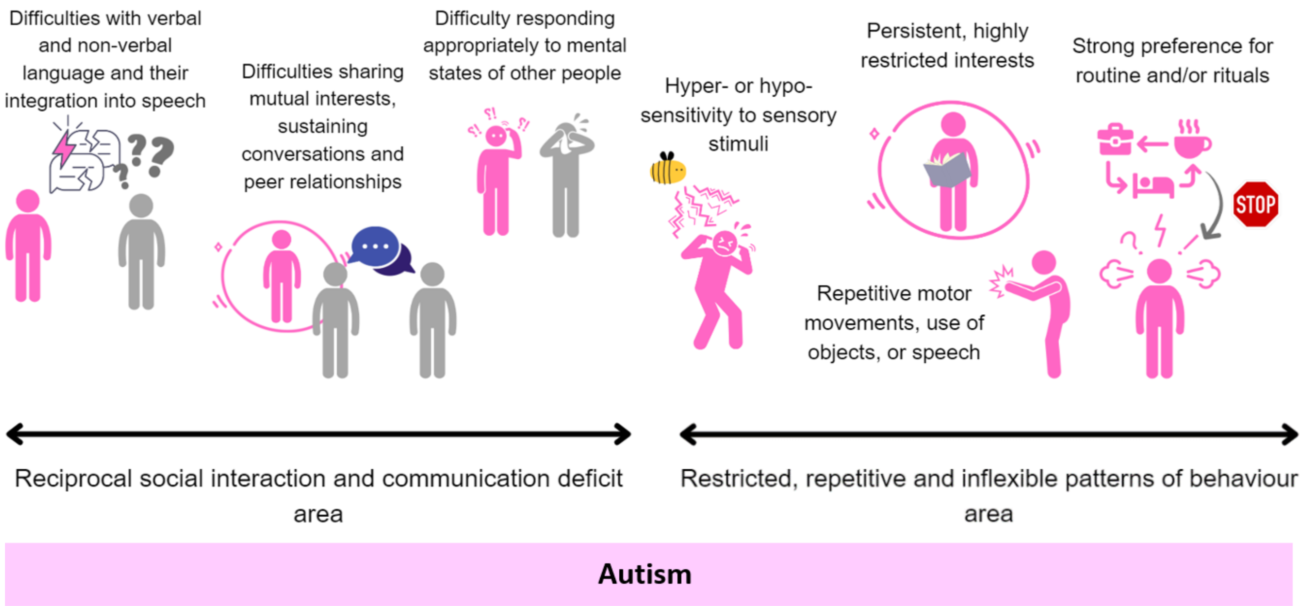
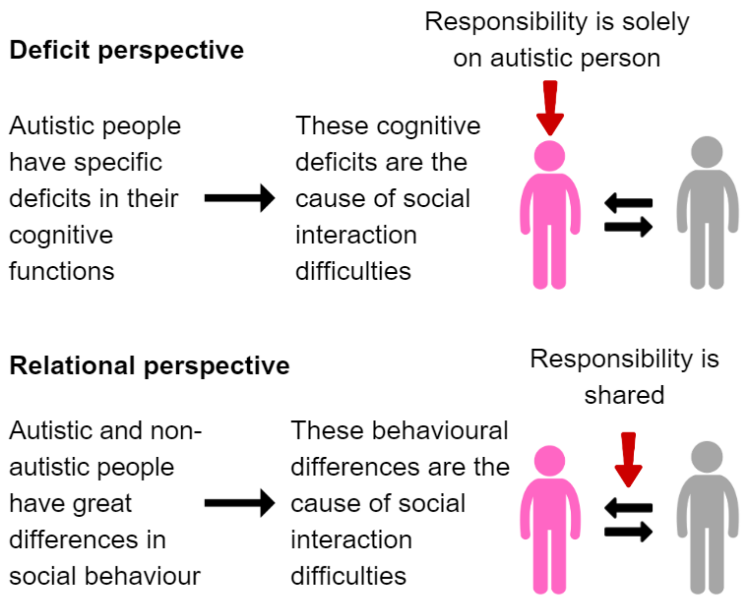


Figure 1 (original). The symptomology of autism. Autism is defined by a range of symptoms that can be grouped into two large areas: social communication and interaction deficit area, and restricted, repetitive and inflexible patterns of behaviour area. This figure shows a range of symptom presentations in each area with a brief description of each presentation (WHO, 2018). In the figure, the autistic person is represented in pink, and non-autistic people are represented in grey.

Historically, these symptoms were thought to be caused by deficits in certain cognitive functions. Several cognitive theories of autism have been developed but none have been fully supported by research to date (Fletcher-Watson & Happé, 2019). A new perspective on autistic social interaction and communication was proposed by Milton (2012) in his theory called the Double Empathy Problem. The Double Empathy Problem suggests that communication difficulties experienced by autistic people are the result of autistic and non-autistic people having vastly different dispositional outlooks on the lifeworld. According to this theory, when people with different dispositional outlooks interact (as is the case with autistic and non-autistic people), their social expectations and, as a result, social behaviours do not match, which leads to communication breakdowns and disruption of interaction. However, when individuals have similar dispositional outlooks (as in the case with two autistic or two non-autistic people), their social expectations and social behaviours are aligned which makes their interaction successful. Thus, the Double Empathy Problem views autism not from a deficit perspective, but from the perspective of relational differences and emphasises shared responsibility when, as in the case of autism, these differences are too big (see Figure 2A) (Lin *et al.* 2022).

A



B

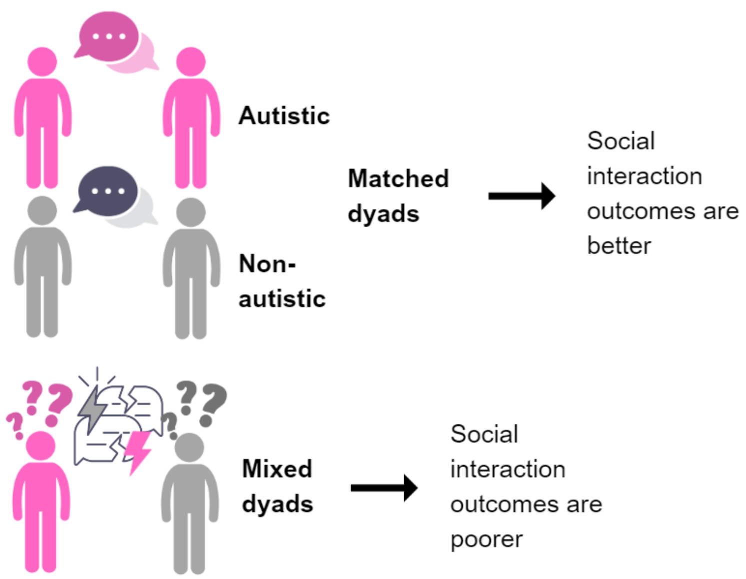


Figure 2 (original). Re-framing autism-associated social interaction and communication difficulties from a deficit perspective to a relational perspective (A and B). Panel A briefly summarises two perspectives highlighting the shift in responsibility from the individual level in deficit perspective to the dyad level in relational perspective. Panel B illustrates how different dyads (matched autistic, matched non-autistic and mixed) should affect outcomes of social interaction from a relational perspective, with mixed dyads theoretically exhibiting poorer outcomes (Lin *et al.* 2022). In the figure, autistic individuals are represented in pink, and non-autistic individuals are represented in grey.

One way to empirically test the Double Empathy Problem is to examine whether interactions differ when two interlocutors are either both autistic or non-autistic (in so-called “matched dyads”) or when one interlocutor is autistic and the other is non-autistic (in so-called “mixed dyads”). If the theory of the Double Empathy Problem is supported, the interactions between matched and mixed dyads should be different with mixed dyads exhibiting poorer interaction outcomes (see Figure 2B). Indeed, Crompton *et al* (2020a) found that information transfer across chains of matched autistic and matched non-autistic dyads (hereafter referred to simply as “autistic” and “non-autistic” or “matched” when referring to both) was more effective than information transfer across chains of mixed dyads, indicating that breakdowns in communication (and loss of information, as a result) occur more often when interlocutors are of different neurotypes. Furthermore, Crompton *et al* (2020b) found that rapport, a term which describes the subjective experience of mutual attentiveness, positivity and coordination during social interaction (Tickle-Degnen & Rosenthal, 1990), was also rated higher for matched rather than mixed dyads by both autistic and non-autistic adults. However, what differences between autistic and non-autistic social behaviours contribute to these communication breakdowns and decreased rapport remains largely unknown.

Aims and Hypothesis

The aim of my research project was to examine the potential role of verbal communicative behaviour in observed differences in interactions and rapport between matched and mixed dyads. For my analysis I chose two variables: verbal engagement and the rate of verbal backchannel response. I analysed interactions in autistic, non-autistic and mixed dyads to investigate (1) whether these variables differed depending on the type of dyad and (2) how these differences affected rapport.

Verbal Engagement

People of different neurotypes are less motivated to engage with each other. In their study, Sasson *et al* (2017) found that non-autistic adults and children preferred to engage with their non-autistic peers rather than their autistic peers. Similarly, Crompton *et al* (2020c) found that autistic adults felt more understood and less emotionally depleted when engaging with other autistic people. Furthermore, Chen *et al* (2021) found that both autistic and non-autistic adolescents preferred to engage more with peers of the same neurotype. It can be assumed that decreased preference for engagement within mixed dyads leads to decreased rapport. Thus, I hypothesised that in both autistic and non-autistic dyads participants would verbally engage with their conversational partners a higher proportion of time than participants in mixed dyads, which would lead to higher rapport.

Rate of Verbal Backchannel Response

Verbal backchannel response occurs when a listener produces short utterances, called backchannels, during a conversation while their conversational partner is speaking. The primary function of backchannels is to signal the listener's attention, interest and/or understanding (Wehrle *et al.* 2020). In English, typical expressions used as backchannels include ‘yeah’, ‘uh-huh’, ‘mm-hm’, ‘right’ and ‘okay’ (Ward & Tsukahara, 2000). It was found that differences in the realisation of backchannel response could lead to misunderstandings and communication breakdowns in cross-cultural interactions. In particular, the rate of backchannel response was found to play an important role in establishing rapport between two non-autistic interlocutors (Cutrone, 2005). It is possible that the rates of verbal backchannel response differ between autistic and non-autistic people, and this discrepancy leads to decreased rapport. Indeed, Wehrle *et al* (2024) found that in autistic dyads backchannels are less frequent and less diverse compared to non-autistic dyads. Similarly, Rifai *et al* (2022) found that the rate of backchannel response is much higher in non-autistic dyads compared to both autistic and mixed dyads. Thus, I hypothesised that in non-autistic dyads participants would backchannel to their conversational partners more frequently than in mixed or autistic dyads, and that higher rate of backchannel response would lead to higher rapport only in non-autistic dyads.

Methods

Ethics

This study was carried out in accordance with the British Psychological Society's Code on Human Research Ethics. Experimental procedures were reviewed and approved by the University of Edinburgh Psychology Research Ethics Committee, and all participants provided written informed consent prior to participating in the study.

Participants

A total of 72 participants were enrolled and randomly assigned to one of three groups: autistic (24 autistic individuals), non-autistic (24 non-autistic individuals), and mixed (12 autistic and 12 non-autistic individuals). Each group was further divided into 3 chains of 8 people (see Figure 3). In mixed chains, autistic and non-autistic people alternated, with all three chains starting with non-autistic people.

