


## SHORT REPORT

# Specificity of the short-story task for autism diagnosis when controlling for depression

Irina Jarvers<sup>1</sup>  | Johannes Pfisterer<sup>2</sup> | Katrin Döhnel<sup>2</sup> | Lore Blaas<sup>2</sup> |  
Manuela Ullmann<sup>2</sup> | Berthold Langguth<sup>2</sup> | Rainer Rupprecht<sup>2</sup> | Monika Sommer<sup>2,3</sup>

<sup>1</sup>Department of Child and Adolescent Psychiatry and Psychotherapy, University of Regensburg, Regensburg, Germany

<sup>2</sup>Department of Psychiatry and Psychotherapy, University of Regensburg at the Bezirksklinikum Regensburg, Regensburg, Germany

<sup>3</sup>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](mailto:irina.jarvers@ukr.de)

## Abstract

Securing an accurate autism-spectrum-condition diagnosis, particularly among women, remains challenging for autistic adults. Building upon previous research highlighting the short-story task (SST) as a promising tool for detecting fiction-based mentalizing difficulties in autistic adults, this study expands its scope. We investigated the SST's discriminative capacity across three distinct groups: autistic individuals ( $n = 32$ ), nonautistic individuals without mental health problems ( $n = 32$ ), and nonautistic individuals with clinical depression ( $n = 30$ ). All three groups differed significantly from each other in their SST mentalizing score with the nonautistic group having the highest scores, the nonautistic but depressed group having medium scores and the autistic group showing the lowest scores. Receiver operator curve (ROC) analysis reaffirmed the SST's efficacy as a discriminator. Moreover, a linear regression analysis identified the SST mentalizing score, the SST comprehension score, and the number of books read per month as significant predictors of autism-spectrum-condition diagnosis. These findings bolster the SST's potential as a valuable adjunct in autism diagnostics, highlighting its discriminatory ability across diverse samples.

## Lay Summary

Finding out if someone is autistic, especially if they are a woman, can be really difficult. A new test called the short-story task seems to be a new promising diagnostic tool. The short-story task looks at how well people understand the thoughts and feelings of nonautistic people within a story, and it seems to be good at telling the difference between autistic and nonautistic people. In our study, we looked at three groups of people: those who are autistic, those who are not autistic and without mental health problems, and those who are not autistic but suffer from depression. We found that the short-story task was good at figuring out who was autistic, who was not and who had depression. We also found that how well someone did on the test, and how many books they read each month were all linked to whether they were autistic. This means that the short-story task could be a useful tool for autism diagnostics.

## KEYWORDS

fiction-based mentalizing, depression, diagnostics, mentalizing, short-story task

## INTRODUCTION

While obtaining an official diagnosis of autism spectrum condition (ASC) often brings a sense of relief (Mitchell

et al., 2021), the diagnostic process is associated with various challenges for autistic adults (de Broize et al., 2022). Notably, autistic women, who frequently employ camouflaging techniques and have developed coping

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Autism Research* published by International Society for Autism Research and Wiley Periodicals LLC.

strategies into adulthood (Hull et al., 2020), are particularly prone to not receiving a formal ASC diagnosis (Huang et al., 2020). The current diagnostic procedures have limitations in capturing the real-life difficulties autistic individuals face when interacting with nonautistic peers and understanding their thoughts and beliefs (Jarvers et al., 2023). Dodell-Feder et al. (2013) introduced a promising tool, the short-story task (SST), which demonstrated effective discrimination between autistic and nonautistic adults without mental health problems (Jarvers et al., 2023). However, challenges in mentalizing abilities extend beyond autism; meta-analytical evidence from Bora and Berk (2016) indicates that individuals with depression also exhibit reduced performance on mentalizing tasks, correlated with the severity of depressive symptoms. Furthermore, depression is among the most reported comorbidities in autistic adults (Albantakis et al., 2018; van Heijst et al., 2020). To validate the SST as a reliable tool for ASC diagnosis, it must specifically target mentalizing differences associated with ASC and not depression. In this study, we aim to further assess the specificity of the SST by extending a previously published dataset of autistic and nonautistic adults (Jarvers et al., 2023) to include a sample of individuals diagnosed with clinical depression, a group known for decreased mentalizing abilities. We hypothesized that individuals with depression will achieve lower scores than the nonautistic group without mental health problems but better than the autistic group on the SST. Additionally, we expect mentalizing performance on the SST to significantly predict ASC diagnosis.

