

(Deictic) relational responding is not related to autism traits in the general population

Maura Nevejans^{a,*}, Jamie Cummins^{a,b,c}, Jan De Houwer^a, Emiel Cracco^a, Jan R. Wiersema^a

^a Department of Experimental Clinical and Health Psychology, Ghent University, Belgium

^b Institute of Marketing and Business Administration, University of Bern, Switzerland

^c Institute of Psychology, University of Bern, Switzerland

ARTICLE INFO

Keywords:

Relational frame theory
Deictic framing
Autism
Relational responding

ABSTRACT

Relational Frame Theory (RFT) proposes that a specific type of relational responding, namely “deictic” relating, is critically involved in perspective-taking and Theory of Mind (ToM). Therefore, it has been proposed that presumed ToM difficulties experienced by individuals with autism may be attributed to difficulties with deictic relating. However, rigorous studies testing whether altered deictic relating is associated with autism are lacking. To this end, we conducted two experiments ($N_{\text{exp1}} = 134$, $N_{\text{exp2}} = 120$) in which we tested the relationship between autism traits and performance on different types of relational responding, including deictic relating, in a general population sample. In Experiment 1, we investigated the relationship between autism traits and performance on eight types of relational responding via an improved version of the Relational Abilities Index (RAI). In Experiment 2, we adjusted the deictic relating scale to include stimulus functions more directly relevant to perspective-taking. Contrary to our predictions, deictic relating did not correlate significantly with autism traits. Moreover, none of the other RAI subscales correlated reliably with autism traits. These findings question the claims made by RFT about the relationship between autism and relational responding and warrant a reassessment of the accuracy of this position.

The ability to take another person's perspective is considered essential for successful human social interaction (Galinsky et al., 2005; Todd & Galinsky, 2014). Cognitive psychologists often study perspective-taking by focusing on a skill commonly referred to as Theory of Mind (ToM), defined as the ability to attribute mental states to oneself and others (Premack & Woodruff, 1978). Because of its importance for social interaction, differences in ToM have been proposed as an explanation for the social difficulties observed in individuals with autism¹ (Baron-Cohen et al., 1985; Frith et al., 1994; Long et al., 2025). In line with this, numerous studies have found reduced ToM abilities in children and adults with autism (e.g., Baron-Cohen et al., 1985; Callenmark et al., 2014; Happé, 1994; Livingston et al., 2024), and lower ToM abilities to be associated with more autism traits in clinical as well as in neurotypical samples (Livingston et al., 2024; Nijhof et al., 2017). However, not all studies found evidence for altered ToM in autism (for a critical review, see Gernsbacher & Yergeau, 2019), and the processes

underlying ToM remain unclear (Quesque & Rossetti, 2020).

Relational Frame Theory (RFT; Hayes et al., 2001) is a behavioral-analytic theory of human cognition, which proposes that arbitrarily applicable relational responding (AARR) is central to complex human behavior, including language, reasoning, and, importantly, perspective-taking and ToM (Hughes & Barnes-Holmes, 2016). AARR refers to the ability to derive relations between stimuli that share no physical features and to do so without direct training or instruction (Hayes et al., 2001; Hughes & Barnes-Holmes, 2016). Moreover, AARR is seen as a form of generalized operant behavior, implying that it is a learned and trainable skill (Healy et al., 2000; Ming et al., 2014).

RFT typically categorizes relational responding into various types of relating, referred to as “relational frames”, such as frames of coordination (e.g., A is the same as B), opposition (e.g., A is opposite to B), or temporality (e.g., A comes after B). RFT theorists have claimed that one of these, deictic relating (DR), “provides the core behavioral unit that

* Corresponding author.

E-mail address: maura.nevejans@ugent.be (M. Nevejans).

¹ We use an abbreviated version of the diagnostic term and refer to a person with a diagnosis of autism spectrum disorder, as a person with autism. With this, we do not intend to take a stance in the ongoing person-first versus identity-first debate, in which there is currently no consensus (Flowers et al., 2023) and acknowledge and respect different language preferences (De Laet et al., 2023).

may be involved in behaviors referred to as ToM in the mainstream psychological literature” (Hempkin et al., 2024, p. 109), although recent work suggests that ToM might involve more than responding in accordance with basic deictic frames (Hempkin et al., 2024; Kavanagh et al., 2020). Specifically, the theory assumes that perspective-taking is the ability to respond to a family of deictic frames, which specify relations in terms of the perspective of the speaker (Barnes-Holmes et al., 2004; Hayes et al., 2001). Three deictic relations are central to DR: I versus YOU, HERE versus THERE, and NOW versus THEN. According to RFT, DR emerges through a history of responding to questions involving these critical frames, for example, “What are you doing over there?” or “Where were you yesterday?” (Barnes-Holmes et al., 2004; Hayes et al., 2001; McHugh et al., 2004). Each time such questions are asked, the physical context changes, whereas the relational properties of I versus YOU, HERE versus THERE, and NOW versus THEN in the question remain constant (Hayes et al., 2001). By answering these questions, children learn to describe their own perspective in relation to others, thereby abstracting the underlying relational properties (Barnes-Holmes et al., 2004; Rehfeldt et al., 2007), which serves as a foundation for developing perspective-taking skills.

Based on the proposed link between DR and ToM (Barnes-Holmes et al., 2004), RFT research has often been directed at training DR in children (Barron et al., 2019; Gilroy et al., 2015; Jackson et al., 2014) and adolescents (Belisle et al., 2016; Lovett & Rehfeldt, 2014) with autism, to target their assumed perspective-taking deficits. To do so, most training studies use the so-called “Barnes-Holmes protocol” (see McHugh et al., 2004), in which DR is assessed across three levels of relational complexity: simple, reversed, and double reversed. In each trial, the experimenter gives the participant certain information, such as “You are sitting there on the blue chair. I am sitting here on the black chair.” After this description, the experimenter could ask the participant where they are sitting, which requires them to respond by the I-YOU and HERE-THERE deictic frames. In a reversal trial, there is an additional statement about a reversal in one of the deictic relations, for example, “If I were you and you were me” (I-YOU reversal). An example of a double reversal, considered to have the highest relational complexity, is “If here were there and there were here and if I were you and you were me”, in which both the HERE-THERE and I-YOU deictic frames are reversed (Barnes-Holmes et al., 2004; Rehfeldt et al., 2007).

Although several studies have focused on training DR in individuals with autism, surprisingly, only one study has attempted to test whether individuals with autism actually perform worse on DR. Rehfeldt et al. (2007) used the Barnes-Holmes protocol to compare DR between children with and without autism. However, whereas the researchers claimed there were differences between the groups, the descriptive differences between the groups were not statistically significant. Furthermore, the study had a small sample size ($N = 18$) and was therefore likely underpowered.

Apart from the small sample size, the Barnes-Holmes protocol itself is also suboptimal for several reasons. Firstly, there are redundant trials: for each trial, participants must respond to two questions based on the information provided; however, the response to the second question is whichever of the two options was not picked for the first question. Consequentially, if a participant answered the first question correctly, they would also answer the second correctly; if they answered the first incorrectly, they would also answer the second incorrectly. Secondly, the measure has not been subjected to psychometric examination, particularly its reliability. If a measure is not internally reliable (i.e., it correlates poorly with itself), then it cannot correlate strongly with other measures. Furthermore, it is not clear whether the trials in the Barnes-Holmes protocol have sufficient variation in difficulty to assess differences in perspective-taking skill levels (Cummins, 2023).