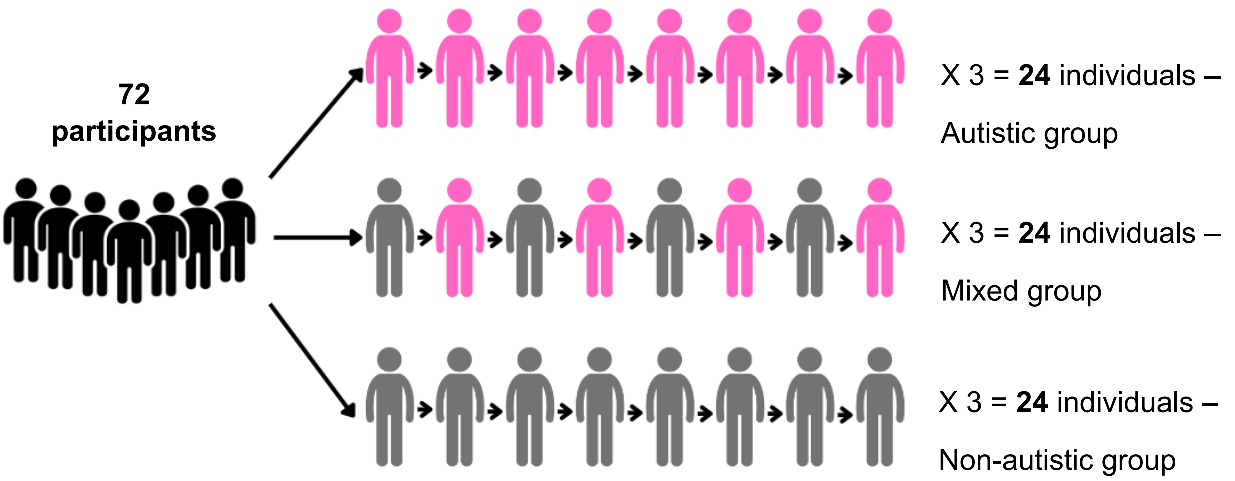


Figure 3 (original). Study design. In total 72 participants were recruited and assigned to one of the three types of groups: autistic, non-autistic or mixed. Each group consisted of 24 individuals divided into 3 chains of 8 people. In the figure, autistic individuals are represented in pink, and non-autistic individuals are represented in grey.

Due to the nature of chosen variables, the first participant in each chain was excluded from the analysis (more details on that below). Additionally, due to video loss, 4 participants from the autistic group were excluded, leaving the data unbalanced with a total of 59 participants: 17 participants in the autistic group, 21 participants in the non-autistic group and 21 participants in the mixed group (12 autistic and 9 non-autistic participants).

All groups were matched on age, gender, years of education and intelligence quotient (IQ) (see Table 1). All participants were native English speakers and did not have a diagnosis of social anxiety disorder. To assess IQ, all participants completed the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI—II) yielding Full-Scale IQ-2 (FSIQ-2) scores within the typical range (Wechsler, 2011). Non-autistic participants scored < 32 on the Autism-Spectrum Quotient

(Baron-Cohen *et al.* 2001) indicating low levels of autistic traits and autistic participants had a clinical diagnosis of autism or if self-diagnosed scored > 72 on the Ritvo Autism-Asperger Diagnostic Scale-Revised (RAADS-R) (n [self-diagnosed participants] = 4) indicating a high chance of being autistic (Ritvo *et al.* 2008).

Table 1. Descriptive statistics on demographics of participants and group comparisons using one-way analysis of variance (ANOVA) and Fisher's Exact Test. The three groups (autistic, non-autistic and mixed) were compared on age, gender, years of education, intelligence quotient (IQ), autism quotient and age of autism diagnosis. Categorical variables ("gender") were compared using Fisher's Exact Test. (All test assumptions were met.) Numerical variables were compared using one-way ANOVA. (In all cases, all test assumptions were met, with the exception of "years of education", where the normality assumption was slightly violated.) For each group, the mean is shown with standard deviation in parentheses. Each comparison shows the corresponding test statistic with degrees of freedom and p-value.

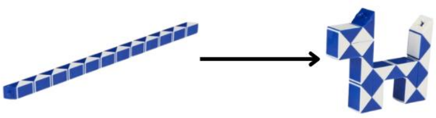
	Autistic (n = 17)	**Mixed** (n = 21)	**Non-autistic** (n = 21)	**Comparisons**
Age	38.41 (13.06)	36.80 (10.62)	39.71 (14.37)	One-way ANOVA, $F_{2,56} = 0.273$, $p = 0.762$
Gender	12 F, 2 M, 3 NB	16 F, 5 M	18 F, 3 M	Fisher's Exact Test, $p = 0.14$
Years of education	17.97 (2.71)	16.90 (1.99)	18.00 (1.55)	One-way ANOVA, $F_{2,56} = 1.814$, $p = 0.172$
IQ WASI—II, FSIQ-2	115.94 (16.78)	117.76 (14.29)	115.29 (12.28)	One-way ANOVA, $F_{2,56} = 0.165$, $p = 0.848$
Autism quotient	36.00 (6.15)	29.10 (13.82)	13.00 (5.81)	One-way ANOVA, $F_{1,23} = 0.0139$, $p = 0.907$
Age of autism diagnosis	31.44 (11.66)	30.89 (10.20)	NA	One-way ANOVA, $F_{1,23} = 0.0139$, $p = 0.907$

M = Males; F = Females; NB = Non-binary; IQ = Intelligence Quotient; WASI—II = Wechsler Abbreviated Scale of Intelligence, Second Edition; FSIQ-2 = Full-Scale IQ-2; ANOVA = analysis of variance; NA = Not Applicable; p = p-value.

Dyadic Interactions

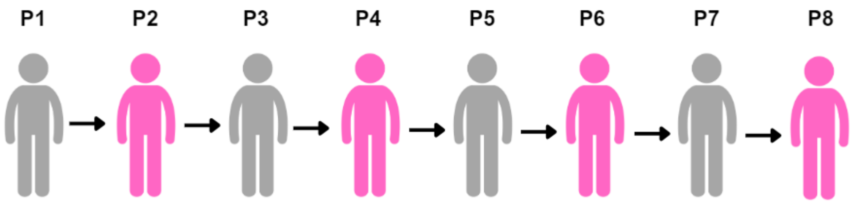
In each chain adjacent participants interacted with each other, thus forming a series of dyadic interactions. In each dyadic interaction, one participant played the role of demonstrator and the other participant played the role of observer. The demonstrator was asked to show the observer how to make a dog from a Rubik's cube (see Figure 4A). Once the demonstrator was finished with the task (regardless of whether they succeeded in the task or not), they were asked to leave the room and the next participant was invited. The next participant became the new observer, and the previous observer was asked to now be the demonstrator. Each chain contained 8 such dyadic interactions, with the first interaction occurring between the researcher and the first participant in the chain. The last participant in the chain did not interact with anyone as a demonstrator and was asked to simply make the dog without anyone present in the room. Thus, with the exception of the last participant, each participant in the chain participated in two dyadic interactions, playing first the role of observer and then of demonstrator (see Figure 4B). There was no strict time limits or specific guidelines as to how these interactions should be approached. Each dyadic interaction was recorded and later video coded for analysis.

A



B

Single Chain



Dyadic interactions within single chain

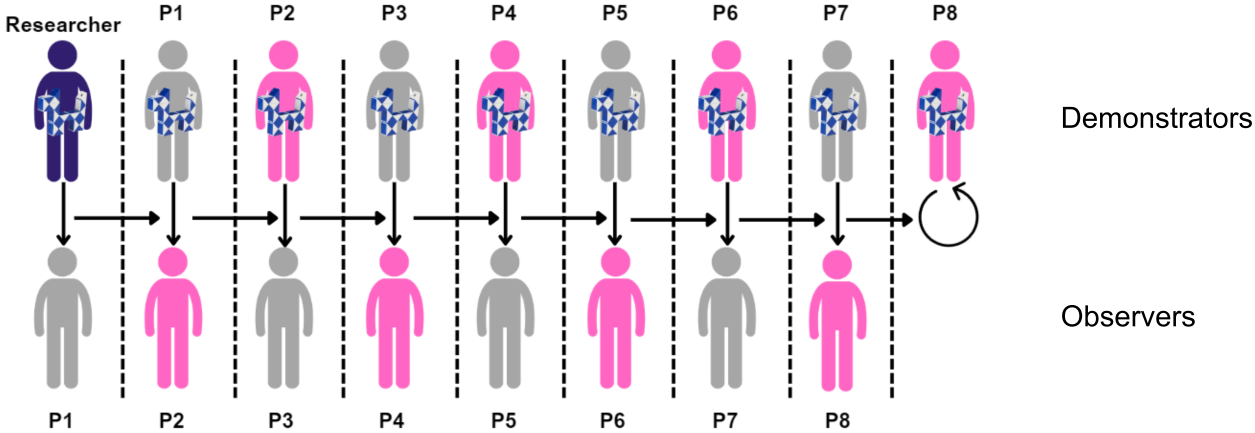


Figure 4 (original). Design of dyadic interactions (A and B). Panel A shows the task that participants as demonstrators were asked to complete in each dyadic interaction – show how to make a dog from a Rubik's cube. Panel B shows how a series of dyadic interactions within each chain were structured using a mixed chain as an example. In the figure, autistic individuals are represented in pink, and non-autistic individuals are represented in grey. All participants (P) are numbered from 1 to 8 (i.e., P1, P2, P3, …, P8).

Video Coding

Each dyadic interaction was coded separately using ELAN (The Language Archive, 2023). ELAN is a coding software which facilitates coding of variables in duration or frequency (Brugman *et al.* 2004). All variables (with the exception of task-related interaction) were coded and later analysed from the observers' perspective only. The first participant in each chain was excluded from the analysis because their demonstrator was the researcher (see Figure 4B). Before coding, I wrote a coding scheme that defined all variables. Below is a brief description of each variable (for a full coding scheme see *Appendix Methods*).

Task-oriented interaction was measured in seconds, starting from the moment when participants began the task and ending when the participant playing the role of demonstrator was finished with the task.

Verbal engagement was defined as the total length of task-oriented interaction in which the observer interacted verbally with their demonstrator and was measured in seconds. All parts of a task-oriented interaction in which the observer produced verbal utterances (i.e., questions, comments, or backchannels) less than 10 seconds apart were considered verbal engagement (including the time between those utterances). All parts in which the observer was silent or produced separate utterances, that were more than 10 seconds apart, were not considered verbal engagement. Since autistic people have a preference for restrictive behaviours (Kuenssberg *et al.* 2011), they are less likely to change roles assigned to them. For this reason, any parts of the interaction where the observer switched roles with the demonstrator (i.e., verbally explained to them how to make a dog from a Rubik's cube) were excluded to prevent result confounding. Prior to statistical analysis, the length of verbal engagement was divided by the length of task-oriented interaction (which did not include time when the observer switched roles with their demonstrator) to determine the proportion of time each observer verbally engaged with their demonstrator in each dyad.

Verbal backchannel response was measured as the count of the verbal backchannels produced by the observer during task-oriented interaction. Verbal backchannels were divided into four categories (see *Appendix Methods* for more details). Regardless of the category, any single verbal backchannel was equal to one (e.g., observer saying 'Yeah' was equivalent to them saying 'That makes sense' and both would count as equivalent individual backchannel). Prior to statistical analysis, the count of verbal backchannel response was divided by the length of task-oriented interaction to determine the rate with which each observer backchanneled to their demonstrator in each dyad.

Rapport

After the interaction, all demonstrators were asked to rate their experience of rapport with their observers on a five-dimensional rapport scale (Crompton *et al.* 2020b), where overall rapport ranged from 0 to 500 and was calculated from five equally weighted components: ease, enjoyment, success, friendliness, and awkwardness.

Statistical Analysis

Prior to analysis, the assumption was made that even though adjacent dyads shared participants (see Figure 4B), this would not affect the independence of the collected data. This assumption was made based on the fact that in all cases (with the exception of task-oriented interaction), data was collected only from the observer's perspective, not from the perspective of the entire dyad, and within each chain, observers and demonstrators were never repeated. In the case of task-oriented interaction, data was collected from the prospective of the entire dyad. However, since the participants played different roles in adjacent dyads and the length of the interaction was rather depend on the demonstrator's ability to make a dog, I made the assumption that the data was independent in this case too.

