Journal of Evidence-Based Psychotherapies,
Vol. 20, No. 2, September 2020, 79-100.

MIND READING IN FILMS TASK TO ASSESS SOCIAL COGNITIVE DEFICITS IN AUTISM SPECTRUM CONDITIONS

Soraya Tahazadeh[1], Usha Barahmand[2], Fereshteh Yaghooti[3], Mohamad Ali Nazari[4]*

[1] Department of Educational Sciences and Psychology, Tabriz University, Tabriz, Iran.
[2] Department of Psychology, Queens College, City University of New York, U.S.A.
[3] Department of Psychology and Education of Exceptional Children, Science and Research Branch, Islamic Azad University, Tehran, Iran.
[4] Department of Psychology, Faculty of Education and Psychology, Tabriz University, Tabriz, Iran.

Abstract

Background and Objectives. Various tasks with a variety of stimuli have been devised to measure aspects of theory of mind. In the present study, with due consideration of Iranian culture, we introduce a sensitive video-based test, called Mind Reading in Films Task (Films Task), for the evaluation of complex emotions and mental states. This new social ecological task for mindreading comprises several short film scenes, some measuring cognitive theory of mind and some measuring affective theory of mind.

Method. This study included two experiments. In experiment 1, the validity of the newly devised Films Task was investigated by comparing the responses of 342 students to the Films Task and to the Reading the Mind in the Eyes Task (Eyes Test). In experiment 2, the predictive power and sensitivity of the Films Task was assessed. Twenty adults with high functioning autism spectrum conditions (ASC) were compared with 20 matched healthy controls in terms of their responses to the Eyes Task, Films Task and Empathy Quotient questionnaire. The ROC curve was used to determine the best cut-off point and the diagnostic value.

Results. Our findings substantiate the discriminative capacity of the Films Task to distinguish individuals with autism spectrum conditions from their healthy non-clinical counterparts. Limitations: Intelligence and comorbid psychiatric conditions were not controlled, limiting the utility of the measure.

Conclusions. Results imply the potential utility of the Films Task as a viable alternative to the Eyes Task in measuring individual differences in social cognitive ability in the general population.

Keywords: autism, theory of mind, social cognition, empathy, complex emotions

* Correspondence concerning this article should be addressed to Dr. Usha Barahmand.
E-mail address: usha.barahmand@qc.cuny.edu

Introduction

The mindblindness theory proposes that impairments in social cognition in autism spectrum conditions (ASC) stem from varying degrees of impairment in theory of mind (ToM) abilities (Baron-Cohen, Leslie, & Frith, 1985). ToM is a multifaceted metacognitive ability that yields abstract, propositional knowledge about the mental states of others, thereby facilitating the implementation of social interactions (Fortier, Besnard, & Allain, 2018). ToM refers to the ability to infer and reason about the beliefs, thoughts or emotions of others in order to predict their behavior (Lombardo & Baron-Cohen, 2011). Coricelli (Coricelli, 2005) distinguishes between cognitive ToM, making inferences about the mental states of others with no emotional involvement, and emotional ToM, the cognitive ability needed to understand the emotional feelings of others.

A construct linked to ToM is empathy, defined as not only the ability to understand what others think, but also to experience it with no confusion within oneself (Decety & Lamm, 2006). Empathy has been proposed as a broader neurocognitive construct and extends the ToM hypothesis (Baron-Cohen, 2002). Like ToM, empathy plays a vital role in human relationships and enables an individual to make sense of and predict the behavior of another (Dziobek et al., 2008; Grove, Baillie, Allison, Baron-Cohen, & Hoekstra, 2014; Smith, 2006; Sucksmith, Allison, Baron-Cohen, Chakrabarti, & Hoekstra, 2013). Generally, researchers agree that empathy comprises both cognitive and affective components. Cognitive empathy is conceptualized as the ability to understand and identify the emotions and emotional experiences of another, without necessarily resonating, whereas affective empathy can be defined as responding to other people's mental states with an appropriate emotion (Baron-Cohen & Wheelwright, 2004; Chakrabarti & Baron-Cohen, 2006) Recent studies have shown that the term cognitive empathy overlaps with ToM and the terms are used interchangeably (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).