## METHODS

### Participants

In total, 32 autistic individuals (ASC group,  $M_{\text{age}} = 30.34$  years,  $SD_{\text{age}} = 11.32$  years, range = 18–55, 37.50% female), 32 nonautistic individuals without mental health problems (CON group,  $M_{\text{age}} = 31.13$  years,  $SD_{\text{age}} = 11.09$  years, age range = 19–52, 46.90% female), and 30 individuals with depression (DEP group,  $M_{\text{age}} = 37.40$  years,  $SD_{\text{age}} = 13.72$  years, age range = 19–62, 56.70% female) participated in this study. Data for the autistic and control sample were taken from Jarvers et al. (2023) and extended by the sample of individuals with depression.

### Task and material

All incorporated measures are commonly used and show sufficient reliability and validity. Participants' nonverbal intelligence quotient (IQ) were evaluated using the Culture Fair Test-20 (CFT-20; Weiß, 2006), while their

verbal IQ was assessed with the Mehrfachwahl-Wortschatztest-B (MWT-B; Merz et al., 1975).

Participants who suffered from depression additionally filled out the Hamilton Depression Scale (HAM-D; Hamilton, 1960) and the Beck-Depression-Inventory II (BDI-II, Beck et al., 1961) as measures of depressive symptomatology.

Furthermore, participants were queried about their monthly consumption of fiction books, excluding nonfiction. Responses predominantly fell within the 0–3 range. To facilitate analysis, categories were established, ranging from less than one book per month (1), between 1 and 2 books per month (2), and more than 2 books per month (3).

For the SST, we utilized the protocol outlined by Dodell-Feder et al. (2013). Participants were instructed to read the German translation of “The End of Something” (translated by E., Horschitz-Horst, A., & Ceram, C. W.) and specifically directed to focus on the dynamic between the story's two characters. The narrative depicts a couple undergoing a breakup, initiated by the man's disinterest in continuing their relationship. The absence of explicit mental state descriptions within the narrative renders it conducive to eliciting mentalizing responses from participants.

Following the reading of the story, participants were asked to summarize its plot. If participants spontaneously incorporated descriptions of characters' mental states into their summaries, they received one point, otherwise zero points. Subsequently, participants had to answer four comprehension questions, eight mentalizing questions, and a final comprehension question. Each question offered the opportunity to earn 0, 1, or 2 points, allowing for a maximum of 10 points for comprehension and 16 points for mentalizing. Detailed rating instructions for each question were adapted from the supplementary materials accompanying the original SST (Dodell-Feder et al., 2013).

### Experimental procedure

Upon arrival at the laboratory, participants were briefed on the experimental protocol and requested to provide written, informed consent. Following this, participants completed a demographic questionnaire, providing details such as age, sex, and years of education. Subsequently, participants underwent assessments of their verbal and nonverbal IQ using the MWT-B and CFT-20, respectively. Individuals with depression additionally filled out the BDI-II and HAM-D.

Following the completion of these assessments, participants were presented with the short story authored by Ernest Hemingway and instructed to read it attentively. Subsequently, they responded to a series of 14 questions pertaining to the narrative. Their responses were audio-recorded for subsequent analysis and rating.

