RESEARCH ARTICLE

"Why do they do it?": The short-story task for measuring fiction-based mentalizing in autistic and non-autistic individuals

Irina Jarvers[1] [ID] | Katrin Döhnel[2] | Lore Blaas[2] | Manuela Ullmann[2] |
Berthold Langguth[2] | Rainer Rupprecht[2] | Monika Sommer[2,3]

[1]Department of Child and Adolescent Psychiatry and Psychotherapy, University of Regensburg, Regensburg, Germany

[2]Department of Psychiatry and Psychotherapy, University of Regensburg at the Bezirksklinikum Regensburg, Regensburg, Germany

[3]Department of Psychology, Ludwig-Maximilians-University of Munich, Munich, Germany

Correspondence
Irina Jarvers, Department of Child and Adolescent Psychiatry and Psychotherapy, University of Regensburg, Universitätsstraße 84, Regensburg, Bavaria, 93053, Germany.
Email: irina.jarvers@ukr.de

Abstract

This study aimed to validate the short-story-task (SST) based on Dodell-Feder et al. as an instrument to quantify the ability of mentalizing and to differentiate between non-autistic adults and autistic adults, who may have acquired rules to interpret the actions of non-autistic individuals. Autistic ($N = 32$) and non-autistic ($N = 32$) adult participants were asked to read "The End of Something" by Ernest Hemingway and to answer implicit and explicit mentalizing questions, and comprehension questions. Furthermore, verbal and nonverbal IQ was measured and participants were asked how much fiction they read each month. Mentalizing performance was normally distributed for autistic and non-autistic participants with autistic participants scoring in the lower third of the distribution. ROC (receiver operator curve) analysis revealed the task to be an excellent discriminator between autistic and non-autistic participants. A linear regression analysis identified number of books read, years of education and group as significant predictors. Overall, the SST is a promising measure of mentalizing. On the one hand, it differentiates among non-autistic individuals and on the other hand it is sensitive towards performance differences in mentalizing among autistic adults. Implications for interventions are discussed.

Lay Summary

In this study, we investigated how well interpreting the actions of characters in a short story (short story task) can help to identify autistic adults, as well as subtle differences among non-autistic adults. Interpreting a character's actions in a story is more similar to social interaction in real life and may therefore be better suited to identify autistic individuals who struggle with interpreting the actions of non-autistic individuals. The short story task could differentiate between autistic and non-autistic adults with very high accuracy. Overall, the task is a promising means to aid diagnostic procedures for autistic adults and may aid them in receiving the help that is needed.

KEYWORDS
diagnostics, explicit mentalizing, implicit mentalizing, spontaneous mentalizing, theory of mind

INTRODUCTION

Inferring the beliefs, desires, and intentions of another person is one of the most crucial abilities that we possess. In order to make social interaction possible, it is important to understand that mental states of our conversation partners influence their behavior. This ability is commonly known as "Theory of Mind" (ToM) or "mentalizing." Many tasks have been developed to assess ToM in children (Gopnik & Astington, 1988; Wimmer &

Perner, 1983) and in autistic individuals who were shown to perform worse on ToM tasks than non-autistic individuals (Baron-Cohen et al., 1999; Happé, 1994; White et al., 2011). As a result, ToM has been considered a core deficit in autism (Baron-Cohen et al., 1985; Leslie, 1987), one that was believed to explain a majority of social interaction problems. However, while many studies examined whether autistic individuals struggle to infer the mental states of non-autistic individuals, few examined the opposite relationship, namely, whether non-autistic individuals are able to infer the mental states of autistic individuals (Mitchell et al., 2021). Addressing this discrepancy, Milton (2012) introduced the double empathy problem, which suggests that while autistic individuals struggle to infer the mental states of non-autistic individuals, problems in social interaction also result from being misunderstood by non-autistic individuals in return, that is, difficulties with mentalizing go both ways.

As of now, mentalizing tasks mostly served the purpose of assessing very specific inferences, for example, false beliefs, which have been considered the critical test for having a ToM (Baron-Cohen et al., 1986; Wimmer & Perner, 1983) or the decoding of emotions from facial expressions (Baron-Cohen et al., 2001, 2015; Harms et al., 2010). However, both tasks can be solved by autistic adults (Döhnel et al., 2012; Sommer et al., 2007; Spek et al., 2010) whereas gaps in interaction style between autistic and non-autistic adults have been shown to increase with age, suggesting that problems with social interaction may become worse with time (Davis & Crompton, 2021) even if performance on existing tasks improves. Thus, tasks that closely resemble real-life interaction are necessary.

The task most closely related to real life interaction is the Faux Pas task (Baron-Cohen et al., 1999; Stone et al., 1998). Here, everyday situations are described that either do or do not include a social mistake. Participants are asked whether a social mistake has occurred and to describe it. In a study by Zalla et al. (2009), autistic adults were able to identify social rule violations, but could not explain why they were violations. Adding to this, Thiébaut et al. (2016) found that autistic adults commonly overcompensated by identifying non-faux-pas situations as faux-pas. This discrepancy may indicate that task performance does not index mentalizing abilities in real life situations. In a real-life faux pas setting, individuals may not be able to identify what exactly constitutes the faux-pas. This applies to both, autistic and non-autistic individuals. Overall, even though ToM tasks were developed in order to identify mentalizing problems in autistic individuals, the detection of subtle differences between non-autistic or well compensated autistic adults (Begeer et al., 2011; Begeer et al., 2012; Happé, 1994) remains challenging. Furthermore, existing tasks rarely reflect real-life problems in social interaction with non-autistic adults, which are commonly described as burdening by autistic individuals, whereas they feel more at ease

and understood when interacting with other autistic individuals (Crompton et al., 2020).

In summary, current mentalizing tasks are (1) only able to address a subpart of mentalizing, that is, through faux pas, (2) frequently passed by autistic adults, and (3) rarely reflect autistic adults' difficulties in social interaction. Due to the nature of the available tasks, some well-compensated autistic adults who still struggle with everyday social interaction may not be diagnosed and therefore not receive the help they need, that is, interventions or trainings (Padrón et al., 2022). Already receiving a diagnosis can greatly reduce the feeling of isolation and provide a sense of belonging (Mitchell et al., 2021). Even though mentalizing is not part of official diagnostic criteria, difficulties with social interaction are and may be easily underestimated through autistic adults' compensation techniques (i.e., camouflaging; Cook et al., 2021). A new task is needed that is easy to administer and more sensitive for detecting subtle difficulties in mentalizing. A more sensitive task may be able to differentiate among non-autistic adults and help to identify autistic adults.