Statistical tests that were used to analyse data were one-way analysis of variance (ANOVA), bootstrap method, simple linear regression and Fisher's Exact Test. For one-way ANOVA, in each case the assumptions of normal distribution and equal variance were checked by visually examining the quantile-quantile (QQ) plot and the residual plot, respectively. Tukey's test was used in post-hoc analyses and for adjustment for multiple comparisons. The code for the bootstrap method was used from Johnston and Faulkner paper (2021) with minor modifications to compare means instead of medians. For simple linear regression, in each case the assumptions of linearity and equal variance were checked by visually examining the scatterplot of the two variables and the residual plot, respectively. Independence and normality of errors were checked by visually examining the residual plot and normal probability plot, respectively. For Fisher's Exact Test, I used Freeman-Halton extension (Freeman & Halton, 1951) to apply the test to a 2x3 contingency table. I also used Benjamini–Hochberg method to adjust for multiple comparisons (Benjamini & Hochberg, 1995). I also differentiated between mild and extreme outliers with the former defined as outliers positioned within the outer fences of the boxplot, and the latter defined as outliers positioned outside the outer fences of the boxplot. The outer fences for each boxplot were calculated as the difference between upper (or lower) quartile and the interquartile range of the boxplot multiplied by three (Halldestam, 2016).

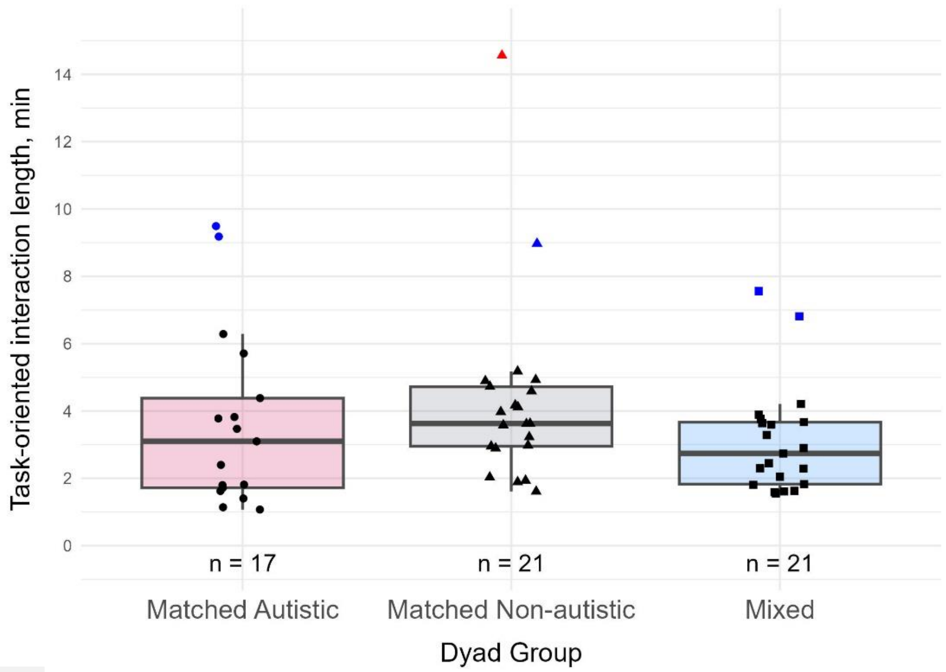
When reporting my results, I used the guidelines from paper by Ganesh and Cave (2018) which recommends interpreting p-values in terms of strength of evidence rather than statistical significance, and paying more attention to effect sizes and 95% confidence intervals (CIs). However, I want to acknowledge that there is an ongoing debate about what is the most meaningful and honest way to statistically analyse and present data (Wasserstein *et al.* 2019).

Results

Task-oriented Interaction

I tested whether there was a difference in the length of task-oriented interaction between different groups of dyads using one-way ANOVA (see Figure 5A). There was one extreme and several mild outliers in the untransformed data. After logarithmic transformation, only one mild outlier remained (see Figure 5B). Since one-way ANOVA is sensitive to outliers (Halldestam, 2016), I analysed first untransformed and then log-transformed data to confirm my results. In the first case, the assumptions of equal variance and normal distribution were not met. In the second case, both assumptions were fully met.

A



B

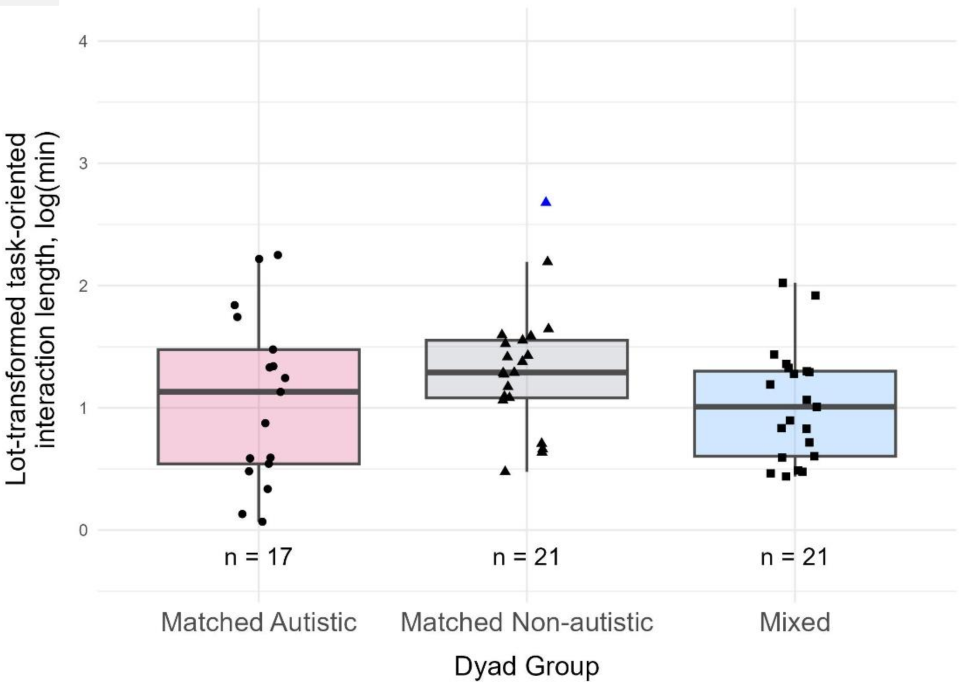


Figure 5. Differences in the length of task-oriented interaction between different groups of dyads (A and B). Panel A shows the length of task-oriented interaction in each dyad group (autistic – n = 17, non-autistic – n = 21, and mixed – n = 21), expressed in minutes (min). Similarly, panel B also shows the length of task-oriented interaction expressed in minutes but with logarithmic transformation (log(min)) to meet all necessary assumptions for comparison. The bold line on the boxplots indicates the median for each dyad group, and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. Points represent the length of task-oriented interaction in each individual dyad. Autistic dyads are shown as circles, non-autistic dyads are shown as triangles, and mixed dyads are shown as squares. Mild outliers are represented as blue points, and extreme outliers are represented as red points. Points that are not outliers are shown in black.

One-way ANOVA without logarithmic transformation showed insufficient evidence that task-oriented interaction length varied between different groups of dyads (Figure 5A; $F_{2,56} = 1.2087$, p-value = 0.306). On average, task-oriented interaction length in autistic dyads was 3.66 min (± SE: 0.591; 95% CIs 2.47 to 4.84), in non-autistic dyads – 4.39 min (± 0.532; 3.32 to 5.45), and in mixed dyads – 3.23 min (± 0.532; 2.16 to 4.29). Similarly, one-way ANOVA with logarithmic transformation provided also insufficient evidence of a difference between groups (Figure 5B; $F_{2,56} = 1.7569$, p-value = 0.182).

Rapport

I analysed whether there was a difference in demonstrators' rapport between different groups of dyads following their interactions with the observers using one-way ANOVA (see Figure 6). There was one extreme outlier and several mild outliers in the data and no data transformation methods helped to get rid of them. However, the data met both assumptions of normal distribution and equal variance. Therefore, I continued to analyse the untransformed data. Afterwards, I also analysed the data from which the extreme outlier was removed to confirm my results. In that case, all test assumptions were also met.

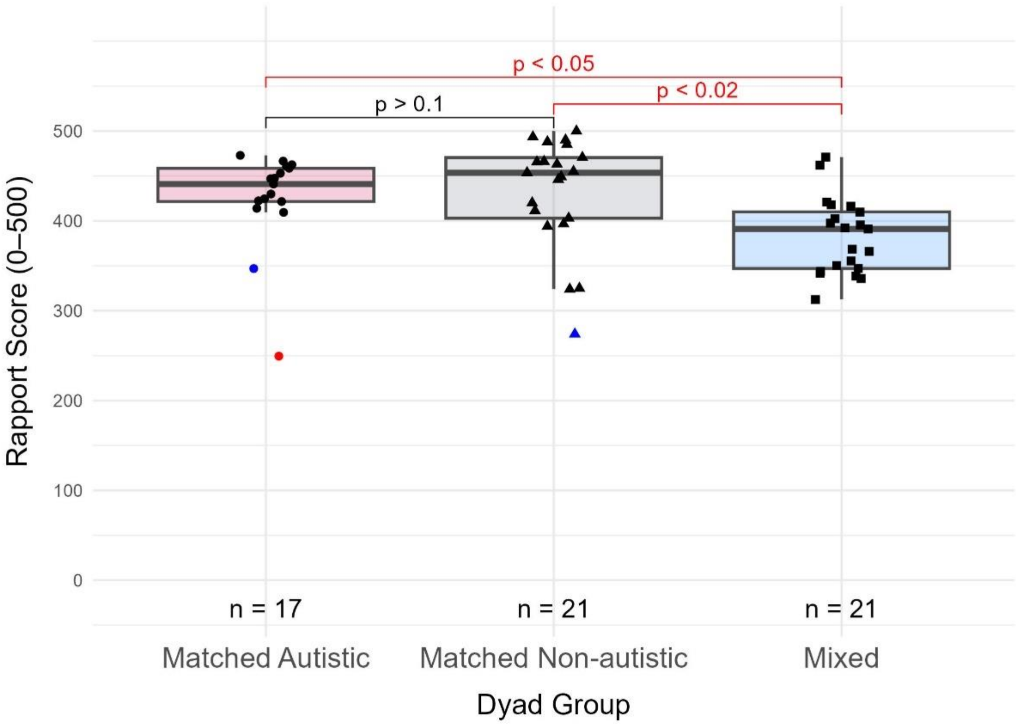


Figure 6. Differences in demonstrators' rapport in different groups of dyads. After each interaction, each demonstrator was asked to rate their rapport with the their observer and final score for each demonstrator's rapport was calculated on a scale ranging from 0 to 500. The bold line on the boxplots indicates the median for each group of dyads (autistic – n = 17, mixed – n = 21 and non-autistic – n = 21), and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. The brackets above the boxplots indicate the p-values from pairwise Tukey post-hoc tests. Points represent demonstrator's rapport in each individual dyad. Autistic dyads are shown as circles, non-autistic dyads are shown as triangles, and mixed dyads are shown as squares. Mild outliers are represented as blue points, and extreme outliers are represented as red points. Points that are not outliers are shown in black.

One-way ANOVA revealed strong evidence that demonstrators' rapport varied between different groups of dyads (Figure 6; $F_{2,56} = 5.1813$, p-value = 0.00862). On average, demonstrator's rapport in autistic dyads was 425 (± SE: 12.9; 95% CIs 399 to 451), in mixed dyads it was 383 (± 11.6; 359 to 406) and in non-autistic dyads it was 432 (± 11.6; 409 to 455). Post-hoc analysis provided strong evidence that the difference in rapport between non-autistic and mixed dyads was 49.31 (± SE: 16.4; t ratio = 3.002, p-value = 0.0110; 95% CIs 9.76 to 88.86), and insufficient evidence for a difference between non-autistic and autistic dyads (estimated difference of 6.87 ± 17.4; t ratio = 0.396, p-value = 0.918; - 34.94 to 48.68). There was moderate evidence that the difference between autistic and mixed dyads was 42.44 (± 17.4; t ratio = 2.444, p-value 0.0459; 0.63 to 84.25). However, after removing the extreme outlier from the data, post-hoc analysis provided strong evidence that the difference between autistic and mixed dyads was 53.42 (± 15.9; t ratio =

3.365, p-value < 0.01; 15.2 to 91.7), while the results comparing non-autistic and autistic dyads, and non-autistic and mixed dyads remained approximately the same.