## Statistical analysis

The data were analyzed using SPSS 28 (IBM Corp, 2023). Initially, Kruskal–Wallis tests were employed to compare basic demographics, mentalizing, and comprehension performance across the three participant groups. Subsequently, area under the curve (AUC) receiver operating characteristic (ROC) analyses were conducted to assess the discriminatory ability of the SST. AUC values falling between 0.50 and 0.70 were indicative of poor discrimination, 0.70–0.80 of acceptable discrimination, 0.80–0.90 of excellent discrimination, and AUC values exceeding 0.90 implied superior discrimination (Shallcross & Ahner, 2020). As a second and final step, two regression models were computed. The first model involved a linear multivariable regression predicting mentalizing performance, with group membership (autistic or depression), spontaneous mentalizing in the SST, number of books read, nonverbal IQ, and years of education serving as predictors. Additionally, SST comprehension scores were

incorporated into this model due to the pronounced group differences observed compared with the original study. The second regression model was a binomial logistic regression predicting ASC diagnosis with SST mentalizing and comprehension scores, number of books read, nonverbal IQ, and years of education as predictors. Statistical significance was set at  $\alpha = 0.05$ .

## RESULTS

Overall, the SST showed an internal consistency of  $\alpha = 0.73$ , with  $\alpha = 0.77$  for the mentalizing score. Demographic characteristics of the sample can be found in Table 1. The DEP group had an average score of 26.67 (SD = 10.29) on the BDI-II and 19.83 (SD = 5.24) on the HAM-D. There was no significant correlation between the BDI-II score for the nonautistic group with depression and the comprehension ( $r = 0.03$ ,  $p = 0.813$ ) or mentalizing score of the SST ( $r = 0.15$ ,  $p = 0.213$ ).

**TABLE 1** Demographic characteristics of participants and linear regression predicting mentalizing performance.

Variable	ASC group	CON group	DEP group
Years in school	13.09 (2.97)	14.64 (4.14)	14.35 (3.14)
Nonverbal IQ	116.97 (15.14)	121.17 (14.64)	115.63 (15.72)
Verbal IQ	113.39 (13.62)	114.13 (14.41)	108.13 (14.65)
Number of Books			
Less than 1 a month	15 (34.40%)	21 (65.60%)	20 (66.70%)
Between 1 and 2 a month	11 (34.30%)	7 (21.90%)	7 (23.30%)
More than 3 a month	6 (18.80%)	4 (12.50%)	3 (10.00%)
SST			
Comprehension score	9.31 (1.12)	9.56 (0.76)	8.63 (1.43)
Mentalizing score	5.44 (2.17)	9.81 (2.95)	7.77 (3.54)
Spontaneous mentalizing	9 (28.00%)	9 (28.00%)	8 (26.70%)

Dependent variable	Predictor	B	SE	$\beta$	t	p
SST Mentalizing	DEP	−1.33	0.70	−0.18	−1.90	0.060
	ASC	−4.00	0.67	−0.56	−5.98	<0.001
	Spontaneous mentalizing	1.21	0.63	0.16	1.92	0.058
	Number of books read	0.71	0.27	0.22	2.62	0.010
	Nonverbal IQ	0.01	0.02	0.05	0.55	0.581
	Years of education	0.15	0.09	0.15	1.78	0.079

Dependent Variable	Predictor	B	SE	Exp(B)	Wald	p
ASC diagnosis	SST comprehension	0.64	0.29	1.89	4.67	0.031
	SST mentalizing	−0.56	0.14	0.57	17.43	<0.001
	Number of books read	0.58	0.27	1.79	4.60	0.032
	Nonverbal IQ	0.01	0.02	1.01	0.13	0.716
	Years of education	−0.08	0.10	0.92	0.67	0.413

*Note:* Gymnasium (higher level education, 8–9 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). DEP = group of individuals with depression, CON = group of nonautistic individuals without mental health problems, ASC = group of autistic individuals, DEP = group of non-autistic individuals with depression.

Abbreviation: SST, short-story task.