Dodell-Feder et al. (2013) developed the short-story task to (1) create a measure that is sensitive to individual differences in ToM ability, (2) includes mental states of varying complexity, and (3) uses stimuli that are representative of the real world. Furthermore, the social context had to be used to make an appropriate mental state inference. In the original study, participants were asked to read "The End of Something" by Ernest Hemingway (Dodell-Feder et al., 2013). This specific story was chosen because Hemingway does not describe or mention the beliefs and intentions of his protagonists but rather gives a description of the actions only, making it suitable for a ToM task. After reading the story, participants answered comprehension questions as well as explicit mentalizing questions and one implicit mentalizing question. Participants also performed other ToM measures including the interpersonal reactivity index (IRI), a measure of individual differences in empathy (Davis, 1983) and the eyes task (Baron-Cohen et al., 2001). Finally, the verbal and nonverbal intelligence quotient (IQ) was measured. Dodell-Feder et al. (2013) found an approximately normal distribution in mentalizing scores, ranging from 2 to 14 points out of 16. Comprehension was not related to the mentalizing score; however, IQ was positively correlated with better mentalizing performance. Moreover, the results of the SST correlated well with the fiction subscale of the IRI and the eyes task.

If the SST can be validated as a reliable instrument for quantifying individual differences in mentalizing, the SST may become helpful in the diagnostic process of the autism spectrum condition (ASC). The diagnostic assessment of ASC is typically a long and extensive process, involving several appointments, and various measures (Jones et al., 2014). Especially in autistic adults symptoms are more subtle and easily overlooked (Rogers et al., 2016). In addition, most adults seeking a diagnostic

assessment of ASC have a number of comorbidities (Arnold et al., 2019; Tromans et al., 2018) which complicates differential diagnostics (Wigham et al., 2019). Also, female sex has been shown to reduce the likelihood of getting diagnosed with ASC (Huang et al., 2020), probably because of autistic females' ability to camouflage their difficulties (Hull et al., 2020; Tubío-Fungueiriño et al., 2021). Overall, the median age from first clinical presentation to receiving a diagnosis of ASC has been shown to be up to 11 years, for both, males and females (Fusar-Poli et al., 2022), and autistic adults are commonly misdiagnosed (Au-Yeung et al., 2019).

Due to the complexity of the diagnostic process in ASC, functional and valid measures are crucial. For children, the Autism Diagnostic Observation Schedule – 2 (ADOS-2; Carr (2013)) and the parent-interview Autism Diagnostic Interview - Revised (ADI-R; Lord et al. (1994)) are considered to be the diagnostic gold standard. However, the ADOS-2 has been shown to be highly variable in coding among clinicians (Kamp-Becker et al., 2018) and specificity and sensitivity are often poor (Conner et al., 2019). For adults, the ADI-R is often not suitable, because parents are deceased or unable to remember the asked details. Additionally, the ADOS is not sensitive enough to detect well compensated deficits.

Nevertheless, receiving an official diagnosis of ASC can be a crucial turning point in the life of many autistic individuals (Tan, 2018), often described as emotionally relieving (Huang et al., 2020) and providing a sense of belonging (Mitchell et al., 2021). Autistic adults struggle to maintain jobs and are commonly overeducated for their positions (Frank et al., 2018) due to a negative first impression (DeBrabander et al., 2019). An official diagnosis of ASC has been shown to improve this impression, in addition to elucidation of peers and colleagues (Sasson & Morrison, 2019). Overall, receiving an accurate diagnosis of ASC can be crucial for getting adequate support and the well-being of autistic individuals, thus, emphasizing the need for more sensitive measures. As a majority of social interaction problems in ASC stem from interaction with non-autistic individuals, a mentalizing-based task that is sensitive to camouflaging may be an important addition to the diagnostic process. Thus, even though mentalizing is not part of official diagnostic criteria, a mentalizing-based task can be a great aid in identifying autistic individuals and helping them to find their community.

The first goal of the current study was to replicate the results of Dodell-Feder et al. (2013) in a German non-ASC population. The second goal was to identify whether the SST is a good measure of performance among autistic adults and whether it differentiates them from non-autistic adults. Additionally, regular frequency of fiction reading, verbal and nonverbal IQ and years in the educational system were examined as possible performance predictors. It was hypothesized that SST performance would be normally distributed for non-autistic participants, similarly to Dodell-Feder et al. (2013). For the ASC group, it was expected that they would, on average, perform worse than the non-ASC group and have a left-shifted distribution.

METHODS

Participants

In total, 32 autistic individuals (M_{age} = 30.34 years, range = 18–55) and 32 non-autistic individuals (M_{age} = 31.13 years, age range = 19–52) participated in this study. Autistic participants were recruited over clinicians and mailing lists for autistic individuals and only autistic adults with an official diagnosis of ASC were included. Non-autistic participants were recruited over flyers and mailing lists for employees of the clinic or the university. Demographic information is depicted in Table 1. All participants gave informed, written consent and the study was approved by the ethics committee of the University of Regensburg (Nr.16-101-0148). All participants received monetary compensation for participating.

Task and material

Participants' nonverbal IQ was assessed with the help of the Culture Fair Test-20 (CFT-20; Weiß, 2006) and the nonverbal IQ was assessed with the Mehrfachwahl-Wortschatzschatztest-B (MWT-B; Merz et al., 1975). The CFT-20 consists of two parts with four subtests each (continuation of series, classification, matrices and topological inference) and is conducted within a set time. The first part consists of 56 items and the second part of 45 items. The CFT-20 has demonstrated high reliability (r = 0.87), high internal consistency (Cronbach's alpha = 0.95) and high factorial validity as well as construct validity in correlation with other IQ tests (r = 0.57 – r = 0.73). The MWT-B is a short 37-item test with each item consisting of four German words. Participants have to choose which word among the four options provided is a real word. The MWT-B has demonstrated high reliability (r = 0.94) and high validity in correlation with other verbal IQ tests (r = 0.80–0.86). Additionally, participants were asked about the number of fiction books they read a month, excluding nonfiction. As answers mostly ranged between 0–3, categories were created ranging from zero books per month (0), less than one book a month (1), between 1 and 2 books a month (2) and more than 2 books a month (3).

