

Evaluation of the theory of mind in autism spectrum disorders with the Strange Stories test

Avaliação de teoria da mente nos transtornos do espectro do autismo com a aplicação do teste *Strange Stories*

Renata de Lima Velloso¹, Cintia Perez Duarte², José Salomão Schwartzman³

ABSTRACT

Objective: To evaluate the theory of mind in autism spectrum disorders (ASD) and control individuals by applying the Strange Stories test that was translated and adapted to the Portuguese language. **Method:** Twenty-eight children with ASD and 56 controls who were all male and aged between 6 and 12 years participated in the study. **Results:** There were significant differences between the median scores of the groups for each of the 12 stories of the test and for the sum total of all the median scores. The median scores for all stories were significantly greater in the control group than those in the experimental group (children with ASD). In addition, the protocol had excellent internal consistency. **Conclusion:** The theory of mind skills assessed with the Strange Stories test indicated alterations in children with ASD compared with children in the control group.

Keywords: child developmental disorders, pervasive, autism, theory of mind.

RESUMO

Objetivo: Avaliação de habilidades de Teoria da Mente em indivíduos com Transtornos do Espectro do Autismo (TEA) e indivíduos-controle, com a aplicação do teste *Strange Stories*, traduzido e adaptado para a Língua Portuguesa. **Método:** Participaram do estudo 28 crianças com TEA e 56 crianças-controle, todas do sexo masculino e na faixa etária entre seis e 12 anos. **Resultados:** Foram observadas diferenças significativas entre os escores médios dos grupos em cada uma das 12 histórias do teste e na soma dos escores de todas as histórias. Os escores médios registrados para todas as histórias foram significativamente maiores no grupo-controle do que no grupo experimental (crianças com TEA). Observou-se ótima consistência interna do protocolo. **Conclusão:** As habilidades de Teoria da Mente avaliadas pelo teste *Strange Stories* se mostraram alteradas no grupo de crianças com TEA quando comparadas às crianças do grupo-controle.

Palavras-chave: transtornos globais do desenvolvimento, autismo, teoria da mente.

In an attempt to identify and characterize the most common symptoms that make up autism, Wing and Gould¹ have described the following triad of very specific behaviors and characteristics of the disorder: severely handicapped social interactions, severe difficulties in verbal and nonverbal communications, and absence of imaginative activities, with the presence of repeated and stereotyped behaviors. Currently, the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)² and the International Classification of Diseases

(ICD-10)³ use the term pervasive developmental disorders (PDD) to characterize these conditions. Rutter⁴ has emphasized that the increase in autism diagnoses in the last 40 years demands the introduction of the idea of autism spectrum disorders (ASD). In this study, the term ASD will be used to refer to subjects with the diagnoses of autism, Asperger's syndrome, and PDD not otherwise specified (PDDNOS) to take into account the similar manifestations that are encountered in subjects with these diagnoses.

¹Fonoaudióloga, Mestre e Doutora em Distúrbios do Desenvolvimento, Universidade Presbiteriana Mackenzie;

²Psicóloga e Mestre em Distúrbios do Desenvolvimento, Universidade Presbiteriana Mackenzie;

³Médico Neuropediatra, Doutorado em Neurologia pela Universidade Federal de São Paulo, Professor titular do Programa de Pós-Graduação em Distúrbios do Desenvolvimento, Universidade Presbiteriana Mackenzie.

Instituição: Clínica de Transtornos do Espectro do Autismo – TEA MACK, Programa de Pós-Graduação em Distúrbios do Desenvolvimento, Centro de Ciências Biológicas e da Saúde - Universidade Presbiteriana Mackenzie.

Correspondence: Renata de Lima Velloso; Rua Piaui 181 / Higienópolis; 01241-001 São Paulo SP - Brasil; E-mail: relimaveloso@yahoo.com.br

Conflict of interest: There is no conflict of interest to declare.

Received 19 March 2013; Received in final form 06 June 2013; Accepted 13 June 2013.