These issues relating to psychometrics were discussed in a recent paper by Cummins (2023) in the context of another measure that assesses DR (as well as seven other frames) in adults: the Relational Abilities Index (RAI; Colbert et al., 2017, 2020; Cummins et al., 2023).

Cummins (2023) demonstrated that the trials in the RAI were insufficiently varied in difficulty to capture individual differences in relational responding skills in adults. To resolve this issue, Cummins and De Houwer, 2026 developed a new version of the RAI with improved psychometric properties. This improved version of the task, which required selecting correct statements given some relational information, rather than simply (dis)affirming a given statement as in the older versions (Colbert et al., 2017, 2020; Cummins et al., 2023), showed greater variation in difficulty, implying that the measure could assess a wider range of ability levels.

Across two experiments, the current research builds on this work to examine the relationship between relational responding, including DR, and autism traits in a large sample from the general population. In both experiments, we used a dimensional approach to autism by assessing autism traits in the general population. In line with this approach, research suggests that autism traits do not occur solely within a discrete category, but rather represent a continuum dimension across the general population, with clinically diagnosed individuals being at the extreme tail of that distribution (Constantino & Todd, 2003; Happé & Frith, 2020; Robinson et al., 2011). Initial research within this dimensional approach focused on undiagnosed family members of individuals with autism, and found that they often display a pattern of subclinical yet qualitatively similar characteristics to the core features of autism, now referred to as the Broader Autism Phenotype (BAP; Piven et al., 1997; for a review, see Rubenstein & Chawla, 2018). Since then, research has increasingly focused on examining the correlates of higher autism trait scores in the general population rather than solely in diagnosed individuals or their family members (Happé & Frith, 2020). By doing so, studies have shown that (subclinical) autism traits are also associated with differences in, for example, sensory sensitivity (Robertson & Simmons, 2013) and social cognition (Gökçen et al., 2016; Goris et al., 2020; Hermans et al., 2009; Nijhof et al., 2017) in the general population. Genetic research further indicates substantial overlap between genetic influences on subclinical autism traits and those on diagnosed autism (Happé & Frith, 2020). Hence, applying the dimensional approach within the general population can yield valuable insights into autism and its associated traits, and aligns with findings from genetic studies.

We assessed autism traits in the general population through two self-report questionnaires, the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001) and the Comprehensive Autistic Trait Inventory (CATI; English et al., 2021). Although the AQ is a widely used instrument for assessing autism traits in the general population, previous research has reported several psychometric limitations, including its limited fit with the more recent (two-domain) conceptualization of autism (Kloosterman et al., 2011), and concerns regarding its applicability across different populations, such as women (Belcher et al., 2023; Hechler et al., 2025). Therefore, we included the recently developed CATI as a secondary measure of autism traits. This self-report questionnaire aligns with the two-domain conceptualization of autism, shows sound psychometric properties, and includes items designed to identify autism traits in women (English et al., 2021).

1. Experiment 1

In Experiment 1, we examined the relationship between scores on the eight subscales of the improved RAI (Cummins and De Houwer, 2026) and scores on the AQ and CATI in a large general population sample. By using the improved RAI and a large sample size, this experiment addresses key limitations of previous research on DR in autism. Our primary research question (RQ) concerned the relationship between DR and autism traits (RQ1). We predicted that participants’ accuracy on the deictic subscale of the RAI would correlate negatively with AQ scores, implying that individuals with more autism traits would have lower performance on the deictic subscale.

Apart from the focus on differences in DR in autism (Hempkin et al., 2024), several studies have focused on training other types of relating in

children with autism to improve verbal abilities (Dunne et al., 2014; Kent et al., 2017; O'Connor et al., 2009), and some claim that difficulty with relating is one of the core problems for many individuals with autism (Ming et al., 2014). However, previous research used small samples and never directly compared performance on relational responding between individuals with and without autism (Kent et al., 2017). The RAI allowed us to not only assess DR but also other relational frames, allowing for a direct comparison of performance across different frames and their relationship with autism traits. Regarding the other seven RAI subscales (RQ2), we had no specific predictions about their relationship with AQ scores. If autism traits are specifically related to performance on DR, as RFT would predict, we would expect no significant correlation between AQ scores and any of the remaining seven RAI subscales. However, if autism traits are associated with worse performance on relational responding in general, we would expect AQ scores to correlate negatively with all subscales.

1.1. Method

The data and analysis files for Experiment 1 are available on the Open Science Framework (OSF): <https://osf.io/6g7un>. This experiment was preregistered (<https://osf.io/tsjge>) and approved by the ethical committee of the Faculty of Psychology and Educational Sciences of Ghent University (reference number 2024-062).

1.1.1. Participants

We preregistered a maximum sample size of 150, matching the number of participants included in the final analysis sample after exclusions in a previous study (Cummins and De Houwer, 2026). Accordingly, 150 English-speaking UK residents who completed the RAI in Cummins and De Houwer, 2026 were invited to participate in this study through Prolific Academic. One hundred thirty-four participants (76 female, 57 male, 1 non-binary, $M_{\text{age}} = 31.70$, $SD_{\text{age}} = 6.24$) completed the study and were paid £5 (£10/hour) for their participation. Participants had, on average, completed 16.54 years of education ($SD = 2.87$) since the first year of primary school, not counting repeated years.

1.1.2. Materials

Relational Abilities Index. The RAI included 128 trials equally divided across the eight subscales: opposition/coordination, difference, quantity, temporal, containment, deictic, analogical, and arithmetical relations. Before the main task, participants completed three practice trials. On all trials, participants read 2-to-4 premises about the relationship between arbitrary nonsense stimuli, for example: "CUG is the same as JOM, JOM is opposite to VEK" (opposition subscale). Below the sentences, there were 2-to-4 buttons with additional statements about the relation between the non-words (e.g., "VEK is the same as CUG"). On some trials, participants were instructed to select all buttons with true statements, whereas on other trials, they were asked to select all buttons with false statements, based on the given premises. Each trial differed in the number of given premises (2, 3, 4), the number of response options (2, 3, 4), the number of true vs. false statements, and whether they had to indicate true or false statements (select all *true* statements; select all *false* statements). Some subscales followed a different structure. Specifically, the analogy and arithmetical subscales always included two premises, and the deictic subscale included two premises and a reversal statement for one (single reversal) or two (double reversal) of the three deictic relations (I-YOU, HERE-THERE, NOW-THEN). Further, in the deictic subscale, a maximum of two of the additional statements could be true. To confirm their choice, participants clicked on the "submit" button, after which the next trial started. There was always at least one correct option. If participants had not pressed the submit button before the response deadline of 60 s, no response was registered and the next trial started automatically. A trial was coded as correct if all correct statements (i.e., either all *false* or all *true* statements, depending on the instruction) were selected and all incorrect statements were unselected.

The trial was coded as incorrect if participants made one or more mistakes or did not press the submit button before the trial deadline. The subscale accuracy scores were calculated by taking the average accuracy across all 16 trials within each subscale. Fig. 1 represents an example trial of the deictic subscale in the RAI.