With respect to ASC, one study suggested that it is the integration of the cognitive and affective ToM that is impaired and not the facets per se (Shamay-Tsoory, Tomer, Yaniv, & Aharon-Peretz, 2002), while several other studies have reported only cognitive empathy is impaired in ASC with affective empathy relatively intact (Dziobek et al., 2008; Jones, Happé, Gilbert, Burnett, & Viding, 2010; Mazza et al., 2014).

Various tasks with a variety of stimuli have been devised to measure the facets of ToM. These include false belief tasks (Wimmer & Perner, 1983), Strange Stories (Happé, 1994), Reading the Mind in the Eyes task (Baron-Cohen, Wheelwright, Hill, et al., 2001), The Awkward Moments Test (Heavey, Phillips, Baron-Cohen, & Rutter, 2000), Silent Animations (Castelli, Frith, Happé, & Frith, 2002), Reading the Mind in Films task (Golan, Baron-Cohen, Hill, & Golan, 2006), Reading the Mind in Voice test (Golan, Baron-Cohen, Hill, & Rutherford, 2007),

Awareness of Social Inference Test (Mathersul, McDonald, & Rushby, 2013), and The Strange Stories Film Task (Murray et al., 2017).

Most studies investigating emotion recognition deficits in ASC and focusing on the recognition of the so-called basic emotions from discrete emotional stimuli such as dynamic or static facial expressions or voice recordings have not yielded conclusive findings, while others employing the Eyes test or tasks assessing the ability to detect mental states from contextual cues have shown that individuals with ASC display deficits in ToM (Golan et al., 2006). Other research has shown that the ability to succeed in such ToM tasks varies with age, IQ, and the nature of the task (Begeer, Malle, Nieuwland, & Keysar, 2010).

Atypicalities in facial emotion recognition have been reported in ASC (Lozier, Vanmeter, & Marsh, 2014). The Reading the Mind in the Eyes Test (Eyes Test) is an advanced ToM task in which individuals are required to attribute mental states and recognize complex facial emotions from photographs displaying only the eye area of the face (Baron-Cohen et al., 2015). The Eyes Test has been shown to be strongly related to measures of emotion perception ability (Petroni et al., 2011) and used as a measure of emotion recognition (Guastella et al., 2010).

The Eyes test has been evaluated in over 250 studies to date, and has been shown to have sound psychometric properties (Khorashad et al., 2015; Prevost et al., 2014; Vellante et al., 2013). Individuals with autism consistently perform less well relative to controls (Baron-Cohen, Wheelwright, Hill, et al., 2001; Demurie, De Corel, & Roeyers, 2011; Kaland, Callesen, Møller-Nielsen, Mortensen, & Smith, 2008; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007) indicating defitis in social cognition in ASC.

The psychometric properties of the Persian version of the Eyes Test have been investigated and revealed acceptable test-retest reliability but poor internal consistency (Khorashad et al., 2015). It is of interest to note that in many Iranian subcultures, avoiding direct eye contact during social interactions is considered respectful and is learned in the process of socialization. Accordingly, people rely on facial expressions, body movements, or tone of voice to decipher the mental states of others. Therefore, we decided to design and develop a tool using naturalistic social stimuli an alternative to eye reading as a measure of theory of mind and social cognition ability. Golan, et al. (2006) used the Reading the Mind in Films to assess the ability of adults with ASC to recognize complex emotions and mental states. They reasoned that short film footages of complex social interactions may better represent real life socioemotional interactions and, therefore, provide a more valid measure of the ability of individuals with ASC to interpret complex emotions. They argued that in everyday social interactions, one makes judgments of others' mental states and emotions by integrating visual and auditory input with contextual cues into a coherent whole. However, we believe that the footages that Golan et al., (2006) used in their test is not congruent with Eastern cultures and to better represent socio-economical interactions typical in Eastern culture, a modified tool is essential. The purpose of this study was to develop a sensitive tool to assess cognitive and affective ToM in both clinical and non-clinical samples in Iran.