Individuals with ASD face many difficulties, and throughout the years, these have been attributed to social, affective, and cognitive deficits. The two main theories that are currently discussed in relation to the primary prejudices of subjects with ASD are developmental and cognitive theories. Among the cognitive theories, the theory of mind is often used to explain the cognitive characteristics of individuals with ASD. Uta Frith⁵ and Simon Baron-Cohen⁶ have suggested that cognitive deficits are the primary deficits in the development of individuals with ASD. The authors have described difficulties or inabilities of these individuals to comprehend the point of view of others, to have a mental image, or to form a self-representation of reality. This model is called the theory of mind^{5, 6}. Baron-Cohen⁷ has used the term “mindblindness” to describe the deficits of individuals with ASD according to the theory of mind. Due to their intelligence and preserved speech, individuals with Asperger’s syndrome can make use of strategies to compensate for their difficulties according to the theory of mind, and these strategies should not be confused with adequate social competence⁷.

It was with this in mind that Happé^{8,9} created the advanced test of the theory of mind Strange Stories (SS), in which little skits or stories about everyday situations involving people saying things that they do not mean literally, are presented to the autistic child. Other studies have replicated the use of this test¹⁰⁻¹³, and they have found that, relative to the control group, individuals with ASD do not respond adequately to the stories, even when they exhibit positive results in other theory of mind tests.

This study was developed with the objectives of evaluating the theory of mind test in Brazilian children with ASD and comparing children with ASD with controls (children without the disorder). The age of the children ranged from 6 to 12 years. The advanced SS test, which was translated and adapted to the Portuguese language, was used in the study.

METHOD

This study was evaluated by the Ethics Committee of the Mackenzie Presbyterian University who approved all the procedures that were proposed for the research (process CEP/UPM n. 1295/10/2010 and CAAE n. 0103.0.272.000-10).

Twenty-eight children who were all male and who had ASD (experimental group) and 56 children of the same gender without the disorder (control group) who were aged between 6 and 12 years participated in this study. The following criteria were used for inclusion in the experimental group.

1. Clinical diagnosis of PDD according to the criteria proposed by DSM-IV² and ICD-10³.
2. Absence of a low intelligence quotient (IQ). To determine this, children were subjected to a neuropsychological

evaluation, and the Wechsler intelligence scale for children, 3rd edition (WISC-III)¹⁴ was applied. Only the vocabulary (verbal scale) and cube (execution scale) subtests were used to reduce the WISC-III test. Sattler¹⁵ has published a table that converts the estimated IQs to the corresponding global IQs using these two subtests, and this table was used in this study. To determine the absence of intellectual deficiencies, an IQ of 70 was used as the cutoff point for inclusion.

3. Speaks a response, even if only single words.

Children were excluded from the study if they had listening and/or visual deficiencies or any other development disorders as well as possible neurological or psychiatric comorbidities.

All children were selected and included in this study after their parents or proxy gave their consent to be included.

Children with ASD were selected randomly after we divulged the search for subjects for the experimental part of the study. This investigation and other data collection were conducted at the Autism Spectrum Disorders Clinic, which is linked to the Postgraduate Program in Developmental Disorders at Mackenzie Presbyterian University. In the clinic, children were subjected to neuropsychological and speech-therapy evaluations as well as physical and neurological exams. The subjects were identified and removed from the routine clinical standardized protocol. The data were collected in a long interview with the parents and/or proxy of the individual being evaluated.

The control group consisted of 56 male children who were selected according to the same inclusion criteria, except those diagnosed with ASD and were assigned to the experimental group. Children who complained of learning difficulties or listening and/or visual deficiencies or those that had any other developmental disorders and possible neurological and psychiatric disorders were excluded.

Children without developmental disorders were selected from the general population at a social center that was located in the city of São Paulo. All male children from 6 to 12 years of age who matriculated at this center were part of the control group. A letter of intent and terms of consent were given to the parents and/or proxy of these 56 children in the control group and then submitted to the institution after their approval.