The AQ. The AQ contains 50 items, equally divided into five subscales: communication (e.g., "I enjoy social chit-chat. (reversed)"), social skill (e.g., "I find it difficult to work out people's intentions."), imagination (e.g., "I find making up stories easy. (reversed)"), attention to detail (e.g., "I am fascinated by dates."), and attention switching (e.g., "I like to plan any activities I participate in carefully."). Participants indicated how much they agreed with each statement on a four-point Likert scale, with the options "definitely agree", "agree", "disagree", and "definitely disagree". The AQ showed good internal consistency (English et al., 2021) and acceptable test-retest reliability (Baron-Cohen et al., 2001) in previous research. In line with previous research (Goris et al., 2020; Kloosterman et al., 2011), we used dimensional scores for our analysis, in which item scores range from 1 to 4, as these capture more variability, with higher scores indicating more autism traits. In the typical binary scoring of the AQ, the categories "agree" and "definitely agree" are taken together as a score "1" and "disagree" and "definitely disagree" are taken together as a score "0".

The CATI. The CATI was included as a secondary measure of autism traits and contains 42 items, equally divided into six subscales: social interactions (e.g., "I find it difficult to make new friends."), communication (e.g., "I have difficulty understanding someone else's point-of-view."), social camouflage (e.g., "I rely on a set of scripts when I talk with people."), cognitive rigidity (e.g., "It annoys me when plans I have made are changed."), repetitive behaviors (e.g., "I engage in certain repetitive actions when I feel stressed."), and sensory sensitivity (e.g., "I am sensitive to flickering lights."). Participants indicated how much they agreed with each statement on a five-point Likert scale, with the options "definitely disagree", "somewhat disagree", "neither agree nor disagree", "somewhat agree", and "definitely agree". The CATI showed good internal consistency in previous research (English et al., 2021) and was scored according to the general guidelines, with higher scores indicating more autism traits (English et al., 2021).

1.1.3. Procedure

The study was conducted in two parts. Both the RAI (part 1) and the questionnaires (Part 2) were programmed in JavaScript, using the Lab.js online study builder (Henninger et al., 2022).

Part 1. Part 1 of this experiment was conducted as part of a different study (Cummins and De Houwer, 2026). Participants first read and signed the informed consent and provided demographic information (age and gender). To indicate their gender, they could select between "female", "male", "other", and "prefer not to say". Participants then completed the RAI.

Part 2. The same participants were invited for a follow-up survey via the Prolific website. Participants first read and signed the informed consent, which explicitly mentioned their data would be linked to their RAI data from Part 1, and once again provided their age and gender, as well as how many years of education they had completed since the first year of primary school, not counting repeated years. They then filled out the AQ and CATI in a random order and completed another task, the function transformation task, which was part of a different study and will not be discussed further (Finn et al., 2026). The function transformation task was always presented last. At the end, participants received a short debriefing. The total duration of this part was approximately 25 min.

1.1.4. Analysis

All data were processed and analyzed using R (Version 4.3.1; R Core Team, 2023).

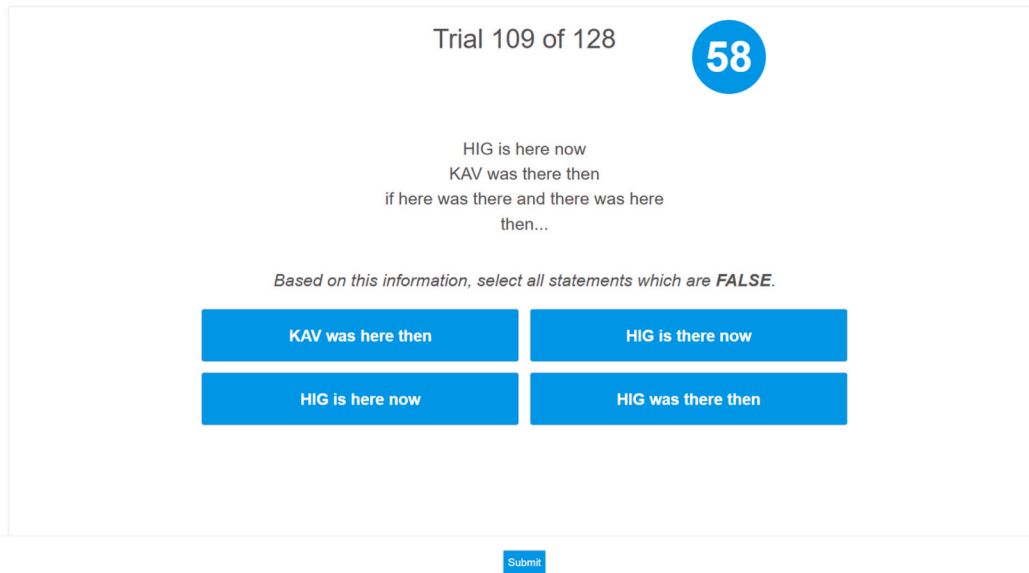


Fig. 1. Example Trial of the Deictic Subscale in the RAI of Experiment 1

Note. This example trial includes a spatial reversal (HERE-THERE), with four response options. In the current trial, participants are required to select false statements (i.e., the two statements on the bottom). The “58” in the circle indicates the time left in seconds.

1.1.4.1. Correlational Analyses. We computed Spearman's rank-ordered correlation between AQ scores and accuracy on the deictic subscale (RQ1) and between AQ scores and accuracy on the remaining seven RAI subscales (RQ2). Finally, we performed the same analyses for the CATI scores. As preregistered, we used Spearman's rank-ordered correlation because the Pearson correlation's assumptions were not met. We report the corresponding 95 % CIs and raw *p* values. We computed Bayes factors (BF) for the correlations of interest using Jeffrey's default BF hypothesis test (Ly et al., 2016), available in the BayesFactor package in R (Morey et al., 2024). The preregistered exploratory analysis of the correlations between accuracy on the deictic subscale and the different AQ subscales, along with its results, can be found in Appendix A.

1.2. Results

1.2.1. Autism trait questionnaires

The AQ. Using dimensional scoring, the total AQ scores ranged from 73 to 180 ($M = 117.13, SD = 16.98$) on 200, with a Cronbach's alpha (α) of .89 (50 items), 95 % CI [.84, .91], indicating good internal consistency. The binary scores ranged from 7 to 48 ($M = 21.30, SD = 7.88$) on 50. Fifteen (11.20 %) participants scored equal to or above the clinical cut-off of 32/50, which may indicate autism (Baron-Cohen et al., 2001).

The CATI. The total CATI scores ranged from 54 to 175 ($M = 117.90, SD = 27.81$) on 210. The CATI showed excellent internal consistency, 42 items, $\alpha = .94$, 95 % CI [.92, .95]. Using the general cut-off of 148/210, 20 individuals (14.93 %) scored above or equal to the cut-off (English et al., 2021). The correlation between the CATI and AQ was strong, $r(132) = .73, p < .001$, 95 % CI [.64, .80].

1.2.2. RQ1: correlation between AQ scores and accuracy on the deictic subscale

There was no significant correlation between participants' AQ scores and their accuracy on the deictic subscale, $r(132) = .05, p = .587$, 95 % CI [-.12, .22]. The BF analysis, $BF_{10} = 0.23$, provided evidence in favor of a null model over a model assuming an effect, indicating that the data were 4.35 times more likely given the null model. Fig. 2 shows the scatterplot for the AQ scores and accuracy on the deictic subscale.