The current study is an attempt to systematically replicate the Golan et al., (2006) study and film scenes depicting emotions of varying levels of complexity were utilized. To distinguish our task from that used by Golan et al., (2006), we name our task the Mind Reading in Films Task (Films Task). First, the validity of the Films Task was assessed by examining the correlation between the Films Task score and the Eyes Test score in the general population. In line with the finding by Baron-Cohen et al. (2001) regarding sex differences in Eye Test scores, we hypothesized that women would score significantly higher than males in both the Films Task as well as on the Eyes Test. Next, the Films Task was administered along with the Eyes Test to adults with high functioning autism in order to test the predictive power and sensitivity of this instrument compared to the Eyes Test. To minimize confounding, a control group matched in terms of age, gender, and education was also included in the research design.

In addition, on account of its likely association with ToM, empathy (cognitive and affective) was also studied. It was anticipated that empathy would correlate with scores on Films Task, thereby enhancing its validity. ROC curves were constructed to visualize the performance of the Films Test and Eyes Test as binary classifiers of the two output classes, the autism group and the healthy controls. For comparison purposes, ROC curves were also constructed to test the predictive ability of EQ and AQ.

Method

Development of the Mind Reading in Films Task (Films Task)

Initially, 40 short scenes sampled from 5 Indian movies and depicting complex mental states and emotions were selected. To reduce the likelihood of any participant having already watched the movies and therefore, the potential impact of memory of the scenes, films from the years 1993 to 2003 were used (any participants who had watched any of the movies were not included in the study).

The selected scenes portrayed various emotions and mental states, such as secretiveness, insolence, doubt, bitterness. Each of the 40 scenes was evaluated by three faculty members of the department of Psychology, at the University of Tabriz, and four mental states were assigned as options to each film scene, with one of them being the target or correct label and the other three being foils. Participants were required to indicate their judgment by selecting any one of the four options. The correct label received a score of one point. Consistent with Golan (Golan et al., 2006), the target word and the foils were selected from the 412 mental and emotional states in six developmental levels outlined in the Baron-Cohen taxonomy of emotions (Baron-Cohen, Golan, Wheelwright, & Hill, 2004) such that they were matched for verbal difficulty. The selected foils matched some but not all of the emotional information in the scene. A glossary of all the descriptive target labels and foils included was prepared for participants to use during judgment of the scenes.

Questions

Each of the 40 film clips was followed by a question aimed either at assessing emotion recognition (question type A) or beliefs and intentions (question type B) of the protagonist. Twenty-four film clips assessed emotion recognition and the question asked at the end of each clip was "At the end of the scene, how is the man/woman feeling?" That is, how does the person feel about what the other party did/said? The response options provided were, for example, annoyed, hate, angry, or surprised. These film scenes were utilized to measure the affective and emotional states (affective ToM) attributed to the protagonist. The remaining 16 film clips were followed by the question, "At the end of the scene, what is the intention of the man/woman?" That is, what does X want the other person to understand from his/her words or reaction? What does he/she want to convey? The response options provided were, for example, secretiveness, sarcasm, determination, blaming. These scenes measured the thoughts, beliefs and intentions attributed to the protagonist (cognitive ToM). The presentation of scenes associated with question sets A and B was randomized. For the assessment of emotions, items of negative, positive and neutral valence were considered: 13 scenes with negative emotional valence, such as bitterness, horror, distress, and 7 scenes with positive emotional valence, such as surprise, happiness, amusement and 4 scenes with neutral emotional valence, such as doubt, confusion, bewilderment were selected. This frequency distribution is similar to that considered by Dziobek et al. (2006). A sample of the questions is presented in Figure 1.

Police: Listen Beta, they have caught Maruf too.