Verbal Engagement

I analysed whether there was a difference in the proportion of observers' verbal engagement between different groups of dyads using one-way ANOVA (see Figure 7). The data fully met the assumptions of equal variance and normal distribution.

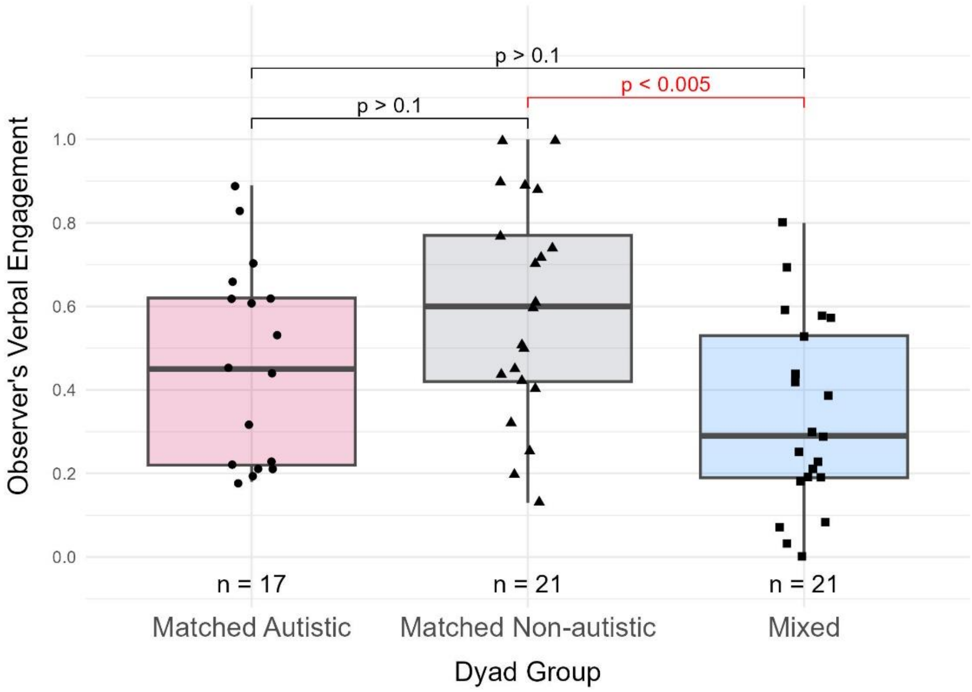


Figure 7. Differences in the proportion of observers' verbal engagement in different groups of dyads. The amount of time each observer was verbally engaged with their demonstator in each dyad was measured and then divided by the total length of task-oriented interaction to determine the proportion of each obsever's verbal engagement. The bold line on the boxplots indicates the median for each dyad group (autistic – n = 17, non-autistic – n = 21 and mixed – n = 21), and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. Points represent proportion of observer's verbal interaction in each individual dyad. Autistic dyads are shown as circles, non-autistic dyads are shown as triangles, and mixed dyads are shown as squares. The brackets above the boxplots indicate the p-values from pairwise Tukey post-hoc tests.

One-way ANOVA revealed strong evidence that the proportion of observers' verbal engagement varied between different groups of dyads (Figure 7; $F_{2,56}$ = 5.9268, p-value = 0.00463). Post-hoc analysis showed strong evidence that observers verbally engaged with their demonstrators a higher proportion of time in non-autistic dyads (mean ± SE: 0.593 ± 0.053; 95% CIs 0.487 to 0.699) compared to mixed dyads (0.335 ± 0.053; 0.229 to 0.441), with a difference of 0.258 (± SE: 0.075; t ratio = 3.443, p-value = 0.0031; 95% CIs 0.078 to 0.438). There was insufficient evidence for a difference in the proportion of observers' verbal engagement between autistic (0.467 ± 0.059; 0.349 to 0.585) and non-autistic dyads, or autistic and mixed dyads with a difference of -0.126 (± 0.079; t ratio = -1.594, p-value = 0.257; -0.317 to 0.064) and 0.132 (± 0.079; t ratio = 1.662, p-value = 0.229; -0.059 to 0.322), respectively.

Additionally, I decided to also test whether there was a difference in the engagement between observers in mixed dyads depending on whether they were autistic or non-autistic. Since data did not meet the normal distribution assumption and groups had fairly different variances (see Figure 8), I used bootstrap method which is robust to both conditions (Johnston & Faulkner, 2021).

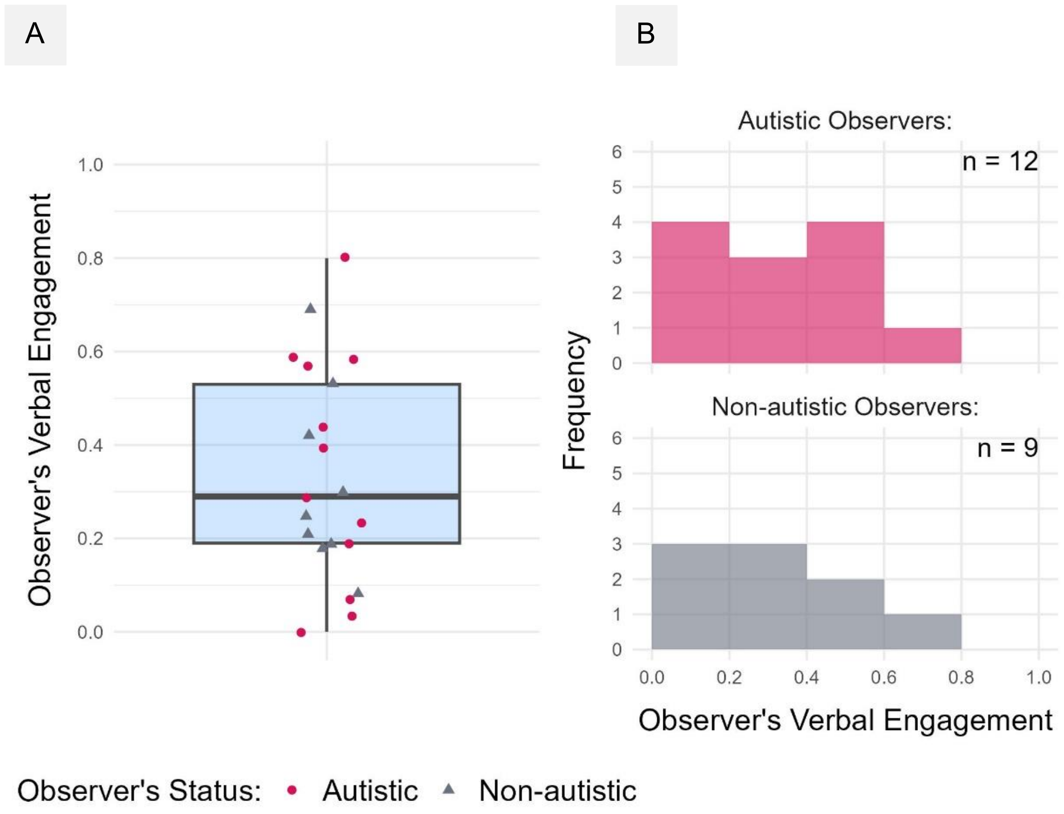
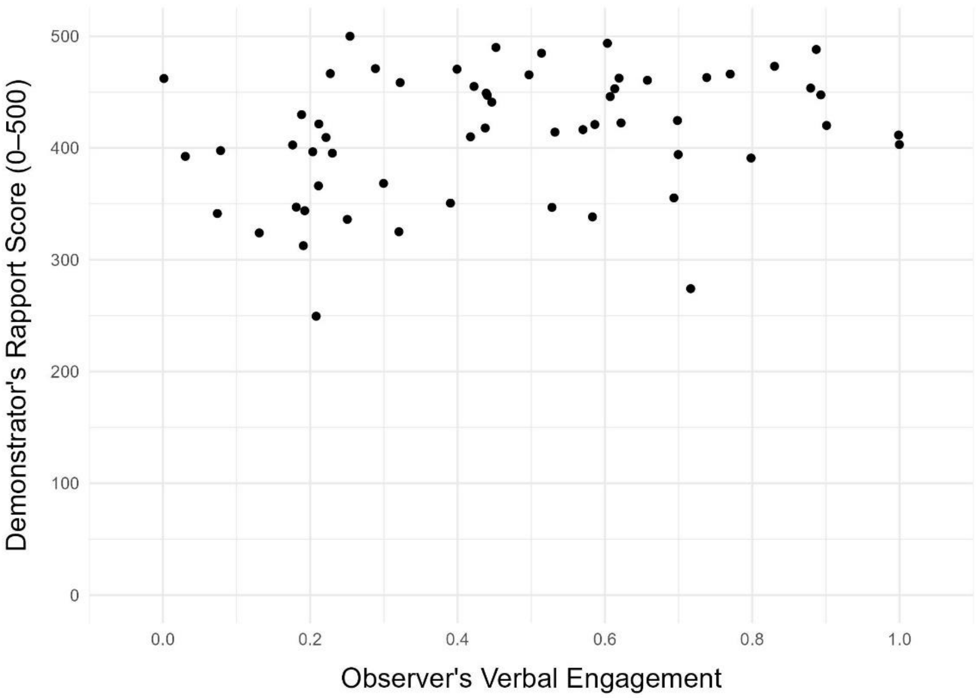


Figure 8. Differences in the observer's engagement in mixed dyads depending on the observer's autism status (A and B). Panel A compares observer's engagement when the observer is autistic (n = 12) and when the observer is non-autistic (n = 9). The bold line on the boxplot indicates the median for each group (autistic and non-autistic), and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. Points represent each observer's engagement. Autistic observers are shown as circles, and non-autistic observers are shown as squares. Panel B shows the distribution of observers' engagement when the observer is autistic and non-autistic.

Bootstrap method did not provide sufficient evidence that the engagement in mixed dyads was associated with the observer being autistic or non-autistic (see Figure 8; $N_{non\text{-}autistic} = 9$, $N_{autistic} = 12$, p-value = 0.717).

Next, I tested whether observer's verbal engagement affected their demonstrator's rapport using simple linear regression (see Figure 9A). The data met all test assumptions. There were several outliers in the data which could affect the slope and intercept estimates (Tedeschi & Galyean, 2023). Therefore, I also analysed data from which the outliers were removed to confirm my results. In that case, all test assumptions were also met.