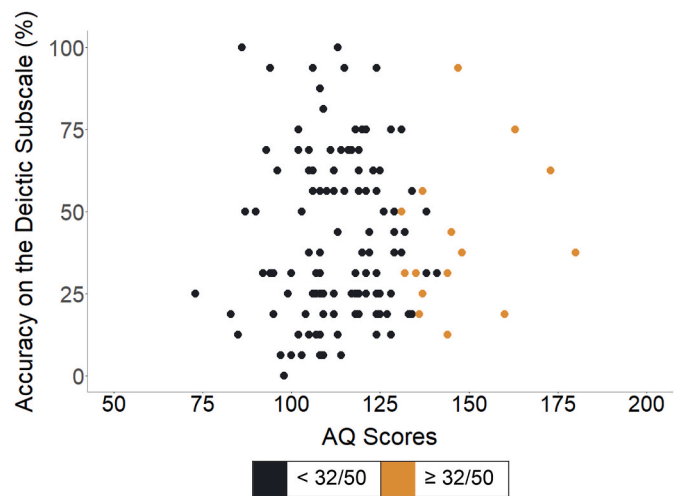


Fig. 2. Scatterplot Including the AQ Scores and Accuracy on the Deictic Subscale in Experiment 1

Note. AQ stands for Autism Spectrum Quotient. Accuracy is expressed in percentage correct responses (%). Orange dots indicate individuals scoring equal to or above the AQ's clinical cut-off of 32/50 when using binary scoring.

1.2.3. RQ2: correlations between AQ scores and accuracy on the other RAI subscales

None of the other RAI subscales correlated significantly with the participants' AQ scores, all $|rs| \leq .16$, all $ps \geq .057$. Table 1 provides an overview of all correlations between the AQ scores and RAI subscales.

1.2.4. Correlations between CATI scores and accuracy on the RAI subscales

The CATI scores did not correlate significantly with the accuracy on the deictic subscale, $r(132) < .01, p = .958$, 95 % CI [-.17, .17]. This was again supported by the BF analysis, $BF_{10} = 0.20$, which indicated that the data were 5.00 times more likely under a null model.

There was a small, significant correlation between the CATI scores and accuracy on the analogy subscale, $r(132) = -.18, p = .042$, 95 % CI [-.34, -.01]. This suggests that higher CATI scores were related to worse performance on the analogy subscale. The BF analysis indicated that the

Table 1
Correlations between the AQ scores and RAI subscales.

RAI subscale	Correlation with AQ	95 % CI	<i>p</i> value
Analogy	-.03	[-.20, .14]	.699
Containment	.07	[-.11, .23]	.454
Deictic	.05	[-.12, .22]	.587
Difference	.16	[-.01, .33]	.057
Arithmetical	.00	[-.17, .17]	.995
Opposition	.08	[-.10, .24]	.385
Quantity	.09	[-.08, .26]	.283
Temporal	.07	[-.10, .24]	.394

Note. RAI stands for Relational Abilities Index. AQ stands for Autism Spectrum Quotient. *p* values and 95 % CIs are uncorrected.

data were only 1.47 times more likely given a model assuming an effect than one assuming no effect. CATI scores did not correlate significantly with any of the other RAI subscales, all $|rs| \leq .10$, all $ps \geq .271$. Table 2 provides an overview of all correlations between the CATI scores and RAI subscales.

1.3. Discussion

Contrary to our primary prediction, we did not find a significant correlation between participants' performance on the deictic subscale of the RAI and their AQ scores, with our BF analysis showing evidence for the absence of a correlation. Similarly, none of the other RAI subscales correlated significantly with AQ scores. Regarding the CATI, we found no correlation between CATI scores and performance on the deictic subscale, but we did find a significant correlation between CATI scores and performance on the analogy subscale. However, our BF analysis prompted us to be cautious, with a BF close to 1.

One potential explanation for the absence of a relationship between DR and autism traits in Experiment 1 is that, although our task had improved psychometric properties compared to previous versions of the RAI (Cummins, 2023), our assessment might have lacked important characteristics of the typical Barnes-Holmes protocol. First, whereas the typical DR assessment includes premises about "I" and "YOU", the deictic subscale in the RAI included non-words. However, pronouns such as I and you are inherently part of perspective-taking, and differences in the use and processing of "I" and "you" have been found in autism (Mizuno et al., 2011). Therefore, the use of pronouns could be essential for the assessment of DR. Second, the premises in the RAI included very abstract scenarios, with statements like "TAK was there then", whereas the Barnes-Holmes protocol includes concrete, more realistic scenarios, such as "You are sitting there on the blue chair", and others have even adapted the protocol to include even more diverse trials, for example, by using children's book stories (e.g., Davlin et al., 2011; Gilroy et al., 2015; Lovett & Rehfeldt, 2014). Furthermore, Montoya-Rodríguez et al. (2018) noted that the repetition of scenarios, which is present in both our task and the original Barnes-Holmes protocol, could cause disengagement. Altogether, these limitations of our

Table 2
Correlations between the CATI scores and accuracy on the RAI subscales.

RAI subscale	Correlation with CATI	95 % CI	<i>p</i> value
Analogy	-.18	[-.34, -.01]	.042
Containment	-.01	[-.18, .16]	.906
Deictic	.00	[-.17, .17]	.958
Difference	.06	[-.12, .22]	.524
Arithmetical	-.03	[-.20, .14]	.759
Opposition	.00	[-.17, .17]	.980
Quantity	.10	[-.08, .26]	.271
Temporal	.02	[-.15, .19]	.804

Note. RAI stands for Relational Abilities Index. CATI stands for Comprehensive Autistic Trait Inventory. *p* values and 95 % CIs for the Spearman's rank-ordered correlations are uncorrected.

task could have influenced its relationship with autism traits. Therefore, we conducted a second experiment with an improved DR measure.

2. Experiment 2

For Experiment 2, we developed a novel DR task, following the structure of the improved RAI but closer to the format of the Barnes-Holmes protocol. Further, each trial described a concrete scenario without repetitions. We also included a description of the emotions of the individuals (I and/or you) in some of the trials, similar to the task used by Lovett and Rehfeldt (2014), to explore potential effects of emotional vs. action scenarios on DR.

As in Experiment 1, we tested the relationship between DR and autism traits (RQ1) and expected to find a negative correlation between accuracy on the DR task and total AQ scores. To examine whether this correlation is specific to DR, we also tested the relationship between AQ scores and performance on the analogy subscale of the RAI (RQ2). We chose to include the analogy subscale because this subscale significantly correlated with CATI scores in Experiment 1. We added the CATI as a secondary assessment of autism traits, with the same predictions for the CATI as for the AQ.

2.1. Method

The materials, data, and analysis files for Experiment 2 are available on the OSF: <https://osf.io/6g7un>. This experiment was preregistered (<https://osf.io/tb4pu>) and approved by the ethical committee of the Faculty of Psychology and Educational Sciences of Ghent University (reference number 2024-062).

2.1.1. Participants

To determine the minimum required sample size, we conducted a simulation-based power analysis for our primary prediction (RQ1). We simulated data based on the AQ scores and the accuracy scores on the deictic subscale from Experiment 1. The power analysis indicated that a sample size $N = 125$ would allow us to detect a correlation of $r = -.30$ with 90.71 % power, using a significance threshold of 0.05 (see <https://osf.io/6g7un> for the full simulation rationale and code).