For the short-story task (SST), the documented description of the SST by Dodell-Feder et al. (2013) was used. Participants were asked to read the German translation of "The End of Something" (translated by E., Horschitz-Horst, A., & Ceram, C. W.) and, as recommended by Dodell-Feder et al. (2013), participants were

T A B L E 1 Demographic characteristics of participants

	ASC group	Non-ASC group	Group comparisons
Sex			
Female	12 (27.50%)	15 (46.90%)	
Male	20 (62.50%)	17 (53.10%)	$t(62) = -1.25, p = 0.108$[b]
Age	30.34 (11.32)	31.13 (11.09)	$t(62) = -1.75, p = 0.042$[b]
School type[a]			
University	6 (18.80%)	9 (28.10%)	
Gymnasium	11 (34.40%)	16 (50.00%)	
Realschule	9 (28.20%)	7 (21.90%)	
Hauptschule	6 (18.80%)	0 (00.00%)	
Years in school	13.09 (2.97)	14.64 (4.14)	$t(62) = -0.28, p = 0.390$[b]
Nonverbal IQ	116.97 (15.14)	121.17 (14.64)	$t(62) = -0.87, p = 0.195$[b]
Verbal IQ	113.39 (13.62)	114.13 (14.41)	$t(62) = -1.76, p = 0.040$[b]
Number of books			$z = -1.43, p = 0.153, d = 0.16$[c]
Less than 1 a month	15 (34.40%)	21 (65.60%)	
Between 1 and 2 a month	11 (34.30%)	7 (21.90%)	
More than 3 a month	6 (18.80%)	4 (12.50%)	
SST			
Comprehension score	9.31 (1.12)	9.56 (0.76)	$z = -0.58, p = 0.561, d = 0.12$[c]
Mentalizing score	5.44 (2.17)	9.81 (2.95)	$z = -5.30, p < 0.001, d = 1.75$[c]
Spontaneous mentalizing	9 (28.00%)	9 (28.00%)	

[a]Gymnasium (higher level education, 8 to 9 years of school after 4 years of elementary school, terminating with the general university entrance qualification), realschule (intermediate secondary school, 6 years of school after 4 years of elementary school), hauptschule (9 years of elementary school).
[b]Equivalence tests were performed, significance indicates equivalence.
[c]Mann–Whitney U tests were performed to examine group differences.

asked to pay attention to the relationship between the two characters. In the story, a couple breaks up because the man is no longer interested in a relationship with the woman and attempts to leave her while they fish together. Because there are no mental state descriptions in the story, it is well suited for mentalizing questions.

After participants read the story, they were asked to summarize the plot of the story. If participants produced spontaneous mental state descriptions, they received one point, otherwise they received 0 points. Then, four comprehension questions, eight mentalizing questions, and one last comprehension question were asked. On all of these questions the participants could gain 0, 1, or 2 points, resulting in a maximum of 10 points for comprehension and 16 points for mentalizing. A rating of 0 points indicated no mental state reference, a rating of 1 indicated consideration of one perspective or partial understanding of a character's mental state, and a rating of 2 indicated consideration of several characters' perspectives and accurate mental state reasoning. Detailed rating instructions for each question can be found in the supplementary material S1 to the original SST (Dodell-Feder et al., 2013). Comprehension questions were, for example: "What does the couple see at the riverbank while rowing to their fishing spot?" Mentalizing

questions, on the other hand, were aimed at intentions and reasons for actions such as: "Why is Nick afraid of looking at Majorie?" or "What does Nick mean by, 'It's no fun anymore.'" The questions, along with detailed rating descriptions, were taken from Dodell-Feder et al. (2013) and translated into German. The English version of the questions can be found in the supplementary material S1.

Experimental procedure

After arriving at the laboratory, participants were informed about the experimental procedure and provided written, informed consent. Then, participants provided general demographic information such as age, sex, and years of education and filled out the MWT-B and the CFT-20 to assess IQ. Lastly, they read the short story by Ernest Hemingway and answered 14 questions. Their answers were audio-recorded and subsequently rated by two independent raters. The experimenter supervising the task served as first rater and an employee who was blind to the status of the participants served as second rater. Inter-rater reliability was computed using the scores given by both raters.

Statistical analysis

Statistical analysis of the data was conducted with SPSS 28 (IBM Corp, 2021). In a first step, distributions of the SST comprehension score, SST mentalizing score and IQ were inspected for outliers (±3 SD of the mean). No outliers had to be removed. Internal reliability of the SST was assessed using Cronbach's alpha. The inter-rater reliability of the comprehension and mentalizing scores was assessed using Kendall's τ correlations with the second rating of the independent judge.

To determine the diagnostic discrimination ability of the SST, area under the curve (AUC) receiver operating characteristic (ROC) analyzes were performed. ROC analysis results were interpreted with $0.50 < \text{AUC} < 0.70$ indicating poor discrimination, $0.70 < \text{AUC} < 0.80$ indicating acceptable discrimination, $0.80 < \text{AUC} < 0.90$ indicating excellent discrimination and $\text{AUC} > 0.90$ indicating superior discrimination (Shallcross & Ahner, 2020). Cut-off scores were chosen through a trade-off between sensitivity and specificity.

In a second step, equivalence tests using the 2 one-sided tests (TOST) procedure via the TOSTER package in R (Lakens, 2013) were computed to examine matching properties in age, sex, years of education, and IQ. A medium sized effect size of $d = 0.50$ and $d = -0.50$ was chosen as smaller effects were not considered to be meaningful in their impact on mentalizing performance. Furthermore, Norman et al. (2003) have argued that health outcomes commonly have a minimally important difference of $d = 0.50$. Mann–Whitney U tests were used to determine group differences in the number of books read per month and in the SST variables. Subsequently, Kendall's τ correlations between the main variables were computed in order to determine relevant control variables for a regression model. A linear regression model predicting SST mentalizing performance was computed with the predictors SST comprehension score, group (ASC, non-ASC), number of books read, spontaneous mentalizing and additional control variables chosen through correlations. Statistical significance was defined as $p < 0.05$, two-tailed for all analyzes.

RESULTS

Internal consistency and inter-rater reliability

Cronbach's alpha was used as a measure for reliability. The SST as a whole achieved an alpha of 0.70, which can be considered acceptable. As there were only five items included in the comprehension score and the intercorrelation of variances was low, internal consistency was difficult to compute. However, the mentalizing score achieved an alpha of 0.73.