## Group differences

The three groups did not differ in baseline characteristics such as age ( $z = 3.96$ ,  $p = 0.138$ ), years of education ( $z = 4.31$ ,  $p = 0.116$ ), verbal IQ ( $z = 4.04$ ,  $p = 0.133$ ), nonverbal IQ ( $z = 2.22$ ,  $p = 0.329$ ), or number of books read ( $z = 3.23$ ,  $p = 0.199$ ). There was also no difference between males and females regarding comprehension ( $z = -1.91$ ,  $p = 0.056$ ) and mentalizing ( $z = -0.74$ ,  $p = 0.458$ ). There was a significant difference in both comprehension ( $z = 9.50$ ,  $p = 0.009$ ) and mentalizing ( $z = 27.10$ ,  $p < 0.001$ ) across the three groups. Post hoc tests revealed significantly lower comprehension scores in the DEP group compared with both the ASC group ( $z = 2.33$ ,  $p = 0.020$ ) and the CON group ( $z = 2.93$ ,  $p = 0.003$ ), but no significant difference in comprehension between the ASC and CON group ( $z = 0.61$ ,  $p = 0.544$ ). Regarding mentalizing performance the ASC group had significantly lower scores compared with both the DEP group ( $z = -2.72$ ,  $p = 0.006$ ) and the CON group ( $z = 5.20$ ,  $p < 0.001$ ). Furthermore, the DEP group had significantly lower scores than the CON group ( $z = 2.40$ ,  $p = 0.017$ ). See Figure 1 for a graphical depiction of performance.

## Receiver operator characteristic analyses

Two ROC curves were computed to assess the SST's ability to differentiate between autistic and nonautistic individuals. The first ROC curve compared autistic individuals with all other individuals, while the second compared autistic individuals specifically with nonautistic individuals with depression. The first model demonstrated an AUC of 0.80 (95% CI [0.70, 0.88],  $p < 0.001$ ), indicating excellent discriminatory capability. Utilizing the previously defined cut-off score of 8 points on the

mentalizing scale of the SST, the test exhibited a sensitivity of 93.70% in identifying participants belonging to the ASC group, alongside a specificity of 54.80%. The second model demonstrated an AUC of 0.70 (95% CI [0.56, 0.83],  $p = 0.005$ ), indicating acceptable discrimination. The previously defined cut-off score of 8 points exhibited a sensitivity of 94.40% and a specificity of 60.00%.

## Predicting SST performance

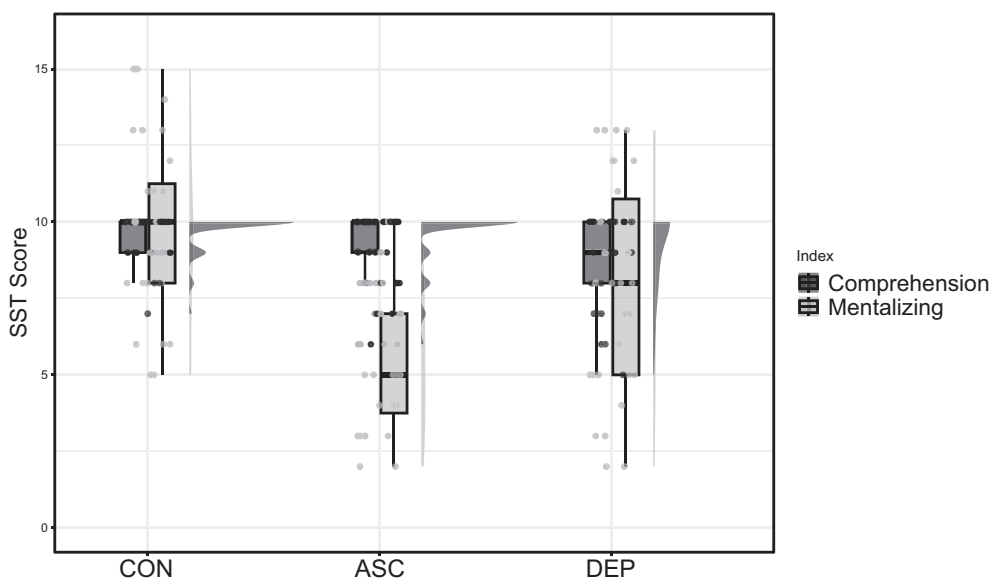
A linear multivariable regression was computed to identify predictors of mentalizing performance on the SST. The model was significant ( $F(85,7) = 10.55$ ,  $p < 0.001$ ) and explained 46.50% of variance in the mentalizing score. Significant predictors were an ASC diagnosis and a larger number of books read per month (see Table 1).