In the first phase of the study, the advanced test of the theory of mind SS⁹ was translated to the Portuguese language. First, two translations were done by independent and Brazilian translators. Second, a synthesis of the two translations was conducted. Third, the two translators proceeded to back-translate the final version. In the fourth and last phase, the final translation and back-translation were undertaken by a committee of 6 bilingual language specialists. The rate of concurrence was calculated at 100%. After all the judges’ suggestions and after a consensus was reached with regard

to the semantic, idiomatic, experimental, and conceptual equivalences, the final version that was used in this study was created.

In its original version, the advanced theory of mind SS is composed of 24 stories that contain different attributes. Latter studies have proposed a smaller number of stories and more variations to these stories¹⁰⁻¹³. For this study, the same 12 stories were used that were used and presented in a recently published study¹⁶ that sought normative data of 140 typical children and that were extracted from the 24 stories that were originally published⁹. Each one of these stories represents an attribute of a mental state. To apply the test, each story is read to a child and then the child is asked two questions to verify the child's interpretation of the mental state of the character. With the exception of the stories about misunderstanding and double-bluff, the questions were always the same. As previously described by the different studies that served as the basis for this study^{10,12,13}, a plastic card with the story printed in 16 size Arial-type font was placed in front of the child while each story was being read by the evaluator so as to exclude the interference of aspects related to memory. These tests were applied to both the experimental group, as well as the control group, in exactly the same manner, and then the scores were calculated.

We followed the same procedures as those used by O'Hare¹⁶. Children were tested individually in a silent room. The evaluator began the evaluation by explaining the following. "Here we have some stories and some questions. I will read you the stories. I want you to listen carefully and answer the questions at the end of each story." As previously stated, each of the 12 stories was selected^{9,16} to represent an attribute of the mental state. The stories were always applied in the same order and were read aloud to children. During the reading of each story and the questions asked by the evaluator, the appropriate printed card was also presented. However, the child was not permitted to read the paper aloud because that could interfere with the evaluation.

For the answers to the first questions, in other words, "Is it true what he (the character says," each story could be read a maximum of two times if the child asked for it to be repeated or if he answered "I don't know" to the question. Children received positive encouragement, but they were not given any signal regarding to the correct answer. When the researcher read the story for a second time, it was read from the beginning. The answer was registered, but the more important data for a more detailed analysis were the answers that were given to the second question, which evaluated the skill of mindfulness. Therefore, after answering the first question, the researcher posed the following second question. "Why did he (the character) say this?" For this second question, the story was re-read to the child until he answered the question or justified his answer.

All the children's answers were recorded to be scored and analyze later. A numerical register was used for the answers to the second question, with the intent of objectively verifying if children had interpreted the story correctly. Thus, if the child answered incorrectly or referred to the physical state (aspects of physical appearance, physical events, or results), the score was 0. When the child gave answers for the mental state that involved thoughts, feelings, desires, traces, and dispositions (i.e., using words such as want, happy, fear, think, play, pretend, lie, fooling someone, or waiting for something), the answer was characterized as a partial or total mental state attribute. When a partial mental state was indicated by the subject, despite including the mental state, the answer was given a score of 1 because the subject did not fully comprehend the exact state. When defining the total mental state, in other words, answers indicating complete comprehension and interpretation were given a score of 2. Therefore, the scoring for each subject could vary between 0 and 24 points, and the greater the score, the better the comprehension and interpretation of the stories.

The story involving double bluff included three questions, and the answer given to the third question was analyzed. Below you will find an example of the story and the points attributed to it based on the study by O'Hare et al¹⁶.

Story 1 (Lie):

Answers:

- Example of an incorrect answer (score 0): "He didn't have a toothache."
- Example of an answer referring to the physical state (score 0): "He had a toothache."
- Example of an answer referring to the partial mental state (score 1): "He doesn't like to go to the dentist."
- Example of an answer referring to the total mental state (score 2): "He doesn't want to go to the dentist."