SST total score ratings of the two raters correlated significantly with an τ of 0.79 ($p < 0.001$). The same applied to the SST mentalizing score ($\tau = 0.80$, $p < 0.001$). The SST comprehension score showed ceiling effects and could therefore not be reliably used for computation, however, separate questions showed correlations between $\tau = 0.75$ and $\tau = 0.90$ ($p < 0.001$). For statistical analysis, the scores of the first rater were used.

SST performance

The non-ASC group achieved a mean performance of 9.81 out of 16 points (SD = 2.95) for the mentalizing questions and 9.56 out of 10 points (SD = 0.79) for the SST comprehension questions. Overall, 9 out of 32 non-autistic participants received a point for spontaneous mentalizing. SST mentalizing was normally distributed according to the Shapiro–Wilk test ($W(32) = 0.1$, $p = 171$) with a minimal skew to the left (skew = 0.11, kurtosis = -0.66). SST comprehension scores were not normally distributed and had a strong skew to the right (skew = -1.88, kurtosis = 2.69) with most controls performing at ceiling.

The ASC group achieved a mean performance of 5.44 out of 16 points (SD = 2.17) for the SST mentalizing questions and 9.31 out of 10 points (SD = 1.12) for the SST comprehension questions. Overall, 9 out of 32 autistic participants received a point for spontaneous mentalizing. SST mentalizing was normally distributed according to the Shapiro–Wilk test ($W(32) = 0.95$; $p = 0.145$) with a slight skew to the left (skew = 0.20, kurtosis = -0.94). SST comprehension scores were not normally distributed and had a strong skew to the right (skew = -1.55, kurtosis = 1.52), showing that also the ASC group performed at ceiling. See Figure 1 for a detailed depiction of SST performance.

To assess the ability of the SST to differentiate between participants in the ASC group and participants in the non-ASC group, an ROC curve was computed for the SST mentalizing score (see Figure 2). The model achieved an $\text{AUC} = 0.88$, thereby showing excellent discrimination. With a cut-off score of 8 points on the mentalizing scale of the SST, the test would achieve a sensitivity of 93.70% to identify a participant in the ASC group and a specificity of 68.80%.

Group differences

To compare matching properties between the ASC and non-ASC group, TOST paired sample t-tests were conducted. The results rejected the hypothesis that the true effect is smaller than $d = -0.5$ or larger than $d = 0.5$ for the variables age and verbal IQ, but failed to reject the hypothesis for the variables sex, years of education and nonverbal IQ. Specifically, results suggested no meaningful differences between the ASC group and the non-ASC group in age and verbal IQ, but failed to confirm no

FIGURE 1 Raincloud plot depicting performance on the SST for the ASC group (on the right) and the non-ASC group (on the left). Performance is depicted separately for the comprehension scale and the mentalizing scale.
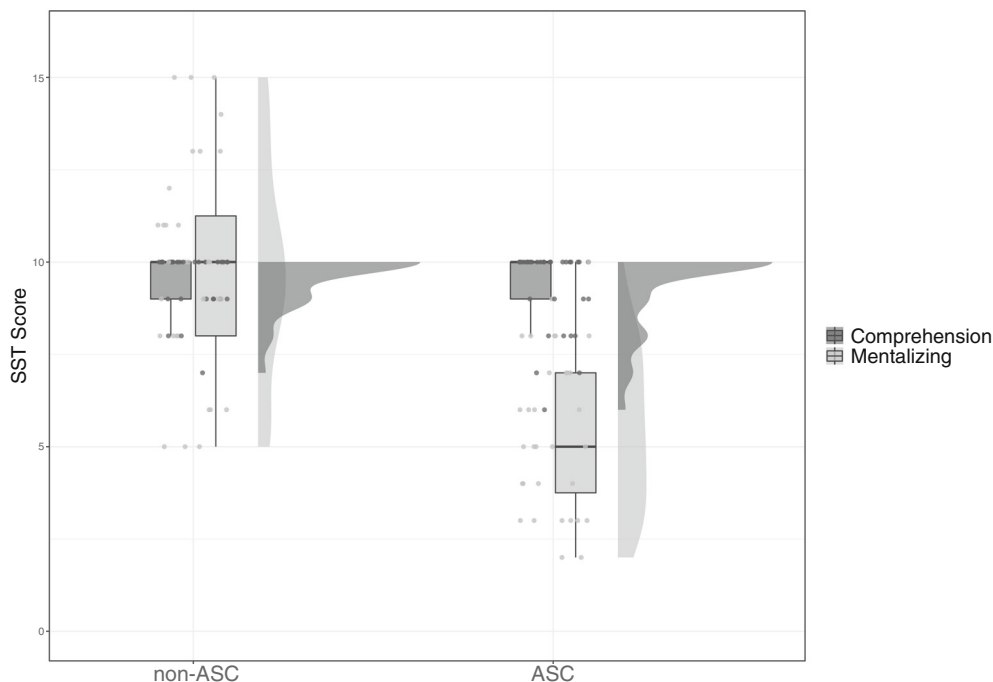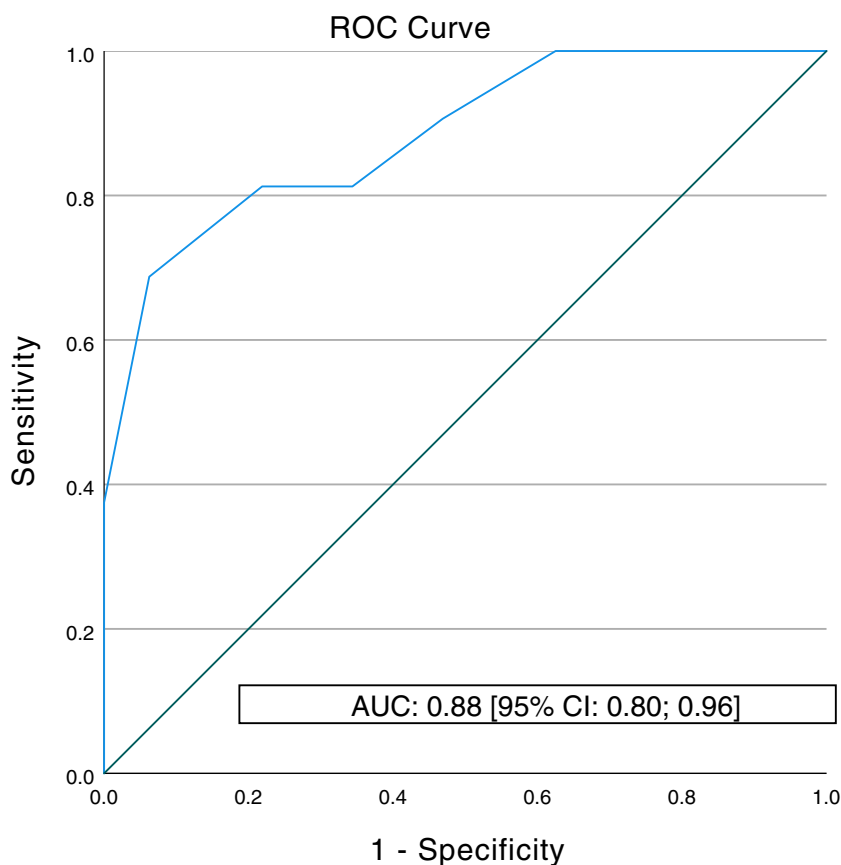