## Predicting ASC diagnosis

A binominal regression was computed to identify predictors contributing to a diagnosis of ASC. The regression was significant ( $\chi^2(5) = 35.80$ ,  $p < 0.001$ , Nagelkerke  $R^2 = 0.44$ ), correctly predicting 82.80% of the cases. The SST mentalizing score, the SST comprehension score, and the number of books read were significant predictors (see Table 1).

## DISCUSSION

The primary objective of the present study was to evaluate the utility of the SST for ASC diagnostics by extending the sample in Jarvers et al. (2023) and comparing performance across three distinct groups: autistic individuals, nonautistic individuals, and nonautistic individuals



**FIGURE 1** Raincloud plot depicting performance on the short-story task (SST) for the CON group (left), the ASC group (middle), and the DEP group (right). Performance is depicted separately for the comprehension scale and the mentalizing scale.

with depression. Given the frequent co-occurrence of autism and depression (van Heijst et al., 2020) and the documented association between depression and diminished mentalizing abilities (Bora & Berk, 2016), individuals diagnosed with clinical depression were included to assess the SST's specificity.

As anticipated, individuals with depression exhibited lower scores on the SST's mentalizing measure compared with the CON group but outperformed the ASC group. This intermediate performance among individuals with depression could be attributed to various factors. Notably, mentalizing difficulties in depression are theorized to stem, in part, from deficits in humor processing and executive functioning (Uekermann et al., 2008), while autistic individuals primarily encounter challenges in comprehending the thoughts and behaviors of nonautistic individuals (and vice versa) (Milton, 2012). Additionally, it is noteworthy that individuals in the DEP group scored significantly lower on comprehension tasks compared with the other two groups. This discrepancy underscores potential executive functioning impairments, which in turn may impact the understanding of the protagonists' implicit mental states and hinder the interpretation of the short story. However, depressive symptomatology was not significantly correlated with mentalizing performance in our present sample with clinical depression, suggesting that the observed reduction in mentalizing performance may not be a direct consequence of depressive symptoms.

ROC analyses demonstrated that with the inclusion of a sample of nonautistic individuals with depression, the SST maintained good discriminatory ability between autistic and nonautistic individuals. However, when only comparing the autistic group with the nonautistic group with depression discriminatory ability reduced to acceptable. While there was no significant correlation between depressive symptomatology and mentalizing performance in the nonautistic group with depression, executive functioning impairments and secondary aspects of clinical depression may contribute to reduced mentalizing performance and thus diminish the discriminatory ability of the SST. Implementing an additional depression screening may enhance the interpretability of the SST in the future.

Further analyses aimed at predicting mentalizing performance and ASC group assignment revealed consistent findings. Specifically, SST mentalizing performance was significantly influenced by ASC group membership but not by DEP group membership. Similarly, variance in ASC group membership was notably explained by SST mentalizing performance, underscoring the task's efficacy in assessing mentalizing abilities across nonautistic individuals.

Additionally, consistent with previous research (Samur et al., 2018), the frequency of monthly book reading emerged as a significant predictor for both mentalizing performance and ASC group assignment. Although directionality cannot be implied from the present results,

a recent investigation by Chapple et al. (2021) highlighted that both autistic and nonautistic individuals reported positive social learning outcomes resulting from reading habits. While this finding may introduce complexity to ASC diagnostics for individuals with extensive reading experience, it also presents an opportunity for leveraging reading habits as a potential tool in the future.

Our study possesses several notable strengths, including the inclusion of a well-characterized sample comprising nonautistic individuals with depression. Additionally, the utilization of a previously collected dataset comprising autistic and nonautistic individuals enhances the robustness of the assessment of the SST. These strengths provide valuable insights into the SST's applicability across diverse populations.

However, certain limitations warrant consideration. The relatively small sample size and the rigid classification of participants into distinct groups represent noteworthy limitations. Furthermore, since the sample with depression was collected after the autistic and nonautistic sample, depressive symptomatology was not assessed for all participants, preventing an examination of the effect of depression in the autistic sample. While the autistic sample was recruited outside of out- and in-patient settings, a certain level of depressive symptomatology may still be present. These constraints may impact the generalizability of the findings and potentially limit the scope of conclusions drawn from the study.