RESULTS

For the statistical analysis of this study, a significance level of 0.05 (5%) was adopted. Throughout the study, all constructed confidence intervals were equivalent to 95%. Parametric statistics were used as the data were quantitative and continuous. The normality of the residuals was verified with this statistical model according to the normality test by Kolmogorov-Smirnov.

To better understand the results, the following SS were used. SS1, lie; SS2, white lie; SS3, misunderstanding; SS4, sarcasm; SS5, persuasion; SS6, contrary emotions; SS7, pretend; SS8, joke; SS9, figure of speech; SS10, double bluff; SS11, appearance/reality; and SS12, forget.

Initially, the median ages and IQs were determined for the experimental and control groups. This comparative analysis

between the groups for the two variables was done with ANOVA, and the results did not indicate any significant differences between the groups, attesting to the homogeneity of the sample.

Table 1 lists the median scores that were determined for each of the 12 stories for the two groups and relative to the total median scores of all stories in each group. Again, this analysis was done with ANOVA (Table 1).

Next, the Pearson correlation was applied separately in each group to measure the relationships among age, IQ, and the sum of the median scores of all stories (Total SS). To validate the correlations found, the correlation test was applied (Table 2).

Lastly, the alpha index was calculated to measure the internal consistency of the protocol or, in other words, of the 12 stories. Therefore, in the control group, the alpha values were

excellent, which indicated that the protocol had an excellent internal consistency (0.953).

DISCUSSION

The results indicated significant differences between the control group and the experimental group in the tasks of theory of mind, even after excluding the possibility of intellectual deficiency. These results demonstrated that the experimental group exhibited results that were much lower than the expected medians for children with typical development. These findings were similar to the results of studies that used the SS test⁸⁻¹³.

Table 1. Comparison of median scores in the experimental and control groups for each story and the total sum of all stories.

Group	Median	Score	Standard deviation	Min	Max	N	CI	p-value	
SS1	ASD	0.64	1	0.56	0	2	28	0.21	<0.001
	Control	1.38	1	0.62	0	2	56	0.16	
SS2	ASD	0.43	0	0.57	0	2	28	0.21	<0.001
	Control	1.18	1	0.69	0	2	56	0.18	
SS3	ASD	0.54	1	0.51	0	1	28	0.19	<0.001
	Control	1.14	1	0.67	0	2	56	0.18	
SS4	ASD	0.25	0	0.44	0	1	28	0.16	<0.001
	Control	1.14	1	0.67	0	2	56	0.18	
SS5	ASD	0.07	0	0.26	0	1	28	0.1	<0.001
	Control	0.96	1	0.69	0	2	56	0.18	
SS6	ASD	0.32	0	0.48	0	1	28	0.18	<0.001
	Control	1.39	1	0.62	0	2	56	0.16	
SS7	ASD	0.29	0	0.46	0	1	28	0.17	<0.001
	Control	1.36	1	0.55	0	2	56	0.14	
SS8	ASD	0.29	0	0.46	0	1	28	0.17	<0.001
	Control	0.96	1	0.66	0	2	56	0.17	
SS9	ASD	0.46	0	0.51	0	1	28	0.19	<0.001
	Control	1.11	1	0.71	0	2	56	0.18	
SS10	ASD	0.11	0	0.31	0	1	28	0.12	<0.001
	Control	0.96	1	0.66	0	2	56	0.17	
SS11	ASD	0.36	0	0.49	0	1	28	0.18	<0.001
	Control	1.14	1	0.59	0	2	56	0.15	
SS12	ASD	0.29	0	0.46	0	1	28	0.17	<0.001
	Control	1.2	1	0.55	0	2	56	0.14	
Total	ASD	4.04	4	1.26	3	7	28	0.47	<0.001
SS	Control	13.9	12.5	6.25	2	24	56	1.64	

SS: Strange Stories; ASD: autism spectrum disorders; Min: minimum; Max: maximum; N: number; CI: confidence interval.