Participants were recruited via Prolific Academic. In total, 164 participants started the experiment through Prolific. Participants who completed the study were reimbursed £8 (£8/hour) for their participation. Once 125 complete submissions had been collected, data were screened for exclusion criteria (see Analysis section), and excluded submissions were replaced through additional recruitment. This process was repeated until we obtained the preregistered minimum required sample size from the power analysis ($N = 125$). However, due to five additional rejections (see Analysis section), the analysis sample contained 120 participants (60 female, 56 male, 4 non-binary, $M_{age} = 29.83$, $SD_{age} = 5.82$). Participants were native English speakers and had, on average, completed 15.60 years of education ($SD = 2.86$) since the first year of primary school, not counting repeated years. All participants signed informed consent before the start of the experiment.

2.1.2. Materials

Analogy subscale of the RAI. The analogy subscale of the RAI was identical to the subscale in Experiment 1, including 16 trials and an attention check.

Deictic relating task. The DR task included 48 trials and three attention checks, with a break every 16 trials, and followed the structure of the RAI. Apart from the attention checks, which had a fixed trial number, all trials appeared in random order. On each trial, participants read two premises describing someone (I or YOU) doing or feeling something at a certain place and time. An example premise could be "I was dancing at a concert earlier today" (see Fig. 3). Half of the trials included scenarios about actions (action trials), whereas the other half included scenarios about feelings (emotion trials). Further, half of the

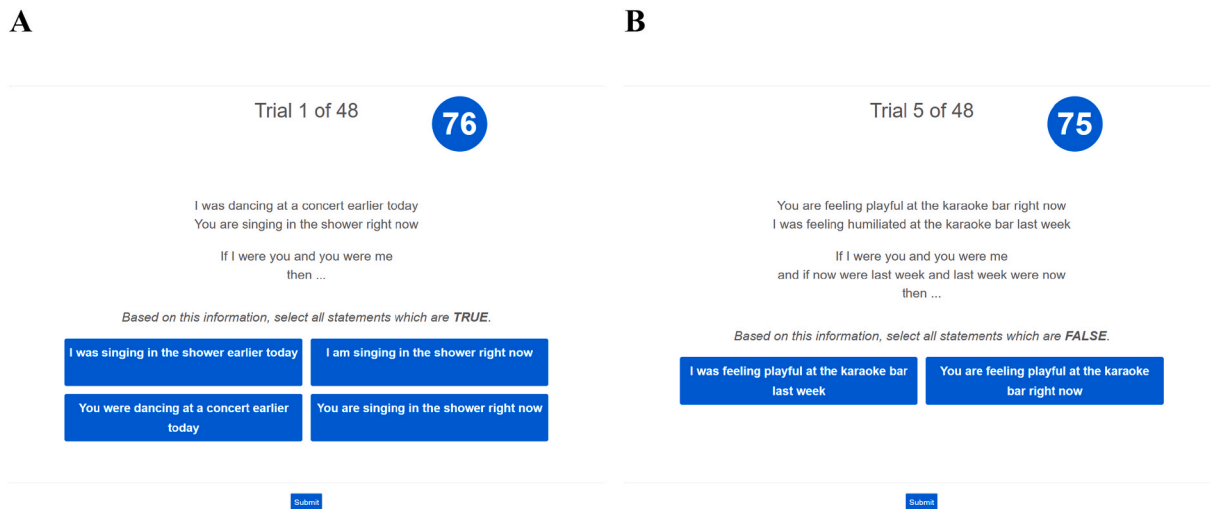


Fig. 3. Two Example Trials of the Deictic Relating Task in Experiment 2

Note. Fig. 3A shows an action trial in the deictic relating task, in which both premises are about an action, with an interpersonal reversal. Fig. 3B shows an emotion trial, including a repetition of the location and an interpersonal-temporal reversal.

trials included two premises describing the same person, or events happening at the same place or at the same time. Hence, trials could include repetition of person, place, or time, or no repetition. Below the premises, participants received an extra statement about a reversal, which could be interpersonal (“If I were you and you were me”), spatial (e.g., “If the garden were the supermarket, and the supermarket were the garden”), temporal (e.g., “If now were yesterday and yesterday were now”), or a combination of two statements (i.e., a double-reversal), resulting in six possible reversals. The combination of these factors created 24 trial conditions, in which the trial type (action trial, emotion trial), repetition (repetition, no repetition), and reversal varied.

Participants selected all buttons with true or false statements from 2- to 4 buttons based on the given information. Trials differed based on the number of buttons (2, 3, 4), the number of true (none true, one true, two true) vs. false (none false, one false, two false, three false) statements, and whether participants had to indicate true or false statements. There was always at least one correct option. If participants had not clicked the submit button before the response deadline of 80 s, no response was registered and the next trial started immediately. The scoring was identical to the scoring of the RAI in Experiment 1. Fig. 3 provides two example trials from the DR task in Experiment 2.

Autism trait questionnaires. The AQ and CATI were again included.

2.1.3. Procedure

Participants first provided informed consent and demographic information, including their age, gender (“male”, “female”, “non-binary”, “not listed”, or “prefer not to say”), and years of education since the first year of primary school, not counting repeated years. If they indicated that their gender was “not listed”, they could provide an alternative.

Participants first completed the RAI analogy subscale, preceded by instructions and two practice trials with elaborate feedback in case of an incorrect answer. They then received new instructions for the DR task and again completed three practice trials before starting the DR task. Finally, participants filled out the AQ and CATI in a randomized order. They were debriefed immediately after the experiment and could contact the researcher via Prolific for more information.

2.1.4. Analysis

All data were processed and analyzed using R (R Core Team, 2023).

Participant exclusions. Forty-four participants were excluded from all analyses following the preregistration plan. In line with previous research using the RAI (Cummins et al., 2023), participants were

excluded because they had incomplete data (22 participants) or because they had more than 20 % of trials with a response time faster or equal to 5 s (11 participants). Eleven participants failed more than one attention check. Initially, we decided to exclude only six participants, deviating from the preregistered plan, because the five remaining participants did not show further signs of low data quality (i.e., fast responding). However, after careful consideration, we chose to follow the preregistered plan and exclude them after all. Note that including these participants does not change any of the results.

Correlational analysis. As preregistered, we computed Spearman's rank-ordered correlations between the AQ scores and accuracy on the DR task (RQ1) and between the AQ scores and accuracy on the analogy subscale (RQ2). We repeated the same analyses with CATI scores. We report the corresponding 95 % CIs, raw *p* values, and Bayes factors, computed using the same method as in Experiment 1 (Morey et al., 2024). We also preregistered several exploratory analyses, for which the rationale, analysis, and results can be found in Appendix B.

2.2. Results

2.2.1. Autism trait questionnaires and relating tasks

Using dimensional scoring, the total AQ scores ranged from 74 to 155 ($M = 115.01$, $SD = 15.51$), with $\alpha = .85$, 95 % CI [.81, .88]. Binary scores ranged from 5 to 40 ($M = 20.58$, $SD = 7.07$). Seven (5.83 %) participants scored equal to or above the clinical cut-off score of 32/50 (Baron-Cohen et al., 2001).

The total CATI scores ranged from 51 to 192 ($M = 118.83$, $SD = 27.31$) on 210, with $\alpha = .94$, 95 % CI [.92, .95]. Seventeen (14.17 %) participants scored equal to or above the general cut-off score of 148. The correlation between the CATI and AQ scores was significant, $r(118) = .69$, $p < .001$, 95 % CI [.58, .77].

The DR task had an average split-half reliability (across 10,000 iterations) of $r = .95$, 95 % CI [.94, .97]. The analogy subscale had an average split-half reliability of $r = .76$, 95 % CI = [.68, .83].