Future research should consider larger and more diverse samples of individuals undergoing ASC diagnostics to comprehensively evaluate the effectiveness of the SST, especially in complex diagnostic cases. Expanding the scope of investigation will enhance our understanding of the SST's applicability and its potential utility in clinical settings.

In summary, the SST demonstrates acceptable discriminative ability in distinguishing between autistic and nonautistic individuals, even when incorporating a sample of nonautistic individuals with depression, who exhibit diminished mentalizing performance. As a fiction-based mentalizing assessment tool that closely resembles real-life mentalizing scenarios, the SST holds promise as a valuable addition to ASC diagnostics. Its ability to capture nuanced differences in mentalizing abilities across diverse populations underscores its potential use in clinical practice. By offering a more ecologically valid assessment of mentalizing skills, the SST may enhance the accuracy and comprehensiveness of ASC diagnostic evaluations.

## ACKNOWLEDGMENTS

We thank all participants for taking the time to participate in our study. Open access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Open Science Framework at <https://>



osf.io/jeadx/?view\_only=fb60af85168a418684ded98714efbc6b, reference number (will be added after peer review).

## 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

- Albantakis, L., Parpart, H., Thaler, H., Krankenhagen, M., Böhm, J., Zillekens, I., & Schilbach, L. (2018). Depression bei Erwachsenen mit Autismus-Spektrum-Störung. *Nervenheilkunde*, 37(9), 587–593. <https://doi.org/10.1055/s-0038-1670568>
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561–571. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- Bora, E., & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, 191, 49–55. <https://doi.org/10.1016/j.jad.2015.11.023>
- Chapple, M., Williams, S., Billington, J., Davis, P., & Corcoran, R. (2021). An analysis of the reading habits of autistic adults compared to neurotypical adults and implications for future interventions. *Research in Developmental Disabilities*, 115, 104003. <https://doi.org/10.1016/j.ridd.2021.104003>
- de Broize, M., Evans, K., Whitehouse, A. J. O., Wray, J., Eapen, V., & Urbanowicz, A. (2022). Exploring the experience of seeking an autism diagnosis as an adult. *Autism in Adulthood*, 4(2), 130–140. <https://doi.org/10.1089/aut.2021.0028>
- 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>
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 23(1), 56–62. <https://doi.org/10.1136/jnnp.23.1.56>
- 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>
- IBM Corp. (2023). *IBM SPSS Statistics for Windows, Version 29.0.2.0*. <https://www.ibm.com/products/spss-statistics>

- Jarvers, I., Döhnell, K., Blaas, L., Ullmann, M., Langguth, B., Rupprecht, R., & Sommer, M. (2023). “Why do they do it?”: The short-story task for measuring fiction-based mentalizing in autistic and non-autistic individuals. *Autism Research*, 16(3), 558–568. <https://doi.org/10.1002/aur.2871>
- Merz, J., Lehl, 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>
- 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>
- 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>
- Uekermann, J., Channon, S., Lehmkämpfer, C., Abdel-Hamid, M., Vollmoeller, W., & Daum, I. (2008). Executive function, mentalizing and humor in major depression. *Journal of the International Neuropsychological Society*, 14(1), 55–62. <https://doi.org/10.1017/S1355617708080016>
- van Heijst, B. F., Deserno, M. K., Rhebergen, D., & Geurts, H. M. (2020). Autism and depression are connected: A report of two complimentary network studies. *Autism: The International Journal of Research and Practice*, 24(3), 680–692. <https://doi.org/10.1177/1362361319872373>
- Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala 2-Revision*. Hogrefe.

**How to cite this article:** Jarvers, I., Pfisterer, J., Döhnell, K., Blaas, L., Ullmann, M., Langguth, B., Rupprecht, R., & Sommer, M. (2024). Specificity of the short-story task for autism diagnosis when controlling for depression. *Autism Research*, 1–6. <https://doi.org/10.1002/aur.3191>