Table 2. Pearson correlation coefficients for age, IQ, and the sums of scores of all stories in the experimental (with ASD) and control groups.

Correlated variables	Experimental group		Control group	
	Coefficient	p-value	Coefficient	p-value
Age versus IQ	-0.40	0.037	-0.02	0.875
Age versus total scores	-0.13	0.493	0.63	<0.001
IQ versus total scores	0.89	<0.001	0.15	0.259

ASD: autism spectrum disorders; IQ: intelligence quotient.

In this study, we verified that there was a greater correlation between IQ and the median total scores of the stories (positive and significant) in children with ASD. These findings were supported by the fact that children with ASD who had a greater intellectual performance were able to create more compensatory strategies for responding to the tests. A similar correlation has been described in a study by Kaland et al¹².

Many studies have indicated that high-functioning autistic individuals, such as those with Asperger's syndrome, exhibit good performances in the usual tasks of the theory of mind. However, when the tests are more complex and involve social context, a prejudice of these individuals is noted^{9,17}. For this reason, a more sophisticated test that involves the skills studied here is necessary.

There was a significantly negative correlation between age and IQ in this same experimental group, and this was not observed in the control group. This correlation might indicate a characteristic that is peculiar to the group with ASD who were selected for this sample and that did not interfere directly with the general results as there was no significant differences between the groups for median age and IQ. In the analysis that was done on the data from the control group, the correlation between age and the total score of the stories was positive and significant. This result was in agreement with the results of a study done by O'Hare et al¹⁶.

Hermelin and O'Connor¹⁸ have suggested that some individuals make use of cognitive mechanisms in order to deal with questions about affective processes and that, many times, these cognitive mechanisms cause these individuals to have insufficient answers to these questions, indicating inadequate responses in situations of social interaction. Thus, they feel strange in these situations. Nonetheless, these strategies do not affect all the social situations that involve the complex theory of mind, as was the case in this study.

The internal consistency of the protocol that was used in this study was measured. In the control group, the values indicated high internal consistency in that the presented results were consistent in children with typical development and that the answers increased adequately with increasing age, indicating the viability of this test for use in the Brazilian population.

During the application of the test, it was observed that many of the participants in the group with ASD supplied correct or partially correct answers but that these answers were given, many times, in an idiosyncratic manner. They mostly presented short answers to the questions and, when the answers were longer, they responded with elements that had little relevance. A study by Loth et al¹⁹ reported that individuals with ASD have difficulty learning elements that are relevant to the context. Very possibly, the individuals

that participated in this research were not attentive to the aspects that were relevant for the adequate interpretation of the mental states in the stories.

Many individuals in the experimental group tended to give literal answers about the stories and physical interpretations of social events. In her studies, Happé^{8,9,20} has described the prejudices of the theory of mind and the difficulty of individuals with ASD to understand nonliteral aspects of language, as they have great difficulty in interpreting something that is not said literally.

Revisiting the aspects of emphasis and relevance that have been described by Sperber and Wilson²¹ and Milosky²², it follows that the understanding of a determinate situation and social context is maximum when that which is most emphasized is also the most relevant. Beginning with the results that were obtained in this research, the possible difficulties that individuals with ASD exhibit in social situations and in activities involving social contexts were discussed because they could present the shortcomings in integrating the cues that are relevant for the context and difficulties in self-representation. These individuals could therefore have difficulties with emphasis and relevance, as they have difficulty in using the specific aspects of context that are required to know what is relevant for the understanding of the immediate task. It is probably very difficult for ASD individuals to focus their attention on what appears to be the most relevant information available, but, instead, their focus is determined by their objectives and by what interests them within the context. It is of extreme importance that the skills and difficulties of ASD subjects are evaluated through adapted tests that contemplate naturalistic situations and that, in this way, present questions for more realistic answers.