FIGURE 2 Receiver operating characteristic (ROC) curve for the identification of ASC versus non-ASC individuals. The area under the curve (AUC) result is depicted in a separate box.

meaningful differences in sex, years of education and nonverbal IQ. See Table 1 for the respective t- and p-values.

The hypothesis whether there were differences between the ASC and the non-ASC group in number of books read and the SST variables were tested via Mann–Whitney U tests, suggesting no difference in the number of books read or the SST comprehension score, but a significant difference in the SST mentalizing score (see Table 1).

Predicting SST mentalizing performance

Kendall's τ correlations between age, SST scores (comprehension, mentalizing), IQ (verbal, nonverbal), years of education, and the number of books read were computed in order to determine relevant control variables for the regression model predicting mentalizing performance. Correlations are depicted in Table 2. Overall, nonverbal IQ and years of education were added as control variables.

Subsequently, a linear regression model was computed in order to predict SST mentalizing scores with the SST comprehension score, spontaneous mentalizing, nonverbal IQ, years of education, number of books read, and group (ASC, non-ASC) as predictors. The regression model was significant ($F(5,57) = 15.22$, $p < 0.001$) and explained a total of 57.20% of the variance in SST mentalizing scores. Significant predictors were group ($t = -6.76$, $p < 0.001$), spontaneous mentalizing ($t = 2.38$, $p = 0.021$), number of books read ($t = 2.38$, $p = 0.021$), and years of education ($t = 2.01$, $p = 0.049$). A depiction of the model is shown in Table 3.

DISCUSSION

The present study aimed to (1) test whether the SST developed by Dodell-Feder et al. (2013) is feasible for identifying subtle mentalizing differences in non-autistic and autistic adults and (2) assess it as an additional measure for social interaction problems within the diagnostic assessment of ASC. For this purpose, the test was administered to 32 German non-autistic controls and 32 autistic participants. For both groups, mentalizing scores were normally distributed with non-ASC scores centered in the upper range of possible scores and ASC scores centered in the lower third. None of the participants in either group achieved a full score, suggesting the lack of a ceiling effect in the SST. Therefore, in contrast to other currently used mentalizing measures such as the faux pas task, the SST is able to identify individual differences among non-autistic and autistic adults. This benefit may be due to more complex scenarios and reality-oriented questions in the SST. As social interaction problems are bidirectional (Davis & Crompton, 2021), with autistic individuals struggling to mentalize in regard to non-autistic individuals and vice versa, a fiction-based task may reveal thinking patterns that do not become apparent in interviews or other mentalizing tasks. In tasks such as the faux pas task, real-life scenarios are heavily simplified and are therefore not a challenge for non-autistic individuals (Baron-Cohen et al., 1999; Stone et al., 1998). Also autistic adults can perform well on faux pas scenarios by applying behavioral rules to short dialog snippets (Thiébaut et al., 2016). In a more complex setting with little information about non-autistic characters' intentions and less dialogue, such behavioral rules are not as easily applicable. Through interpreting the actions of non-autistic fictional characters with few other cues, the SST could therefore reveal a mismatch in social interaction that may not become apparent otherwise.

One might argue that differences on the SST mentalizing score may be due to differences in comprehension,

TABLE 2 Overview of Kendall's τ correlations

Variable	1.	2.	3.	4.	5.	6.	7.
SST comprehension	-						
SST mentalizing	0.16	-					
IQ (nonverbal)	0.42**	0.20*	-				
IQ (verbal)	0.01	0.05	0.13	-			
Years of education	0.21*	0.29**	0.29**	0.24*	-		
Number of books read	0.10	0.09	−01	0.14	−0.01	-	
Age	−0.07	−0.04	.03	0.44***	0.18	−0.04	-

*$p < .05$.
**$p < .01$.
***$p < .001$.

TABLE 3 Overview of the linear regression model predicting SST mentalizing

Dependent variables	Predictor	B	SE	β	t	p
SST mentalizing	Group	−4.04	0.60	−0.60	−6.77	<0.001
	Spontaneous mentalizing	1.66	0.70	0.22	2.38	0.021
	Number of books read	0.67	0.28	0.21	2.38	0.021
	Nonverbal IQ	0.02	0.02	0.10	0.98	0.333
	Years of education	0.17	0.09	0.19	2.01	0.049

Note: $R^2 = 0.57$.

however, both groups showed ceiling performance and did not differ in the SST comprehension score. Furthermore, the SST comprehension score was not a significant predictor for SST mentalizing performance, suggesting that the SST mentalizing score is a genuine measure of mentalizing in regard to non-autistic fictional characters. In addition to the SST comprehension score, also the nonverbal IQ score was not significant as a predictor of SST mentalizing performance. This is not necessarily surprising, as mentalizing in regard to non-autistic individuals has been shown to be independent of IQ in ASC (Chung et al., 2014).

In addition to measuring variation in the two groups, the SST showed excellent performance as a discriminator between them. The SST mentalizing score of 8 points was chosen as a cut-off with individuals scoring below 8 points classified as with increased likelihood of ASC and individuals scoring 8 points or above classified as non-ASC. The cut-off resulted in a sensitivity of 93.70% and a specificity of 68.80% in identifying autistic adults. Both measures of the task fair better than the German version of the fourth module of the ADOS-II, which showed a sensitivity of 84% and a specificity of 28% for autistic adults while using similar sample size as in the present study (Medda et al., 2019). The task is quick to administer, extends on self-reported social interaction difficulties and has demonstrates good sensitivity and specificity in the present sample. This suggests the SST to be a very promising additional measure for autism diagnostics that should be examined further in larger samples. As the story characters portray non-autistic individuals, interpreting their actions can serve as a test for interpreting the actions of a non-autistic individual in everyday life.