A



B

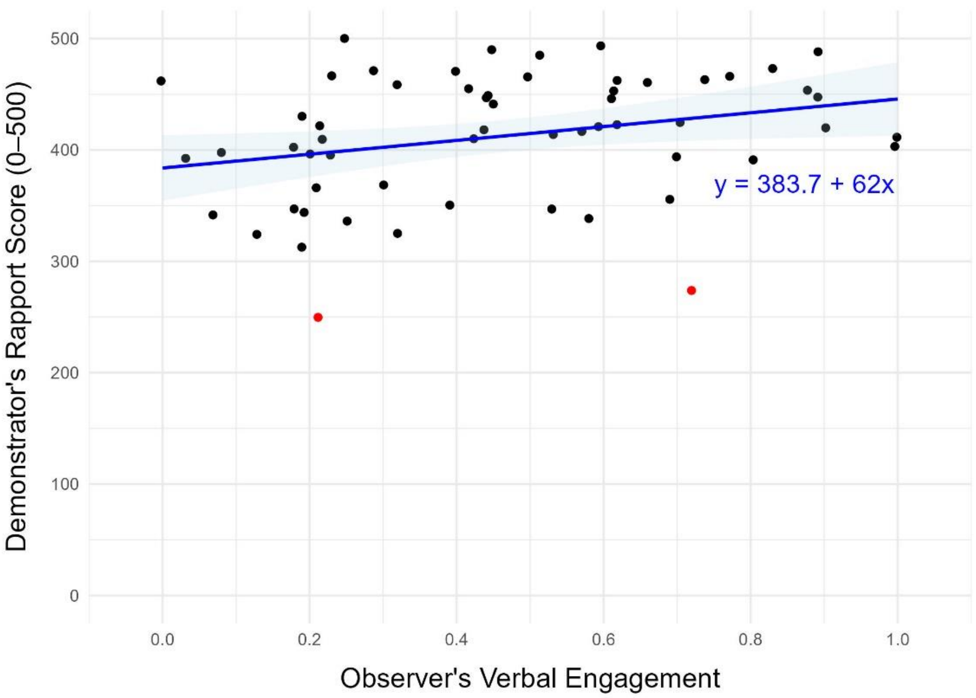
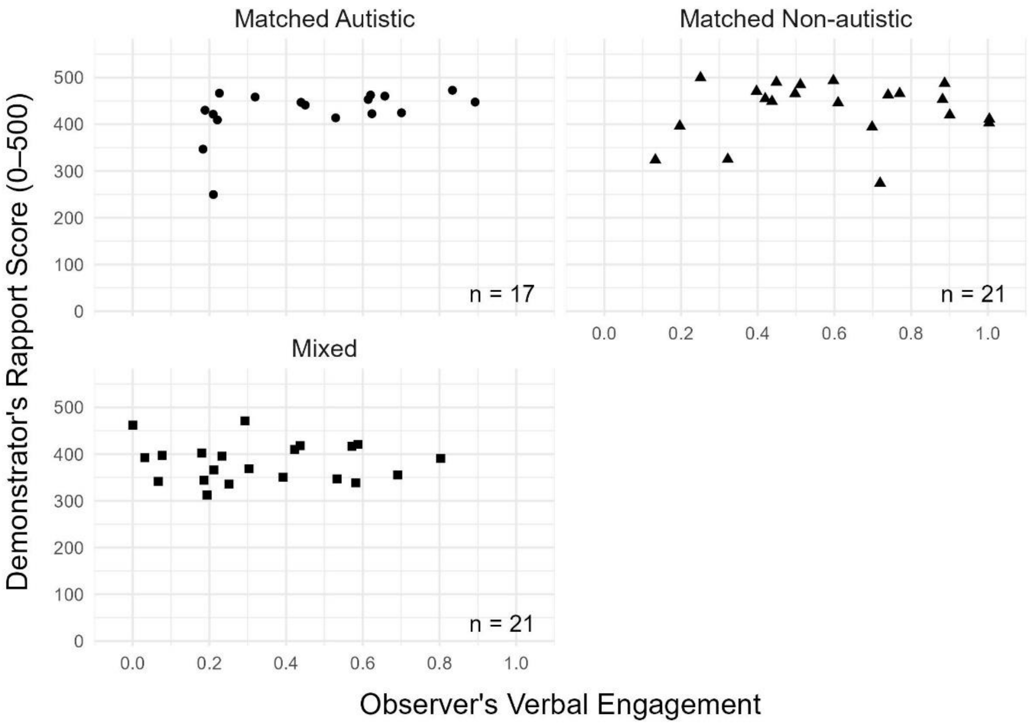


Figure 9. The relationship between observers' verbal engagement and demostrators' rapport (A and B). Panel A shows the association between the observer's verbal engagement and their demonstrator's rapport. Panel B shows demonstrators' rapport as a function of observers' engagement from a sample of 59 dyads: slope (mean ± SE) – 62.00 ± 27.54, and intercept – 383.70 ± 14.67. Points represent individual dyads as crossovers between the observer's verbal interaction and the demonstrator's rapport. Outliers are represented as red points. Points that are not outliers are shown in black.

Simple linear regression provided moderate evidence that demonstrator's rapport increased with the observer's verbal engagement (Figure 9B; $F_{1,57} = 5.066$, p-value = 0.0283, slope ± SE = 62.00 ± 27.54, intercept ± SE = 383.70 ± 14.67). However, when the outliers were removed, it provided strong evidence of a relationship with similar slope and intercept estimates ($F_{1,55} = 5.066$, p-value = 0.0125, slope ± SE = 62.77 ± 24.30, intercept ± SE = 388.64 ± 12.94).

I also examined the relationship between observers' verbal engagement and their demonstrators' rapport within each dyad group separately (see Figure 10A). The non-autistic group failed to meet the assumption of linearity, so it was excluded from the analysis. The autistic group met all assumptions but contained a significant outlier, so it was analysed both with and without that outlier. The mixed group met all assumptions except homoscedasticity when the data were square-root-transformed. I therefore proceeded with the autistic and mixed groups only. However, because of the small sample sizes, the following results should be interpreted with caution.

A



B

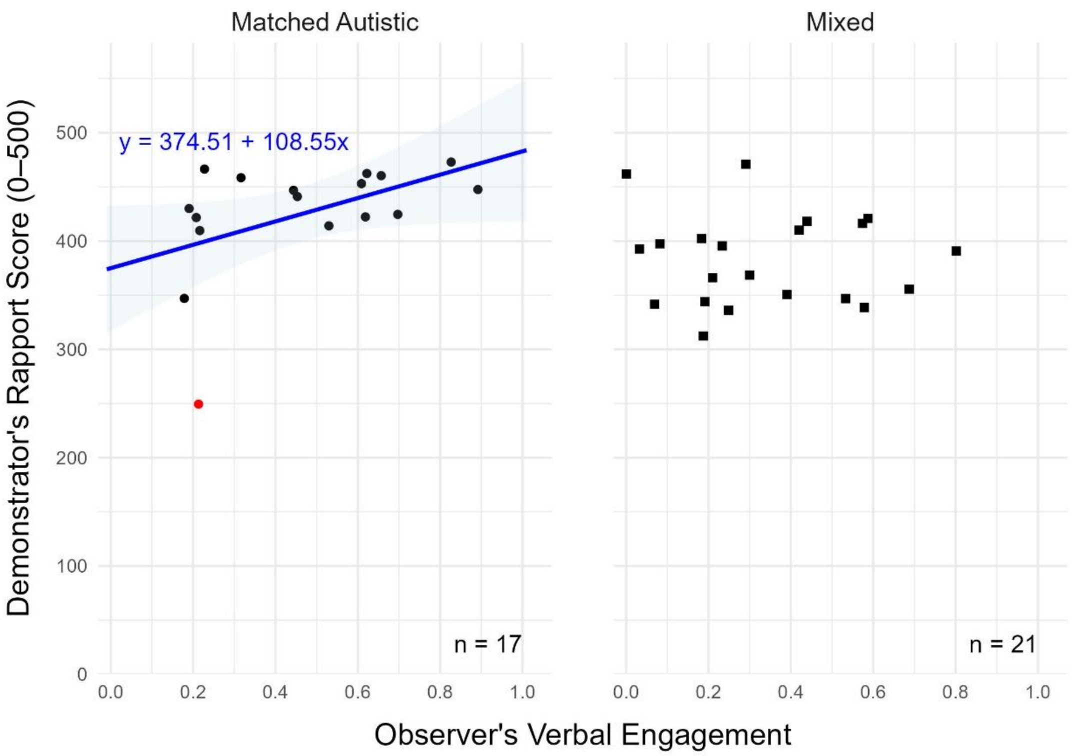


Figure 10. Relationship between observer's engagement and their demostrator's rapport in each group of dyads (A and B). Panel A shows the association between the proportion of each observer's verbal engagement and their demonstrator's rapport in each group: autistic (n = 17), non-autistic (n = 21) and mixed (n = 21). Panel B shows that demonstrators' rapport as a function of observers' proportion of verbal interaction was found only in autistic group (n = 17): slope (mean ± SE) – 108.55 ± 52.07, and intercept – 374.51 ± 27.09. Points represent individual dyads as crossovers between the observer's verbal interaction and the demonstrator's rapport. Autistic dyads are shown as circles, non-autistic dyads are shown as triangles, and mixed dyads are shown as squares. Outliers are shown as red points. Points that are not outliers are shown in black.

Simple linear regression provided weak/moderate evidence of a relationship in the autistic dyad group (see Figure 10B; $F_{1,15} = 4.345$, p-value = 0.0546, slope ± SE = 108.55 ± 52.07, intercept ± SE = 374.51 ± 27.09) and insufficient evidence of a relationship in the non-autistic ($F_{1,19} = 0.1273$, p-value = 0.725) and mixed dyad group ($F_{1,19} = 0.05457$, p-value = 0.819). However, after removing outliers from the data in the autistic dyad group, there was only weak evidence of a relationship ($F_{1,14} = 3.759$, p-value = 0.0729).

Finally, I decided to analyse whether there was a difference in the frequency of role switch between different groups of dyads using Fisher's Exact Test (see Figure 11). To analyse the data, I counted the number of dyads in which a role switch did and did not occur for each group of dyads. Each group in this 2x3 contingency table was mutually exclusive. Thus, all test assumptions were met.

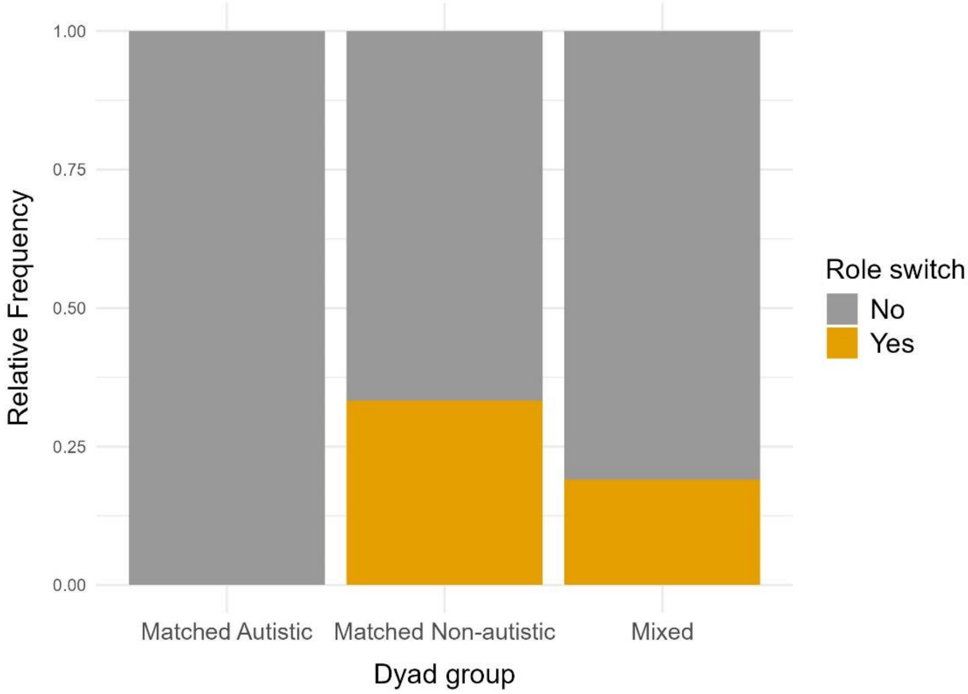


Figure 11. Differences in the relative frequency of role switch across different groups of dyads. The number of dyads in which a role switch occurred was counted within each dyad group and divided by the total number of dyads within that dyad group (autistic – n = 17, mixed – n = 21, and non-autistic – n = 21). This mosaic plot shows the relative frequency of the role switch within each group of dyads: the proportion of dyads in which a role switch occurred is shown in orange, and the proportion of dyads in which a role switch did not occur is shown in grey.