2.2.2. Correlation between AQ scores and accuracy on the DR task

There was no significant correlation between the participants' AQ scores and accuracy on the DR task, $r(118) = .10$, $p = .255$, 95 % CI [-.08, .28]. The BF analysis, $BF_{10} = 0.39$, indicated that the data were 2.55 times more likely in the absence of a correlation. Fig. 4 represents the scatterplot for the AQ scores and accuracy on the DR task.

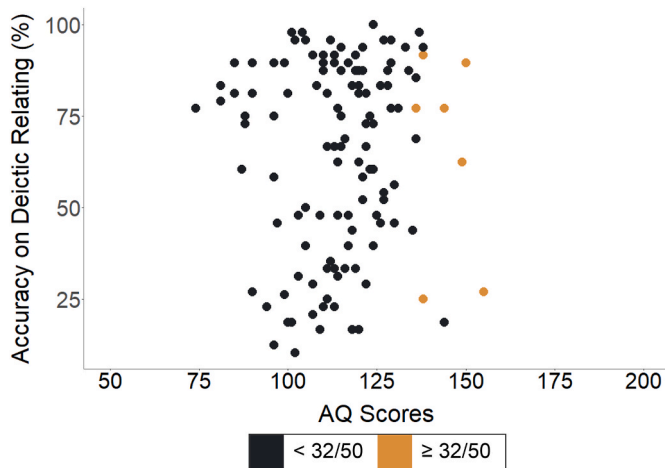


Fig. 4. Scatterplot Including the AQ Scores and Accuracy on the Deictic Relating Task in Experiment 2

Note. AQ stands for Autism Spectrum Quotient. Accuracy is expressed in percentage correct responses (%). Orange dots indicate individuals scoring equal to or above the AQ's clinical cut-off of 32/50 when using binary scoring.

2.2.3. Correlation between AQ scores and accuracy on the analogy subscale

There was no significant correlation between AQ scores and the participants' accuracy on the analogy subscale, $r(118) = .10$, $p = .275$, 95 % CI [-.08, .27]. The BF analysis indicated that the data were 2.68 times more likely under the null, $BF_{10} = 0.37$.

2.2.4. Correlation between CATI scores and accuracy on the DR task and analogy subscale

The CATI scores did not correlate significantly with performance on the DR task, $r(118) = -.09$, $p = .308$, 95 % CI [-.27, .09]. The BF analysis, $BF_{10} = 0.35$, indicated that the data were 2.89 times more likely under the null. Likewise, there was no significant correlation between the CATI scores and the performance on the analogy subscale, $r(118) = -.12$, $p = .199$, 95 % CI [-.29, .06]. The BF analysis, $BF_{10} = 0.46$, indicated that the data were 2.16 times more likely under a null model than a model assuming a correlation.

2.3. Discussion

In line with the results of Experiment 1, we found no significant correlation between DR and AQ or CATI scores in Experiment 2, with the BF analysis supporting this, despite the optimized DR assessment. We further did not observe a significant correlation between performance on the analogy subscale and the AQ scores or the CATI scores in Experiment 2.

3. General discussion

The current preregistered, well-powered study investigated the relationship between autism traits and relational responding in two experiments, using a dimensional approach to autism. In Experiment 1, we did not find a significant correlation between autism traits and accuracy on the deictic subscale of the RAI. Furthermore, apart from the significant correlation between accuracy on the analogy subscale and CATI scores, none of the other RAI subscales correlated with autism traits, which suggests that autism traits are not linked to general differences in relational responding. Previously reported difficulties in relational responding among children with autism – though never formally tested – might have resulted from the children's lower verbal abilities in those studies (Dunne et al., 2014; Kent et al., 2017), rather than from autism itself. In Experiment 2, despite our improved assessment, we again did not find a significant correlation between autism

traits and DR. We also included the analogy subscale, for which there was no significant correlation with autism traits, both for the AQ and the CATI. Because the analogy subscale was identical to the subscale in Experiment 1, we also conducted a post-hoc within-paper meta-analysis combining questionnaire scores and analogy subscale data from both experiments (see Appendix C). This analysis did not reveal a correlation with AQ scores ($r = .03$, $p = .654$) but a small negative correlation with CATI scores ($r = -.15$, $p = .018$).

Most critically, we did not find any relationship between autism traits and DR across both experiments. There are several potential explanations for the absence of this relationship. First, one could argue that DR difficulties might only become apparent in individuals with a formal diagnosis. However, previous research on autism that used a dimensional approach has shown that this approach can provide valuable insights into autism, and that ToM ability is related to autism traits in the general population (Gökçen et al., 2016; Nijhof et al., 2017). Hence, if DR is related to ToM, we would also expect differences in DR performance in individuals with higher autism trait scores. It has been argued, however, that ToM involves more than simple DR (Hempkin et al., 2024; Kavanagh et al., 2020). Nevertheless, even if DR is only one of the relational repertoires contributing to ToM, if it were relevant to autism, we would still expect it to correlate with autism traits in the general population, consistent with findings showing that autism traits are related to other social and non-social autism-relevant characteristics (Gökçen et al., 2016; Goris et al., 2020; Hermans et al., 2009; Nijhof et al., 2017; Robertson & Simmons, 2013). Note also that the current experiments included several individuals scoring above the clinical cut-off for the AQ (Exp1 = 11 %, Exp2 = 6 %) and CATI (Exp1 = 15 %, Exp2 = 14 %), which is more than in previous research (e.g., Goris et al., 2020). Still, further research is warranted to replicate the current findings in a sample of individuals with a formal autism diagnosis.

Second, because our sample consisted of adults, it remains possible that DR is different in children with autism. Having said this, ToM difficulties are still apparent in adults with autism (e.g., Livingston et al., 2024) and are associated with elevated autism traits in the general adult population (Gökçen et al., 2016; Nijhof et al., 2017). Hence, this would mean that, if DR is not different, another explanation is needed for ToM difficulties in adults with autism or elevated autism traits. However, as we did not include children in our study, we cannot exclude the possibility that DR may be different in children with autism. Hence, future research is warranted to investigate whether DR is different in children with autism or elevated autism traits, and linked to their ToM abilities.

Finally, related to the previous point, it is possible that the current operationalizations of DR are not associated with autism traits, nor with perspective-taking abilities. Although some initial studies reported a link between DR and ToM in small samples (e.g., Weil et al., 2011), other studies, including a well-powered study by Taylor et al. (2023), found no generalization of DR training to perspective-taking abilities (Jackson et al., 2014; Lovett & Rehfeldt, 2014). Recent criticism of deictic frames emphasizes theoretical issues with how DR is currently defined (Guinther, 2017; Taylor & Edwards, 2021), arguing that the Barnes-Holmes protocol may function as training in complex sentence structures or logical reasoning rather than perspective-taking (Taylor & Edwards, 2021). If this is true, it could explain the absence of a significant correlation with autism traits in the current study, even for the social subscales of the AQ (see Appendix A). Relatedly, an important limitation of the present and previous studies is that the DR tasks are highly decontextualized and far from real-life perspective-taking. For instance, prior research highlighted differences between DR tasks and ToM measures, such as the absence of references to private events (e.g., thoughts and beliefs) in the Barnes-Holmes protocol, even though private events are central to ToM (Montoya-Rodríguez & Molina-Cobos, 2019). Although our DR task included emotion-related words in some trials, responding to written trials remains far removed from perspective-taking in ecological contexts with real actions and emotions. However, to our knowledge, there are currently no assessments of DR

that are suitable for adults and combine improved ecological validity with sound psychometric properties.