The model computed to predict SST mentalizing performance explained a very large portion of the variance and identified group assignment as a crucial factor with the ASC group showing a 4.04-points decrease in the SST mentalizing score. However, group assignment was not the only factor explaining a significant portion of variance in SST mentalizing performance. Additionally, the presence of spontaneous mentalizing, the average number of books read each month and years of education were significant predictors. In the original study by Dodell-Feder et al. (2013), spontaneous mentalizing was significantly related to several other measures of mentalizing, but no difference in performance between individuals who showed spontaneous mentalizing and those who did not could be observed. In the present study, showing spontaneous mentalizing was associated with a 1.66-point increase on the SST mentalizing score when all other variables are held constant. Thus, spontaneous mentalizing plays a role independent of group assignment. Autistic individuals have commonly shown high performance on explicit mentalizing tasks, but lower performance on spontaneous mentalizing (Senju et al., 2009), however, recent neuroimaging work suggests that also in spontaneous mentalizing autistic individuals cannot be differentiated from non-autistic individuals behaviorally, but that differences in activation patterns in the right temporo-parietal junction are present (Nijhof et al., 2018). In the present study, both groups did not differ in the frequency of spontaneous mentalizing in the SST, however, its presence appears to provide a benefit for answering subsequent questions about mental states of story characters.

Another important aspect was the average number of fiction books read, suggesting that an increase in number of books read per month resulted in an 0.67-point increase on the SST mentalizing score independent of group assignment and other variables. This is in accordance with previous research identifying a relationship between empathy and reading (Bal & Veltkamp, 2013; Mar et al., 2006) and the positive effect of familiarity with reading on mentalizing (Samur et al., 2018). For example, Bal and Veltkamp (2013) found that increased identification with the emotional state of a fictional character is associated with increased empathy over time. Greater immersion in fiction and in the lives of different characters may provide the reader with additional experience of human interaction. Furthermore, developmental evidence suggests that exposing children to fiction books is a significant predictor of better ToM ability, whereas non-fiction was a negative predictor (Mar et al., 2006). However, despite the assumption that children with ASC prefer non-fiction books over fiction books, no difference in book preferences could be observed (Armstrong et al., 2019; Davidson & Ellis Weismer, 2018). As the study was correlational, the directionality of the relationship cannot be determined for sure, however, no difference between the ASC group and the non-ASC group could be identified in reading frequency, hereby suggesting that reduced mentalizing ability does not automatically reflect a dislike for fiction consumption. On the one hand, increased exposure to fiction might result in additional camouflaging for autistic individuals, thereby hiding possible difficulties in everyday social interaction through learned behaviors from books. On the other hand, increased exposure to fiction might also result in more experience with non-autistic interaction patterns and therefore more experience with non-autistic behaviors and their interpretation. Extensive contact between autistic and non-autistic individuals has been shown to reduce mismatches in interaction style (Davis & Crompton, 2021) and similar relationships may hold for reading fiction, even if autistic individuals do not directly participate in the interaction taking place. Thus, additional exposure to fiction books could provide a benefit for SST performance over time, also for individuals with ASC. This suggests that well-camouflaged autistic adults may perform above the cutoff of the SST, resulting in a failure to recognize them as autistic. However, the SST appears to fair better in identifying well-camouflaged adults than previous measures and may therefore serve as an important addition.

Finally, the number of years spent in education played a significant role for the SST mentalizing score, suggesting that each additional year spent in education resulted in a 0.17-point increase in the SST mentalizing score. Despite the effect being small, it may suggest a benefit of prolonging participants' stay in the education system. This might be related to teaching curriculums in schools, in which the analysis of fictional texts and the understanding of character's emotions and actions are an area of focus and explained in great detail. Explanations and analyzes of texts that are less implicit and more rule-oriented, may provide individuals with ASC with tools to assess the mental states of non-autistic individuals. Due to the cross-sectional nature of the study, it is not possible to determine whether prolonged education results in increased mentalizing performance or vice versa. However, if longer education indeed improves mentalizing, it would be important to introduce further support for autistic individuals in order help them achieve higher education. Moreover, more social experience may be gained by a prolonged stay with same-age individuals (Guralnick et al., 2007). Similarly to the effect observed in regard to the number of books read, prolonged education may provide additional experience and explanations for non-autistic behavior that can be used for future encounters, no matter if they are fictional or real. This in turn may lead to a difficulty in recognizing these individuals as autistic as they may score higher than the chosen cutoff of 8 points.

The present study had several strengths, among them the assessment of a fiction-based mentalizing task for autistic and non-autistic individuals, the presence of two raters for reliability measurements and the assessment of the role of fiction and education for mentalizing performance. Nevertheless, limitations are present. The study has a relatively small sample and correlations with low effect sizes might not have been identified despite their presence in the population. However, for the main question assessed via the linear regression model with five predictors, apriori power analyzes revealed the study to be sufficiently powered. Nevertheless, additional work with larger sample sizes is necessary in order to confirm the determined cut-off of 8 points and examine whether sensitivity and specificity remain high. Furthermore, in the SST, spontaneous mentalizing was queried with a single item only and coded dichotomously. Therefore, little variance was possible on the measure and qualitative differences between the two groups could not be assessed. In future studies, the SST may be paired with an additional spontaneous mentalizing measure or several open questions related to the story could be posed to develop a more versatile score. Finally, the SST only assessed mentalizing in regard to non-autistic fictional characters. Even though reported social interaction difficulties usually apply to interpreting the actions of non-autistic individuals, it is important to examine mentalizing from both perspectives. Future studies should develop a short story with autistic fictional characters and autistic communication styles, thereby measuring mentalizing in regard to autistic individuals.