Fisher's Exact Test showed moderate evidence for a difference between groups (Figure 11; p-value = 0.02559). Post-hoc analysis showed moderate evidence that observers switched roles with demonstrators more frequently in non-autistic dyads than in autistic dyads (n = 38; p-value = 0.0324). It showed insufficient evidence for a difference between non-autistic and mixed dyads (n = 42; p-value = 0.484), and autistic and mixed dyads (n = 38; p-value = 0.17).

Rate of Verbal Backchannel Response

I analysed whether there was a difference in the observers' rate of verbal backchannel response between different groups of dyads using one-way ANOVA (see Figure 12). The data met both assumptions of normal distribution and equal variance.

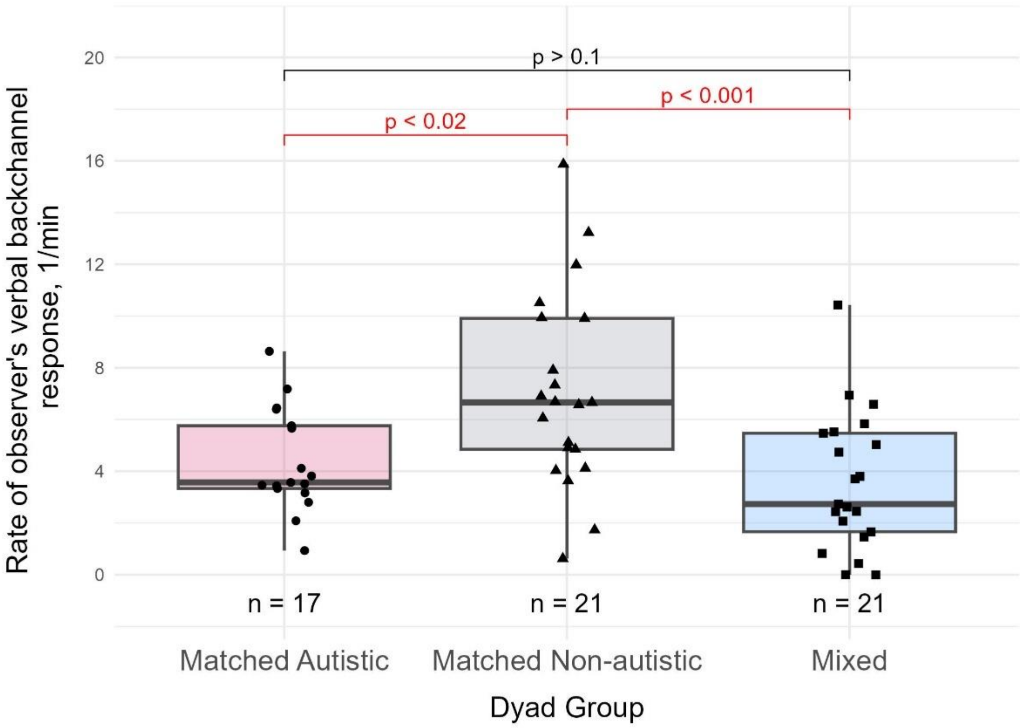


Figure 12. Differences in the rate of observers' backchannel response in different groups of dyads. The number of backchannels produced by each observer in the dyad during task-oriented interaction with their demonstrator was counted and then divided by the length of task-oriented interaction to determine the rate of backchannel response for each observer. The bold line on the boxplots indicates the median for each dyad group (autistic – n = 17, mixed – n = 21 and non-autistic – n = 21), and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. The brackets above the boxplots indicate the p-values from pairwise Tukey post-hoc tests. Points represent the observer's rate of verbal backchannel response in each individual dyad. Autistic observers are shown as circles, and non-autistic observers are shown as squares.

One-way ANOVA revealed very strong evidence that observers' rates of verbal backchannel response differed between different groups of dyads (Figure 12; $F_{2,56}$= 7.7364, p-value = 0.00108). Post-hoc analysis showed very strong evidence that observers backchanneled more frequently in non-autistic dyads (mean ± SE: 6.93 ± 0.642; 95% CIs 5.64 to 8.21) compared to mixed dyads (3.46 ± 0.642; 2.18 to 4.75) with a difference in frequency of 3.46 (± SE: 0.908; t ratio = 3.812, p-value = 0.0010; 95% CIs 1.276 to 5.65). There was also moderate evidence that this was also the case when comparing non-autistic dyads with autistic dyads (4.37 ± 0.714; 2.94 to 5.80) with a difference of 2.55 (± 0.960; t ratio = 2.658, p-value = 0.0272; 0.241 to 4.87). There was insufficient evidence for the difference between autistic and mixed dyads (estimated difference of 0.91 ± 0.960; t ratio = 0.947, p-value = 0.613; -1.40 to 3.222).

I also analysed whether the rate of observer's backchannel response varied in mixed dyads depending on whether the observer was autistic or non-autistic. As before, due to violation of normality and unequal variances between groups (see Figure 13), I used bootstrap method to analyse the data.

Bootstrap method provided insufficient evidence that observer's rate of backchannel response in mixed dyads depended on the observer's autism status (Figure 13; $N_{non\text{-}autistic} = 9$, $N_{autistic} = 12$, p-value = 0.536).

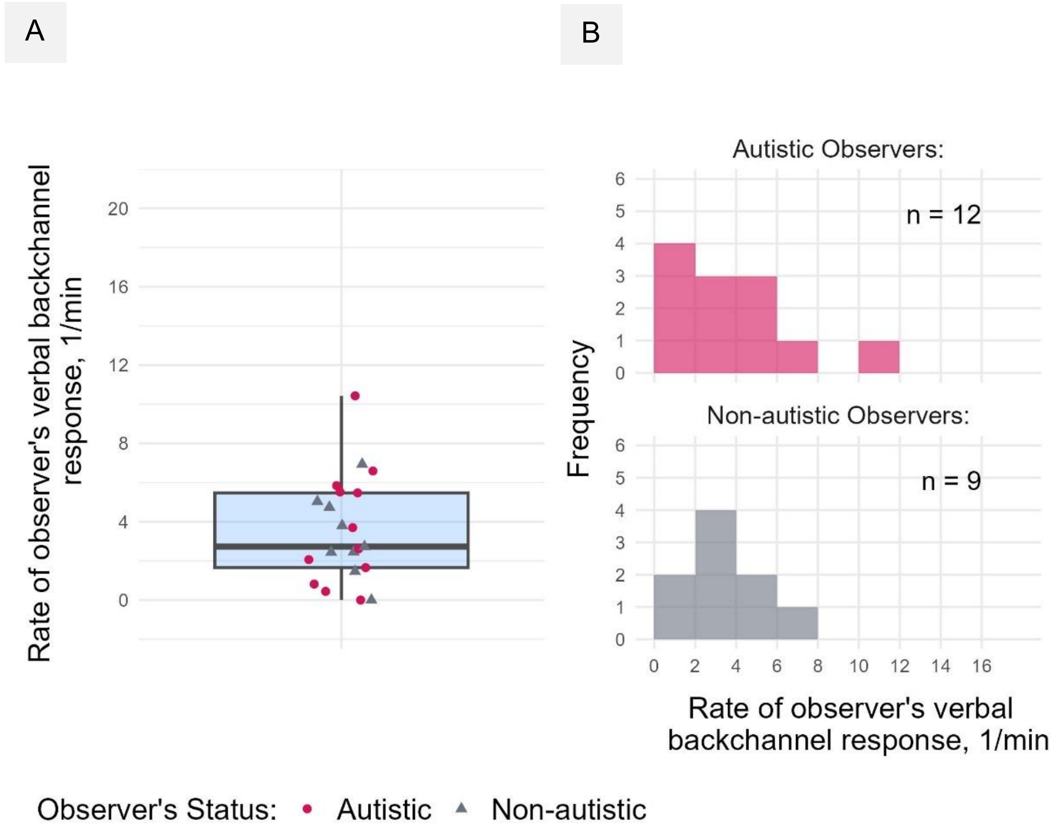
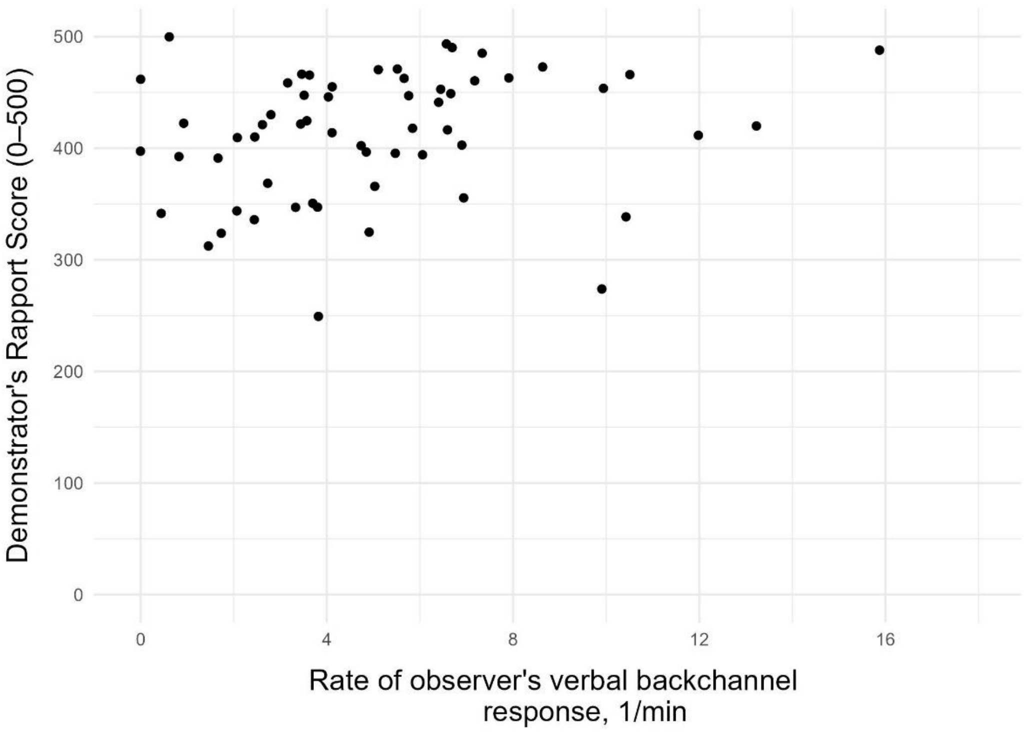


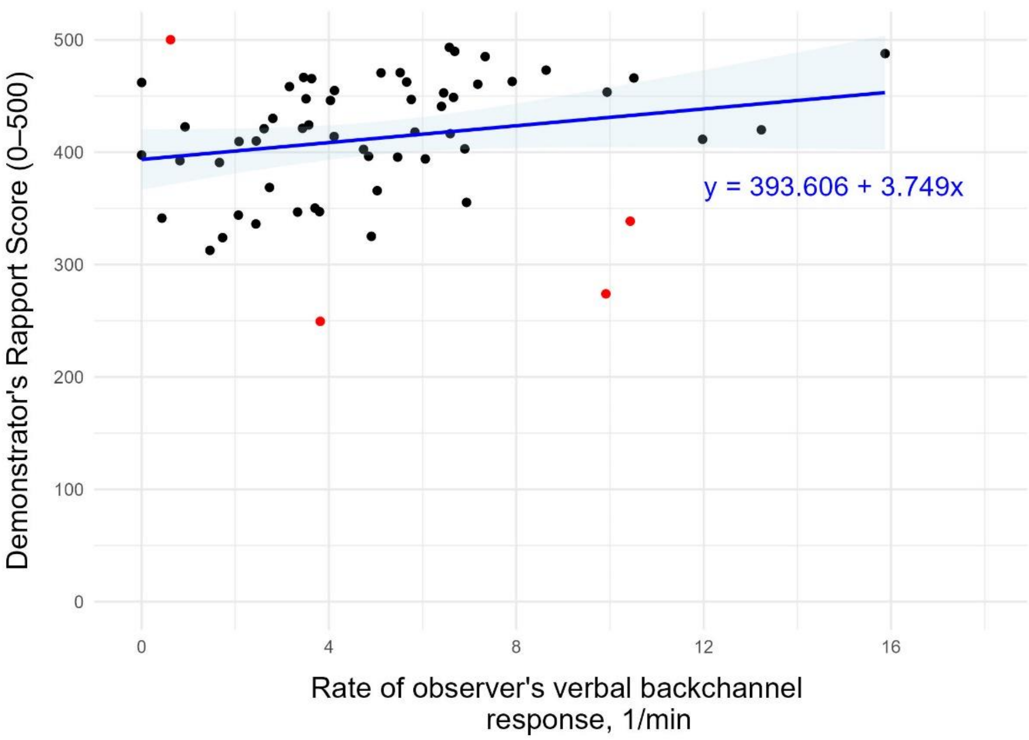
Figure 13. Differences in the rate of verbal backchannel response in mixed dyads depending on the observer's autism status (A and B). Panel B compares the rate of observer's verbal backchannel response when the observer is autistic (n = 12) and when the observer is non-autistic (n = 9). The bold line on the boxplots indicates the median for each group (autistic and non-autistic), and the lower and upper borders of the box indicate the lower and upper quartiles, respectively. Points represent each observer's rate of backchannel response. Panel A shows the distribution of the rate of observers' verbal backchannel response in mixed dyads when the observer is autistic (n = 12) and non-autistic (n = 9).