The translation and adaptation of the advanced test of the theory of mind Strange Stories and its application to children with ASD and those with typical development revealed that the theory of mind in children with ASD who are aged from 6 to 12 years were found to fall short of what was expected in children with typical development of the same age, suggesting that children with ASD have significant alterations in these skills. The results of this study suggested the importance of further research that involves new analyses (studies of reliability and evidence of validity) and discussions.

ACKNOWLEDGMENTS

We acknowledge the contribution of the colleagues of the Autism Spectrum Disorders Clinic, which is linked to the Postgraduate Program in Developmental Disorders in the Mackenzie Presbyterian University.

References

1. Wing L, Gould J. Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *J Autism Dev Disord* 1979;9:11-29.
2. APA - American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders - DSM-IV. Washington, DC: APA; 2002.
3. WHO - World Health Organization. International Classification of Diseases - ICD - 10 (Trad Centro Colaborador da OMS para Classificação de Doenças em Português). 8. ed. São Paulo: Editora Universidade de São Paulo; 2000.
4. Rutter M. Autism research: lessons from the past and prospects for the future. *J Autism Dev Disord* 2005;35:241-257.
5. Frith U. Autism and Asperger's syndrome. London: Cambridge University Press; 1996.
6. Baron-Cohen S, Leslie A, Frith U. Does the autistic child have a "theory of mind"? *Cognition* 1985; 21:37-46.
7. Baron-Cohen S. Mindblindness. Cambridge, MA: MIT; 1995.
8. Happé FGE. Autism: an introduction to psychological theory. London: UCI Press; 1993.
9. Happé FGE. An advanced test of theory of mind: understanding of story characters thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J Autism Dev Disord* 1994; 24:129-154.
10. Jolliffe T, Baron-Cohen S. The Strange Stories Test: a replication with high-functioning adults with autism or Asperger's syndrome. *J Autism Dev Disord* 1999;29:395-406.
11. Brent E, Rios P, Happé F, Charman T. Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism* 2004;8:283.
12. Kaland N, Moller-Nielsen A, Smith L, Mortensen EL, Callesen K, Gottlieb D. The Strange Stories Test: a replication study of children and adolescents with Asperger's syndrome. *Eur Child Adolesc Psychiatr* 2005;14:73-82.
13. Kaland N, Callesen K, Moller-Nielsen A, Mortensen EL, Smith L. Performance of children and adolescents with Asperger syndrome or high-functioning on advanced theory of mind tasks. *J Autism Dev Disord* 2008;38:1112-1113.
14. Wechsler, D. Wechsler Intelligence Scale for Children 3rd UKedn (Psychological Corporation, London, 1992. Trad Casa do Psicólogo Ed. 2002.
15. Sattler J. Assessment of children: WISC-II and WWPSI-R supplement. San Diego: Jerome M. Sattler Publisher; 1992.
16. O'Hare AE, Bremner L, Nash M, Happé F, Pettigrew LM. A clinical assessment tool for advanced theory of mind performance in 5 to 12 years olds. *J Autism Dev Disord* 2009;39:916-928.
17. Kaland N, Moller-Nielsen A, Callesen K, Mortensen EL, Gottlieb D, Smith L. A new advanced test of theory of mind: evidence from children and adolescents with Asperger syndrome. *J Child Psychol Psychiatr* 2002;43:517-556.
18. Hermelin B, O'Connor N. Logico-affective states and non-verbal language. In: Schopler E, Mesibov GM. Communication problems in autism. New York: Plenum Press; 1985:283-310.
19. Loth E, Gómez JC, Happé F. Do high-functioning people with autism spectrum disorder spontaneously use event knowledge to selectively attend to and remember context-relevant aspects in scenes? *J Autism Dev Disord* 2011;41:945-961.
20. Happé FG. Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition* 1993;48:101-119.
21. Sperber D, Wilson D. Relevance: communication and cognition. Cambridge: University Press; 1986.
22. Milosky LM. As crianças na escuta: o papel do conhecimento de mundo na compreensão da linguagem. In: Chapman RS. Processos e distúrbios na aquisição da linguagem. Porto Alegre: Artes Médicas; 1996.