Overall, this study evaluated the SST by Dodell-Feder et al. (2013) in a German population and investigated how well it can identify individual differences in mentalizing among autistic and non-autistic adults in regard to non-autistic fictional characters. Result identified the SST as a reliable task to identify differences in mentalizing abilities with high sensitivity and specificity. A model explaining variance in mentalizing performance, additionally revealed spontaneous mentalizing, number of books read a month and years of education as predictors. The impact of reading fiction and the potential impact of years of education on mentalizing needs to be investigated in more detail in future research.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT
This study was approved by the ethics committee of the University of Regensburg (Nr.16-101-0148).

ORCID
Irina Jarvers https://orcid.org/0000-0002-8414-2895

REFERENCES
Armstrong, R. M., Paynter, J., & Westerveld, M. F. (2019). Fiction or non-fiction: Parent-reported book preferences of their preschoolers with autism spectrum disorder. *Autism & Developmental Language Impairments*, *4*, 2396941519896736. https://doi.org/10.1177/2396941519896736

Arnold, S., Foley, K.-R., Hwang (Jane), Y. I., Richdale, A. L., Uljarevic, M., Lawson, L. P., Cai, R. Y., Falkmer, T., Falkmer, M., Lennox, N. G., Urbanowicz, A., & Trollor, J. (2019). Cohort profile: The Australian longitudinal study of adults with autism (ALSAA). *BMJ Open*, *9*(12), e030798. https://doi.org/10.1136/bmjopen-2019-030798

Au-Yeung, S. K., Bradley, L., Robertson, A. E., Shaw, R., Baron-Cohen, S., & Cassidy, S. (2019). Experience of mental health diagnosis and perceived misdiagnosis in autistic, possibly autistic and non-autistic adults. *Autism: The International Journal of Research and Practice*, *23*(6), 1508–1518. https://doi.org/10.1177/1362361318818167

Bal, P. M., & Veltkamp, M. (2013). How does fiction Reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS One*, *8*(1), e55341. https://doi.org/10.1371/journal.pone.0055341

Baron-Cohen, S., Bowen, D. C., Holt, R. J., Allison, C., Auyeung, B., Lombardo, M. V., … Lai, M.-C. (2015). The "Reading the mind in the eyes" test: Complete absence of typical sex difference in ~400 men and women with autism. *PLoS One*, *10*(8), e0136521. https://doi.org/10.1371/journal.pone.0136521

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

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, 4(2), 113–125. https://doi.org/10.1111/j.2044-835X.1986.tb01003.x

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, 29(5), 407–418. https://doi.org/10.1023/a:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 42(2), 241–251.

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with autism spectrum disorders. *Autism: The International Journal of Research and Practice*, 16(4), 357–366. https://doi.org/10.1177/1362361311434545

Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., & Boer, F. (2011). Theory of Mind training in children with autism: A randomized controlled trial. *Journal of Autism and Developmental Disorders*, 41(8), 997–1006. https://doi.org/10.1007/s10803-010-1121-9

Carr, T. (2013). Autism diagnostic observation schedule. In F. R. Volkmar (Ed.), *Encyclopedia of autism Spectrum disorders* (pp. 349–356). Springer. https://doi.org/10.1007/978-1-4419-1698-3_896

Chung, Y. S., Barch, D., & Strube, M. (2014). A meta-analysis of mentalizing impairments in adults with schizophrenia and autism Spectrum disorder. *Schizophrenia Bulletin*, 40(3), 602–616. https://doi.org/10.1093/schbul/sbt048

Conner, C. M., Cramer, R. D., & McGonigle, J. J. (2019). Examining the diagnostic validity of autism measures among adults in an outpatient clinic sample. *Autism in Adulthood*, 1(1), 60–68. https://doi.org/10.1089/aut.2018.0023

Cook, J., Hull, L., Crane, L., & Mandy, W. (2021). Camouflaging in autism: A systematic review. *Clinical Psychology Review*, 89, 102080. https://doi.org/10.1016/j.cpr.2021.102080

Corp, I. B. M. (2021). *IBM SPSS statistics for windows version 28.0*. IBM Corp Retrieved from https://www.ibm.com/products/spss-statistics

Crompton, C. J., Hallett, S., Ropar, D., Flynn, E., & Fletcher-Watson, S. (2020). '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(6), 1438–1448. https://doi.org/10.1177/1362361320908976

Davidson, M. M., & Ellis Weismer, S. (2018). A preliminary investigation of parent-reported fiction versus non-fiction book preferences of school-age children with autism spectrum disorder. *Autism & Developmental Language Impairments*, 3, 2396941518806109. https://doi.org/10.1177/2396941518806109

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Davis, R., & Crompton, C. J. (2021). What do new findings about social interaction in autistic adults mean for neurodevelopmental research? *Perspectives on Psychological Science*, 16(3), 649–653. https://doi.org/10.1177/1745691620958010

DeBrabander, K. M., Morrison, K. E., Jones, D. R., Faso, D. J., Chmielewski, M., & Sasson, N. J. (2019). Do first impressions of autistic adults differ between autistic and nonautistic observers? *Autism in Adulthood*, 1(4), 250–257. https://doi.org/10.1089/aut.2019.0018

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to assess mental state understanding: A new task for assessing theory of mind in adults. *PLoS One*, 8(11), e81279. https://doi.org/10.1371/journal.pone.0081279

Döhnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., & Sommer, M. (2012). Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *NeuroImage*, 60(3), 1652–1661. https://doi.org/10.1016/j.neuroimage.2012.01.073

Frank, F., Jablotschkin, M., Arthen, T., Riedel, A., Fangmeier, T., Hölzel, L. P., & Tebartz van Elst, L. (2018). Education and employment status of adults with autism spectrum disorders in Germany – A cross-sectional-survey. *BMC Psychiatry*, 18(1), 75. https://doi.org/10.1186/s12888-018-1645-7

Fusar-Poli, L., Brondino, N., Politi, P., & Aguglia, E. (2022). Missed diagnoses and misdiagnoses of adults with autism spectrum disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 272(2), 187–198. https://doi.org/10.1007/s00406-020-01189-w

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59(1), 26–37. https://doi.org/10.1111/j.1467-8624.1988.tb03192.x

Guralnick, M. J., Neville, B., Hammond, M. A., & Connor, R. T. (2007). The friendships of young children with developmental delays: A longitudinal analysis. *Journal of Applied Developmental Psychology*, 28(1), 64–79. https://doi.org/10.1016/j.appdev.2006.10.004

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of storycharacters' 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

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism Spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, 20(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Huang, Y., Arnold, S. R., Foley, K.-R., & Trollor, J. N. (2020). Diagnosis of autism in adulthood: A scoping review. *Autism*, 24(6), 1311–1327. https://doi.org/10.1177/1362361320903128