Finally, I tested whether observer's rate of backchannel rate affected their demonstrator's rapport using simple linear regression (see Figure 14A). The data met all test assumptions. However, there were several outliers in the data, so the following results should be taken with caution.

A



B



C

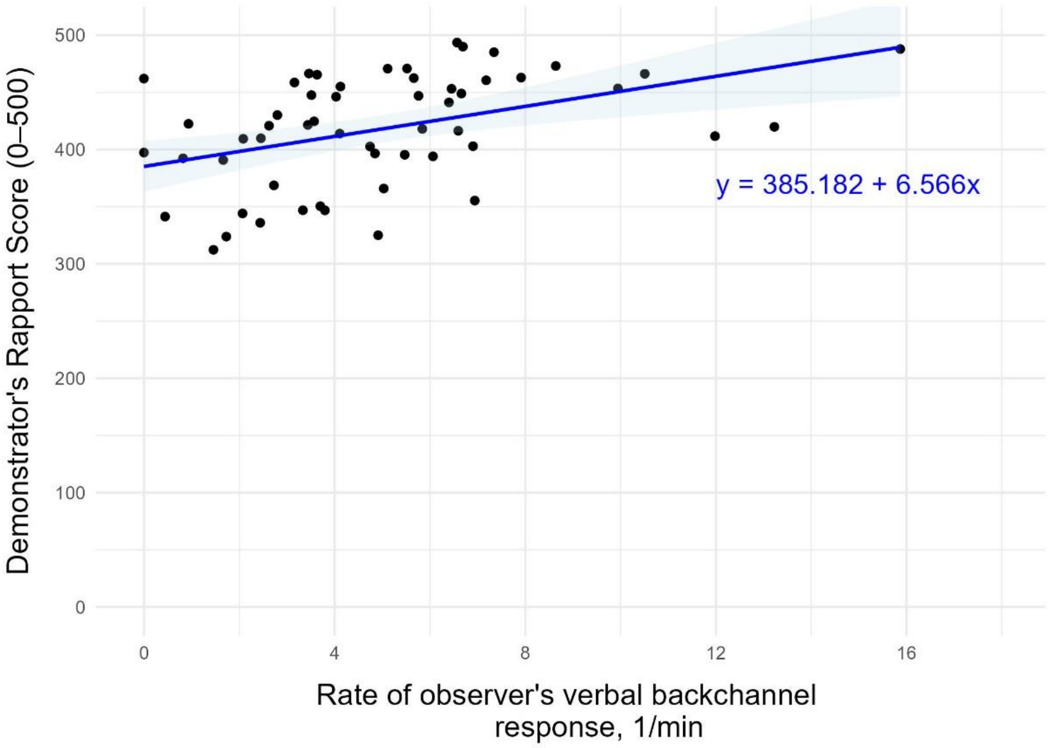
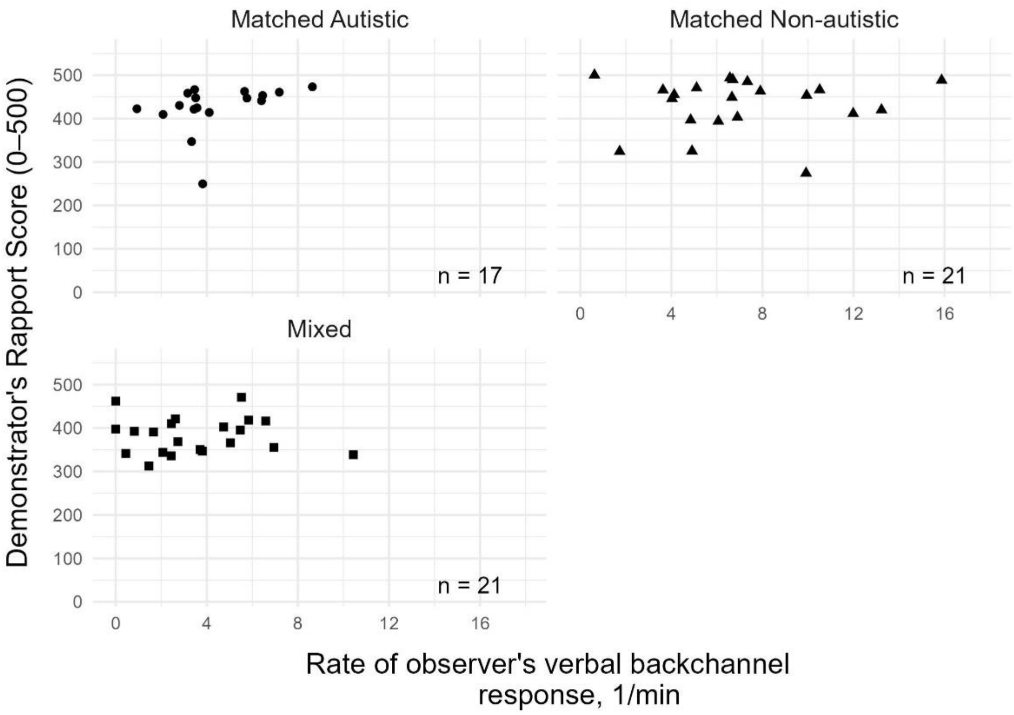


Figure 14. The relationship between observers’ rate of backchannel response and demostrators’ rapport (A, B and C). Panel A shows the association between the rate of each observer’s backchannel response and their demonstrator's rapport. Panel B shows demonstrators’ rapport as a function of observers’ rate of verbal backchannel response from a sample of 59 dyads: slope (mean ± SE) – 6.739 ± 1.888, and intercept – 393.607 ± 13.338. Panel C shows demonstrators’ rapport as a function of observers’ rate of verbal backchannel response without distinct outliers: slope – 3.814 ± 2.252, and intercept – 384.940 ± 10.946. Points represent individual dyads as crossovers between the observer’s rate of backchannel response and the demonstrator’s rapport. Autistic dyads are shown as square points, mixed dyads are shown as triangle points, and non-autistic dyads are shown as circles. Outliers are represented as red points. Points that are not outliers are shown in black.

Simple linear regression provided weak evidence that demonstrator’s rapport increased with the observer’s rate of backchannel response (Figure 14B; $F_{1,57} = 2.868$, p-value = 0.09582, slope ± SE = 3.814 ± 2.252, intercept ± SE = 393.607 ± 13.338). However, when the outliers were removed, it provided very strong evidence of a relationship with a twofold higher slope estimate (Figure 14C; $F_{1,53} = 12.74$, p-value = 0.00077, slope ± SE = 6.739 ± 1.888, intercept ± SE = 384.940 ± 10.946).

I also analysed relationship between observer’s rate of backchannel response and their demonstrator’s rapport within each group of dyads separately (see Figure 15A). The mixed group was excluded from the analysis because the data did not meet the assumption of linearity. In the autistic group there were two distinct outliers that slightly skewed the normality and equal variance assumptions. In the non-autistic group, all test assumptions were met although the assumption of linearity was slightly violated due to multiple outliers.

A



B

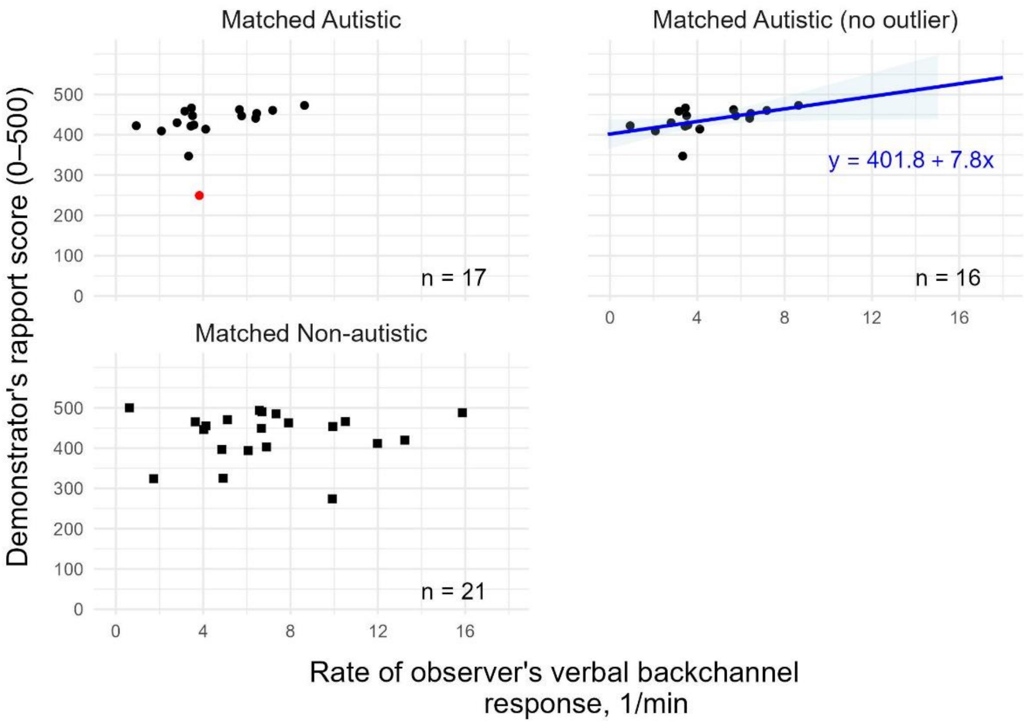


Figure 15. Relationship between observer’s rate of backchannel response and their demostrator’s rapport in each group of dyads (A and B). Panel A shows the association between the rate of each observer’s backchannel response and their demonstrator's rapport in each group: autistic (n = 17), non-autistic (n = 21) and mixed (n = 21). Panel B shows that demonstrators’ rapport as a function of observers’ rate of backchannel response was found only in autistic group and only when distinct outliers were removed (n = 15): slope (mean ± SE) – 6.307 ± 2.044, and intercept – 413.843 ± 10.062. Points represent individual dyads as crossovers between the observer’s rate of backchannel response and the demonstrator’s rapport. Autistic dyads are shown as square points, mixed dyads are shown as triangle points, and non-autistic dyads are shown as circles. Outliers are represented as red points. Points that are not outliers are shown in black.

Simple linear regression did not provide sufficient evidence of relationship in either the autistic or non-autistic dyad group (Figure 15B; autistic: $F_{1,15} = 2.036$, p-value = 0.1741; non-autistic: $F_{1,19} = 0.03409$, p-value = 0.8555). However, once outliers were removed from data in the autistic group, it provided strong evidence that demonstrator’s rapport score increased as with the observer’s rate of backchannel response (Figure 15B; $F_{1,13} = 9.518$, p-value = 0.00869, slope ± SE = 6.307 ± 2.044, intercept ± SE = 413.843 ± 10.062).