To conclude, in contrast to our predictions, we found no significant relationship between autism traits and DR, despite the psychometric soundness of our DR tasks (Experiment 1, Experiment 2) and similarity with typical DR assessments used in RFT research (Experiment 2). Moreover, we found no evidence for any relationship between autism traits and relational responding in general. These findings are not in line with RFT's autism-related claims and highlight the need for a more rigorous and ecologically valid investigation of assumed differences in (deictic) relational responding. Future research should therefore focus on developing reliable and ecologically valid assessments of (deictic) relational responding and investigate whether performance on these measures relates to autism traits and associated (ToM) difficulties.

CRedit authorship contribution statement

Maura Nevejans: Writing – original draft, Writing – review & editing, Visualization, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jamie Cummins:** Writing – review & editing, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jan De Houwer:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Emiel Cracco:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Jan R. Wiersema:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Author note

MN is supported by a doctoral fellowship awarded by the Research Foundation Flanders (11E5225N). JDH is supported by the Methusalem Grant (01M00209) of Ghent University. Correspondence concerning this article should be addressed to Maura Nevejans, Department of Experimental Clinical and Health Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. E-mail: maura.nevejans@ugent.be. Ethical approval was granted by the ethical committee of the Faculty of Psychology and Educational Sciences with reference number 2024-062. This study was preregistered on the Open Science Framework: <https://osf.io/tsjge> (Experiment 1) and <https://osf.io/tb4pu> (Experiment 2). The data and analysis code are available on the Open Science Framework: <https://osf.io/6g7un>.

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jcbs.2026.100977>.

References

- Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-taking and theory of mind: A relational frame account. *Behavior Analyst Today*, 5(1), 15–25. <https://doi.org/10.1037/h0100133>
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”. *Cognition*, 21(1), 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- 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. <https://doi.org/10.1023/A:1005653411471>
- Barron, B. F., Verkuylen, L., Belisle, J., Paliliunas, D., & Dixon, M. R. (2019). Teaching “then-later” and “here-there” relations to children with autism: An evaluation of single reversals and transformation of stimulus function. *Behavior Analysis in Practice*, 12(1), 167–175. <https://doi.org/10.1007/s40617-018-0216-1>
- Belcher, H. L., Uglik-Marucha, N., Vitoratou, S., Ford, R. M., & Morein-Zamir, S. (2023). Gender bias in autism screening: Measurement invariance of different model frameworks of the autism spectrum quotient. *BJPsych Open*, 9(5), Article e173. <https://doi.org/10.1192/bjo.2023.562>
- Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal “I-You” deictic frames. *Journal of Applied Behavior Analysis*, 49(4), 965–969. <https://doi.org/10.1002/jaba.324>
- Callenmark, B., Kjellin, L., Rönqvist, L., & Bölte, S. (2014). Explicit versus implicit social cognition testing in autism spectrum disorder. *Autism*, 18(6), 684–693. <https://doi.org/10.1177/1362361313492393>
- Colbert, D., Dobutowitsch, M., Roche, B., & Brophy, C. (2017). The proxy-measurement of intelligence quotients using a relational skills abilities index. *Learning and Individual Differences*, 57, 114–122. <https://doi.org/10.1016/j.lindif.2017.03.010>
- Colbert, D., Malone, A., Barrett, S., & Roche, B. (2020). The relational abilities index+: Initial validation of a functionally understood proxy measure for intelligence. *Perspectives on Behavior Science*, 43(1), 189–213. <https://doi.org/10.1007/s40614-019-00197-z>
- Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry*, 60(5), 524–530. <https://doi.org/10.1001/archpsyc.60.5.524>
- Cummins, J. (2023). On the measurement of relational responding. *Journal of Contextual Behavioral Science*, 30, 155–168. <https://doi.org/10.1016/j.jcbs.2023.10.003>
- Cummins, J., & De Houwer, J. (2026). *An improved and dynamic tool for assessing relational reasoning skills* [Manuscript in preparation].
- Cummins, J., Nevejans, M., Colbert, D., & De Houwer, J. (2023). On the structure of relational responding. *Journal of Contextual Behavioral Science*, 27, 16–25. <https://doi.org/10.1016/j.jcbs.2022.11.006>
- Davlin, N. L., Anne Rehfeldt, R., & Lovett, S. (2011). A relational frame theory approach to understanding perspective-taking using children's stories in typically developing children. *European Journal of Behavior Analysis*, 12(2), 403–430. <https://doi.org/10.1080/15021149.2011.11434392>
- De Laet, H., Nijhof, A. D., & Wiersema, J. R. (2023). Adults with autism prefer person-first language in Dutch: A cross-country study. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-023-06192-3>
- Dunne, S., Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Murphy, C. (2014). Facilitating repertoires of coordination, opposition distinction, and comparison in young children with autism. *Behavioral Development Bulletin*, 19(2), 37–47. <https://doi.org/10.1037/h0100576>
- English, M., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., Enns, J. T., & Maybery, M. T. (2021). The comprehensive autistic trait inventory (CATI): Development and validation of a new measure of autistic traits in the general population. *Molecular Autism*, 12(1), 37. <https://doi.org/10.1186/s13229-021-00445-7>
- Finn, M., Cummins, J., Nevejans, M., & De Houwer, J. (2026). *A first assessment of the measurement properties of two scales of a function transformation task battery* [Manuscript in preparation].
- Flowers, J., Dawes, J., McCleary, D., & Marzolf, H. (2023). Words matter: Language preferences in a sample of autistic adults. *Neurodiversity*, 1. <https://doi.org/10.1177/27546330231216548>
- Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, 3(2), 108–124. <https://doi.org/10.1111/j.1467-9507.1994.tb00031.x>
- Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations*, 8(2), 109–124. <https://doi.org/10.1177/1368430205051060>
- Gernsbacher, M. A., & Yergeau, M. (2019). Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology*, 7(1), 102–118. <https://doi.org/10.1037/arc0000067>
- Gilroy, S. P., Lohr, E. R., Dodge, J., & Fiorello, C. (2015). Establishing deictic repertoires in autism. *Research in Autism Spectrum Disorders*, 19, 82–92. <https://doi.org/10.1016/j.rasd.2015.04.004>
- Gökçen, E., Frederickson, N., & Petrides, K. V. (2016). Theory of mind and executive control deficits in typically developing adults and adolescents with high levels of autism traits. *Journal of Autism and Developmental Disorders*, 46(6), 2072–2087. <https://doi.org/10.1007/s10803-016-2735-3>
- Goris, J., Brass, M., Cambier, C., Delplanque, J., Wiersema, J. R., & Braem, S. (2020). The relation between preference for predictability and autistic traits. *Autism Research*, 13(7), 1144–1154. <https://doi.org/10.1002/aur.2244>
- Guinther, P. M. (2017). Contextual influence over deriving others' true beliefs using a relational triangulation perspective-taking protocol (RT-PTP-M1). *Journal of the Experimental Analysis of Behavior*, 108(3), 433–456. <https://doi.org/10.1002/jeab.291>
- Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24(2), 129–154. <https://doi.org/10.1007/BF02172093>
- Happé, F., & Frith, U. (2020). Annual research review: Looking back to look forward – Changes in the concept of autism and implications for future research. *Journal of Child Psychology and Psychiatry*, 61(3), 218–232. <https://doi.org/10.1111/jcpp.13176>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A Post-Skinnerian account of human language and cognition*. Kluwer Academic/Plenum Publishers.

- Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, 74(2), 207–227. <https://doi.org/10.1901/jeab.2000.74-207>
- Hechler, F. C., Tuomainen, O., Weber, N., Fahr, F., Karlek, B., Maroske, M., Misia, M., & Caruana, N. (2025). “What does ‘often’ even mean?” revising and validating the Comprehensive Autistic trait inventory in partnership with autistic people. *Molecular Autism*, 16(1), 7. <https://doi.org/10.1186/s13229-025-00643-7>
- Hempkin, N., Sivaraman, M., & Barnes-Holmes, D. (2024). Deictic relational responding and perspective-taking in autistic individuals: A scoping review. *Perspectives on Behavior Science*, 47(1), 107–137. <https://doi.org/10.1007/s40614-024-00397-2>
- Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). Lab.js: A free, open, online study builder. *Behavior Research Methods*, 54(2), 556–573. <https://doi.org/10.3758/s13428-019-01283-5>
- Hermans, E. J., van Wingen, G., Bos, P. A., Putman, P., & van Honk, J. (2009). Reduced spontaneous facial mimicry in women with autistic traits. *Biological Psychology*, 80(3), 348–353. <https://doi.org/10.1016/j.biopsycho.2008.12.002>
- Hughes, S., & Barnes-Holmes, D. (2016). Relational frame theory: The basic account. In *The wiley handbook of contextual behavioral science* (pp. 129–178). <https://doi.org/10.1002/9781118489857.ch9>
- Jackson, M. L., Mendoza, D. R., & Adams, A. N. (2014). Teaching a deictic relational repertoire to children with autism. *Psychological Record*, 64(4), 791–802. <https://doi.org/10.1007/s40732-014-0078-z>
- Kavanagh, D., Barnes-Holmes, Y., & Barnes-Holmes, D. (2020). The study of perspective-taking: Contributions from mainstream psychology and behavior analysis. *Psychological Record*, 70(4), 581–604. <https://doi.org/10.1007/s40732-019-00356-3>
- Kent, G., Galvin, E., Barnes-Holmes, Y., Murphy, C., & Barnes-Holmes, D. (2017). Relational responding: Testing, training, and sequencing effects among children with autism and typically developing children. *Behavioral Development Bulletin*, 22(1), 94–110. <https://doi.org/10.1037/bdb0000041>
- Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. A. (2011). Evaluation of the factor structure of the autism-spectrum quotient. *Personality and Individual Differences*, 50(2), 310–314. <https://doi.org/10.1016/j.paid.2010.10.015>
- Livingston, L. A., Shah, P., & Happé, F. (2024). Linearly integrating speed and accuracy to measure individual differences in theory of mind: Evidence from autistic and neurotypical adults. *Quarterly Journal of Experimental Psychology*, 77(2), 287–297. <https://doi.org/10.1177/17470218231165251>
- Long, E. L., Catmur, C., & Bird, G. (2025). The theory of mind hypothesis of autism: A critical evaluation of the status quo. *Psychological Review*. <https://doi.org/10.1037/rev0000532>
- Lovett, S., & Rehfeldt, R. A. (2014). An evaluation of multiple exemplar instruction to teach perspective-taking skills to adolescents with asperger syndrome. *Behavioral Development Bulletin*, 19(2), 22–36. <https://doi.org/10.1037/h0100575>
- Ly, A., Verhagen, J., & Wagenmakers, E.-J. (2016). Harold Jeffreys's default Bayes factor hypothesis tests: Explanation, extension, and application in psychology. *Journal of Mathematical Psychology*, 72, 19–32. <https://doi.org/10.1016/j.jmp.2015.06.004>
- McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *Psychological Record*, 54(1), 115–144. <https://doi.org/10.1007/BF03395465>
- Ming, S., Moran, L., & Stewart, I. (2014). Derived relational responding and generative language: Applications and future directions for teaching individuals with autism spectrum disorders. *European Journal of Behavior Analysis*, 15(2), 199–224. <https://doi.org/10.1080/15021149.2014.11434722>
- Mizuno, A., Liu, Y., Williams, D. L., Keller, T. A., Minshew, N. J., & Just, M. A. (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. *Brain*, 134(8), 2422–2435. <https://doi.org/10.1093/brain/awr151>
- Montoya-Rodríguez, M. M., Cobos, F. J. M., Montoya-Rodríguez, M. M., & Cobos, F. J. M. (2018). Assessing perspective-taking in children through different formats of deictic framing protocol. In *Behavior analysis*. IntechOpen. <https://doi.org/10.5772/intechopen.74539>
- Montoya-Rodríguez, M. M., & Molina-Cobos, F. J. (2019). Training perspective taking skills in individuals with intellectual disabilities: A functional approach. *Journal of Contextual Behavioral Science*, 14, 1–10. <https://doi.org/10.1016/j.jcbs.2019.08.003>
- Morey, R. D., Rouder, J. N., Jamil, T., Urbaneck, S., Forner, K., & Ly, A. (2024). *BayesFactor: Computation of bayes factors for common designs (Version 0.9.12-4.7) [R]*. <https://doi.org/10.32614/CRAN.package.BayesFactor>
- Nijhof, A., Brass, M., & Wiersema, J. R. (2017). Spontaneous mentalizing in neurotypicals scoring high versus low on symptomatology of autism spectrum disorder. *Psychiatry Research*, 258, 15–20. <https://doi.org/10.1016/j.psychres.2017.09.060>
- O'Connor, J., Rafferty, A., Barnes-Holmes, D., & Barnes-Holmes, Y. (2009). The role of verbal behavior, stimulus name ability, and familiarity on the equivalence performances of autistic and normally developing children. *Psychological Record*, 59(1), 53–74. <https://doi.org/10.1007/BF03395649>
- Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry*, 154(2), 185–190. <https://doi.org/10.1176/ajp.154.2.185>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>
- Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science*, 15(2), 384–396. <https://doi.org/10.1177/1745691619896607>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing [Computer software] Version 4.3.2. <https://www.r-project.org/>
- Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. K. (2007). Assessing relational learning deficits in perspective-taking in children with high functioning autism spectrum disorder. *Psychological Record*, 57(1), 23–47. <https://doi.org/10.1007/BF03395563>
- Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders*, 43(4), 775–784. <https://doi.org/10.1007/s10803-012-1608-7>
- Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., Plomin, R., & Ronald, A. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, 68(11), 1113–1121. <https://doi.org/10.1001/archgenpsychiatry.2011.119>
- Rubenstein, E., & Chawla, D. (2018). Broader autism phenotype in parents of children with autism: A systematic review of percentage estimates. *Journal of Child and Family Studies*, 27(6), 1705–1720. <https://doi.org/10.1007/s10826-018-1026-3>
- Taylor, T., & Edwards, T. L. (2021). What can we learn by treating perspective taking as problem solving? *Perspectives on Behavior Science*, 44(2), 359–387. <https://doi.org/10.1007/s40614-021-00307-w>
- Taylor, T., Sargis, R. J., & Edwards, T. L. (2023). Deictic framing performance fails to generalize to other perspective-taking tasks. *Psychological Record*, 73(3), 419–442. <https://doi.org/10.1007/s40732-023-00550-4>
- Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, 8(7), 374–387. <https://doi.org/10.1111/spc3.12116>
- Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *Psychological Record*, 61(3), 371–390. <https://doi.org/10.1007/BF03395767>