Hull, L., Petrides, K. V., & Mandy, W. (2020). The female autism phenotype and camouflaging: A narrative review. *Review Journal of Autism and Developmental Disorders*, 7(4), 306–317. https://doi.org/10.1007/s40489-020-00197-9

Jones, L., Goddard, L., Hill, E. L., Henry, L. A., & Crane, L. (2014). Experiences of receiving a diagnosis of autism Spectrum disorder: A survey of adults in the United Kingdom. *Journal of Autism and Developmental Disorders*, 44(12), 3033–3044. https://doi.org/10.1007/s10803-014-2161-3

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., Poustka, L., Weber, L., Schmidt, H., Smidt, J., Stehr, T., Roessner, V., Kucharczyk, K., Wolff, N., & Stroth, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child & Adolescent Psychiatry*, 27(9), 1193–1207. https://doi.org/10.1007/s00787-018-1143-y

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review*, 94, 412–426. https://doi.org/10.1037/0033-295X.94.4.412

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685. https://doi.org/10.1007/BF02172145

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, 40(5), 694–712. https://doi.org/10.1016/j.jrp.2005.08.002

Medda, J. E., Cholmkery, H., & Freitag, C. M. (2019). Sensitivity and specificity of the ADOS-2 algorithm in a large German sample. *Journal of Autism and Developmental Disorders*, 49(2), 750–761. https://doi.org/10.1007/s10803-018-3750-3

Merz, J., Lehrl, S., Galster, V., & Erzigkeit, H. (1975). MWT-B - ein Intelligenzkurztest. *Psychiatrie, Neurologie Und Medizinische Psychologie*, 27(7), 423–428.

Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem'. *Disability & Society*, 27(6), 883–887. https://doi.org/10.1080/09687599.2012.710008

Mitchell, P., Sheppard, E., & Cassidy, S. (2021). Autism and the double empathy problem: Implications for development and mental health. *British Journal of Developmental Psychology*, 39(1), 1–18. https://doi.org/10.1111/bjdp.12350

Nijhof, A. D., Bardi, L., Brass, M., & Wiersema, J. R. (2018). Brain activity for spontaneous and explicit mentalizing in adults with autism spectrum disorder: An fMRI study. *NeuroImage: Clinical*, 18, 475–484. https://doi.org/10.1016/j.nicl.2018.02.016

Norman, G. R., Sloan, J. A., & Wyrwich, K. W. (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care*, 41(5), 582–592.

Padrón, I., García-Marco, E., Moreno, I., Birba, A., Silvestri, V., León, I., Álvarez, C., López, J., & de Vega, M. (2022). Multisession anodal tDCS on the right Temporo-parietal junction improves mentalizing processes in adults with autistic traits. *Brain Sciences*, 12(1), 30. https://doi.org/10.3390/brainsci12010030

Rogers, C. L., Goddard, L., Hill, E. L., Henry, L. A., & Crane, L. (2016). Experiences of diagnosing autism spectrum disorder: A survey of professionals in the United Kingdom. *Autism: The International Journal of Research and Practice*, 20(7), 820–831. https://doi.org/10.1177/1362361315611109

Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, 32(1), 130–144. https://doi.org/10.1080/02699931.2017.1279591

Sasson, N. J., & Morrison, K. E. (2019). First impressions of adults with autism improve with diagnostic disclosure and increased autism knowledge of peers. *Autism*, 23(1), 50–59. https://doi.org/10.1177/1362361317729526

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science*, 325(5942), 883–885. https://doi.org/10.1126/science.1176170

Shallcross, N. J., & Ahner, D. K. (2020). Predictive models of world conflict: Accounting for regional and conflict-state differences. *The Journal of Defense Modeling and Simulation*, 17(3), 243–267. https://doi.org/10.1177/1548512919847532

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *NeuroImage*, 35(3), 1378–1384. https://doi.org/10.1016/j.neuroimage.2007.01.042

Spek, A. A., Scholte, E. M., & Van Berckelaer-Onnes, I. A. (2010). Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 40(3), 280–289. https://doi.org/10.1007/s10803-009-0860-y

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, 10(5), 640–656. https://doi.org/10.1162/089892998562942

Tan, C. D. (2018). "I'm a normal autistic person, not an abnormal neurotypical": Autism Spectrum disorder diagnosis as biographical illumination. *Social Science & Medicine*, 197, 161–167. https://doi.org/10.1016/j.socscimed.2017.12.008

Thiébaut, F. I., White, S. J., Walsh, A., Klargaard, S. K., Wu, H.-C., Rees, G., & Burgess, P. W. (2016). Does faux pas detection in adult autism reflect differences in social cognition or decision-making abilities? *Journal of Autism and Developmental Disorders*, 46, 103–112. https://doi.org/10.1007/s10803-015-2551-1

Tromans, S., Chester, V., Kiani, R., Alexander, R., & Brugha, T. (2018). The prevalence of autism Spectrum disorders in adult psychiatric inpatients: A systematic review. *Clinical Practice and Epidemiology in Mental Health: CP & EMH*, 14, 177–187. https://doi.org/10.2174/1745017901814010177

Tubío-Fungueiriño, M., Cruz, S., Sampaio, A., Carracedo, A., & Fernández-Prieto, M. (2021). Social camouflaging in females with autism Spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders*, 51(7), 2190–2199. https://doi.org/10.1007/s10803-020-04695-x

Weiß, R. H. (2006). CFT 20-R: Grundintelligenztest Skala 2-Revision. Hogrefe.

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with autism. *Autism Research: Official Journal of the International Society for Autism Research*, 4(2), 149–154. https://doi.org/10.1002/aur.174

Wigham, S., Rodgers, J., Berney, T., Le Couteur, A., Ingham, B., & Parr, J. R. (2019). Psychometric properties of questionnaires and diagnostic measures for autism spectrum disorders in adults: A systematic review. *Autism*, 23(2), 287–305. https://doi.org/10.1177/1362361317748245

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Zalla, T., Sav, A.-M., Stopin, A., Ahade, S., & Leboyer, M. (2009). Faux pas detection and intentional action in Asperger syndrome. A replication on a French sample. *Journal of Autism and Developmental Disorders*, 39(2), 373–382. https://doi.org/10.1007/s10803-008-0634-y

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.