Discussion

There was no evidence that the length of task-oriented interaction was related to whether the observer and demonstrator were of the same or different neurotypes. This was to be expected, since the length of the interaction was likely largely determined by the demonstrator's ability to make the dog rather than by any other factor. However, the demonstrator's subjective experience during the interaction was related to the observer's neurotype. If the observer was of the same neurotype, the demonstrator experienced higher rapport than if they were of a different neurotype. This is consistent with previous findings of Crompton *et al* (2020b), who found that self-rated rapport was higher for matched rather than mixed dyads.

When examining verbal interaction, I found strong evidence that non-autistic observers verbally interacted with their demonstrators a higher proportion of time in matched rather than mixed dyads. When the demonstrator and observer were of different neurotypes, the observer's length of verbal interaction was reduced. There was no evidence of the same pattern for autistic observers. However, there was also no evidence that in mixed dyads, only autistic observers interacted less with their demonstrators. One might speculate that the autistic observers were more focused on task completion (i.e., understanding how to make a dog) and paid less attention to the affiliative aspects of the interaction, an idea that was recently put forward by Wehrle *et al* (2024). This could explain why no differences were found between autistic and mixed dyads. Interestingly, however, a higher proportion of verbal interaction in the autistic dyads was associated with increased rapport, albeit the evidence for this was weak and rather unreliable due to the small sample size and the presence of an outlier in the data. Considering all groups together, I found moderate evidence of a relationship between the proportion of verbal interaction and rapport. However, the effect size was quite small – on average, a 10% increase in verbal interaction was associated with only a 6.2 unit increase in rapport. Even if verbal interaction is associated with rapport, this association is likely to be quite small and unlikely to explain differences between matched and mixed dyads. Overall, my hypothesis that there is an increased preference for verbal interaction in matched dyads and that this preference leads to higher rapport was not supported by my results.

When examining verbal backchannel response, I found very strong evidence that non-autistic observers verbally backchannelled more frequently to their demonstrators when the demonstrators were also non-autistic. When the demonstrator and observer were of different neurotypes, the rate of observer's verbal backchannel response was decreased regardless of whether the observer was non-autistic or autistic. Moreover, I found moderate evidence that in matched dyads non-autistic observers backchanneled more frequently than autistic observers. This is consistent with the findings of Rifai *et al* (2022). They also found that participants in non-autistic dyads backchanneled more frequently compared to participants in autistic or mixed dyads. I also found that the rate of observer's backchannel response was associated with demonstrator's rapport in autistic dyads but not in non-autistic dyads. Importantly, this association was only present after removing an outlier from the data. Nonetheless, if one assumes that autistic observers were focused primarily on task completion rather than on the affiliative aspects of the interaction, it would follow that as the interaction extended beyond the mere task completion, the rate of observer's backchannel response increased, and this coincided with increased demonstrator’s rapport. Considering all groups together, I found weak evidence of a relationship between the rate of backchannel response and rapport. The evidence was strong, however, after removing a few outliers. Overall, my hypothesis that there is an increased frequency of verbal backchannel response in non-autistic dyads, and it leads to higher rapport was only partially supported by my results.

In addition, I also found moderate evidence that non-autistic observers were more likely to switch roles with their demonstrator than autistic observers. This was expected, since one of the autistic traits is a preference for repetitive and restricted behaviour (Kuenssberg *et al.* 2011).

This study has several significant limitations. First, the task that was given to the participants did not require them to verbally communicate with each other, since the demonstrators were simply asked to show (rather than explain) the observers how to make a dog. Since the demonstrators were involved in the process of making a dog, verbal communication likely played a less significant role than during regular conversation. Second, I only considered verbal communicative behaviour in my analysis, but given the nature of the task, nonverbal communicative behaviour (i.e., mutual gaze, gestures, and posture) may have played a more significant role in building rapport (Tickle-Degnen & Rosenthal, 1990). Third, I did not analyse the demonstrators’ verbal communicative behaviour. As a result, I was unable to determine whether there was any relationship between the verbal interaction of the demonstrator and the observer in different groups of dyads. Additionally, when calculating the rate of verbal backchannel response I used the entire length of the task-oriented interaction rather than the part during which the demonstrator spoke. Finally, participants were allowed to disclose whether they were autistic or non-autistic, which in itself could have affected rapport.

Future research

One potential direction for future research is to test the assumption that autistic people are more focused on completing the task than on the affiliative aspects of social interaction. Recently, Williams *et al* (2021) conducted a qualitative analysis of a series of interactions in which participants were asked to discuss the topic of loneliness. They found that, in most cases, rapport was much higher in autistic rather than mixed dyads. Quantitative analysis of such interactions, in which the role of task is minimised, could potentially test above assumption.

References

Baron-Cohen S, Wheelwright S, Skinner R, Martin J & Clubley E (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders* 31, 5-17.

Benjamini Y & Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57, 289-300.

Brugman H, Russel A & Nijmegen X (2004). Annotating Multi-media/Multi-modal Resources with ELAN. *LREC*, 2065-2068.

Chen YL, Senande LL, Thorsen M & Patten K (2021). Peer preferences and characteristics of same-group and cross-group social interactions among autistic and non-autistic adolescents. *Autism* 25, 1885-1900.

Crompton CJ, Hallett S, Ropar D, Flynn E & Fletcher-Watson S (2020c). 'I never realised everybody felt as happy as I do when I am around autistic people': A thematic analysis of autistic adults' relationships with autistic and neurotypical friends and family. *Autism* 24, 1438-1448.

Crompton CJ, Ropar D, Evans-Williams CV, Flynn EG & Fletcher-Watson S (2020a). Autistic peer-to-peer information transfer is highly effective. *Autism* 24, 1704-1712.

Crompton CJ, Sharp M, Axbey H, Fletcher-Watson S, Flynn EG & Ropar D (2020b). Neurotype-matching, but not being autistic, influences self and observer ratings of interpersonal rapport. *Frontiers in Psychology* 11, 2961.

Cutrone P (2005). A case study examining backchannels in conversations between Japanese–British dyads. *Multilingua* 24, 237-274.

Fletcher-Watson S & Happé F (2019). *Autism: A new introduction to psychological theory and current debate*, 2nd Ed, London UK: Routledge.

Freeman GH & Halton JH (1951). Note on exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika* 38, 141-149.

Ganesh S & Cave V (2018). P-values, p-values everywhere!. *New Zealand Veterinary Journal* 66, 55-56.

Halldestam M (2016). ANOVA-the effect of outliers. Uppsala University, Department of Statistics.

Johnston MG & Faulkner C (2021). A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. *New Phytol* 230, 23-26.

Kuenssberg R, McKenzie K & Jones J (2011). The association between the social and communication elements of autism, and repetitive/restrictive behaviours and activities: A review of the literature. *Research in Developmental Disabilities* 32, 2183-2192.

Lin XB, Lim CG and Lee TS (2022). Social Deficits or Interactional Differences? Interrogating Perspectives on Social Functioning in Autism. *Frontiers in Psychiatry* 13, 823736.

Milton D (2012). On the ontological status of autism: The ‘double empathy problem’. *Disability and Society* 27, 883–887.

Rifai OM, Fletcher-Watson S, Jiménez-Sánchez L & Crompton CJ (2022). Investigating markers of rapport in autistic and nonautistic interactions. *Autism in Adulthood* 4, 3-11.

Ritvo RA, Ritvo ER, Guthrie D, Yuwiler A, Ritvo MJ & Weisbender L (2008). A scale to assist the diagnosis of autism and Asperger’s disorder in adults (RAADS): A pilot study. *Journal of Autism and Developmental Disorders* 38, 213-223.

Sasson NJ, Faso DJ, Nugent J, Lovell S, Kennedy DP & Grossman RB (2017). Neurotypical peers are less willing to interact with those with autism based on thin slice judgments. *Scientific reports* 7, 1-10.

Tedeschi LO & Galyean ML (2023). A practical method to account for outliers in simple linear regression using the median of slopes. *Scientia Agricola* 81, 1-8.

The Language Archive (2023). ELAN (Version 6.6) [Computer software]. Nijmegen: Max Planck Institute for Psycholinguistics. URL: https://archive.mpi.nl/tla/elan [Accessed 26 April 2024]

Tickle-Degnen L & Rosenthal R (1990). The nature of rapport and its nonverbal correlates. *Psychological inquiry* 1, 285-293.

Ward N & Tsukahara W (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of pragmatics* 32, 1177-1207.

Wasserstein RL, Schirm AL & Lazar NA (2019). Moving to a world beyond “p< 0.05”. *The American Statistician* 73, 1-19.

Wechsler D (2011). *Wechsler Abbreviated Scale of Intelligence*, 2nd Ed, San Antonio US: Pearson.

Wehrle S, Vogeley K & Grice M (2024). Backchannels in conversations between autistic adults are less frequent and less diverse prosodically and lexically. *Language and Cognition* 16, 108-133.

Williams GL, Wharton T & Jagoe C (2021). Mutual (mis) understanding: Reframing autistic pragmatic “impairments” using relevance theory. *Frontiers in psychology* 12, 1277.

World Health Organization (2018). International Classification of Diseases, Eleventh Revision (ICD-11). URL: https://icd.who.int/browse11 [Accessed 26 April 2024]

Appendix

Methods

Coding Scheme

Coding scheme used for dyadic interaction coding in ELAN software.

Video coding scheme for demonstration (duration - seconds)

	Included	Excluded
Task-oriented Interaction	Starting from the moment when the investigator leaves the room, and the participants begin speaking about the task. Ending when one of the participants are finished the task: this is indicated by their moving on from discussing the task to discussing another topic at the end of the task, deciding to tell the investigator that they're finished with the task, or when the investigator enters, and the interlocutors confirm that they're finished.	Any interaction in the presence of the investigator, any interaction after the interlocutors have finished discussing how to make a dog but has not yet informed the investigator about it or are waiting for their return.

Video coding scheme for the observer's engagement level (duration - seconds)

	Included	Excluded
No/Minimal Verbal Engagement	The observer is silent or interacts verbally with the demonstrator at a minimal level (i.e., makes no comments or asks for no clarifications, backchannelling if present is spaced and is >10 s apart).	
Verbal Engagement	The observer actively engages with the demonstrator (i.e., makes any (both relevant and irrelevant) comments or asks for clarifications or backchanneling is frequent and is ≤10 s apart).	The gap between backchannels or comments or clarifications is longer than 10 s, any standalone comments or requests for clarifications (not proceeded or followed by any other engagement within 10 s period)
Role Swap	The observer verbally explains or physically shows the demonstrator how to make a dog	The gap between explanation comments or clarifications is longer than 10 s, anything that would be considered an explanation except the participant is not the observer

Video coding scheme for the observer's verbal backchannelling (rate and diversity).

	Included	Excluded
Verbal backchannels	When the participant is the observer and the listener: (1) laughs (breathy and vocal), chuckles (inward and quite), giggles, (2) non-lexical speech elements (e.g., 'mhmm', 'yay', 'aha'), (3) single word lexical speech elements not in response to a direct question (e.g., 'ok', 'right/alright', 'yes/yeah') used to convey understanding or/and attention, (4) short phrases conveying understanding (e.g., 'That makes sense', 'Got it') or attention (e.g., 'I'm following') or repeated phrases (e.g., D: 'So that's how you do it.' O: 'That's how you do it.').	When the participant is the observer and the listener: (1) smiles without laughing (2) gasps or heavy breathing, (3) single word lexical speech elements in response to the direct question (e.g., D: 'Should I do it this way?' O: 'Sure' or 'Ok'), tag question or polar question, (4) Requests for clarification, task-specific/general comments or suggestions, short answers to the direct question (e.g., D: 'Does it makes sense?' O: 'Makes sense.') or anything else that is a part of a conversational turn. Anything that would be considered a backchannel except the participant is not the observer and the listener.