(**A**) What does the convict feel at the end of the scene?
Worry Remorse **Hate** Confusion

Man: Are you not happy and content with living with me?
Woman: Of course, I am.
(**B**) How does the woman sound at the end of the scene?
Satisfied **Sarcastic** Annoyed Complaining

Figure 1. Shows a sample of emotion (A) and intention (B) recognition questions

Pilot Study

A sample of 40 university students (23 males, 17 females) with the mean age of 22.1 years (*SD*=3.03) was randomly selected and individually presented 40 short film scenes (Films Task) for a total duration of 359 seconds (Min=4, Max=15).

Before the test began, participants were provided and familiarized with the glossary of all the descriptive target labels and foils included and informed that they could use them during the test. Each participant was allowed to watch the films without an imposed time limit and select his/her desired answer to the questions posed. Items were excluded if the target word was picked by less than 60% of the participants. This resulted in 4 film clips being dropped.

Experiment 1

Measures

Mind Reading in Films Task (Films Task)

With the exclusion of 4 films in the pilot study phase, the final set of thirty-six short film scenes with a total duration of 325 seconds (*Min*=4, *Max*=15) was used. Each clip was followed by a question that assessed recognition of emotions (set A) or beliefs and intentions (set B) of the protagonist. A glossary of all the descriptive target labels and foils was also included.

Reading the Mind in the Eyes Task (Eyes Test) (Baron-Cohen et al., 2001)

Using a fixed-choice paradigm, the task requires participants to describe the emotional/mental state of a person based on only an image of their eyes. The test consists of 36 grey-scale cropped photos of the area around the eyes of people taken from magazines. Each photo is accompanied by four mental state terms and the participant is instructed to choose within 20 seconds the one word best describes feeling or thinking reflected by the eyes. Only one of the four items is considered to be correct based on judgements provided by a panel of judges in the initial psychometric study. All responses are coded on a binary nominal scale as correct or incorrect yielding a possible total score of 36. The Eyes Test, considered as one of the most effective socio-cognitive tasks available (Pinkham, 2014), has been translated into Persian and used in some studies (Nejati, Zabihzadeh, Maleki, & Tehranchi, 2012). In the current study, the Persian version of the test was used.

Participants and Procedure

For experiment 1, 353 students (197 males, 156 females) majoring in humanities, basic sciences, and technical sciences, with a mean age of 21.6 years (*SD*=2.3) were selected through stratified cluster sampling. The 36 films were played individually for each participant. Participants watched each film clip up to 3 times and then chose his/her response. At this stage, the Eyes Test (Baron-Cohen, Wheelwright, Hill, et al., 2001) was used as well. The order of presentation of the Films Task and Eyes Test was counterbalanced. The duration of these two tests lasted roughly 40 minutes for each participant.

Results

Socio-Demographic Data

Out of the 353 participants, 11 were removed from the sample as they were found to be multivariate outliers. The final sample consisted of 342 participants, among whom 44.2% (*n*=151) were female and 55.8% (*n*=191) were male, with an mean age of 21.6 years (*SD*=2.21, *Min*=18, *Max*=30). The distribution of academic degrees among participants was high school diploma 14.6% (*n*=50), bachelor's degree 76.8% (*n*=232) and master's degree 17.5% (*n*=60). The distribution of majors was humanities from 6 schools 29.8% (*n*=102), basic sciences from 4 schools 31.6% (*n*=108) and the technical sciences from 5 schools 38.6% (*n*=132).

Item Analysis

Using the Mahalanobis Distance (Mahalanobis, 1930) 11 multivariate outliers (Films Task with Eyes Test) were identified and eliminated. After removing 11 participants, the average score of all participants in Films Task was 21.64 (*SD*=3.74, *min*=10, *max*=30). In Table 1, the participants' responses to each question and foils in the Films Task are displayed. Next, to evaluate the quality of each of the 36 items and of the test as a whole and to ensure the required degree of validity in the Films Task, item analysis was undertaken by computing item difficulty, item distractor analysis (Baron-Cohen, Wheelwright, Hill, et al., 2001), item discrimination index, and item discrimination coefficient (Crocker & Algina, 1986).

Table 1. Distribution of Responses to MRFT in Percentages (N=342)

Item	Target	Foil 1	Foil 2	Foil 3	Item	Target	Foil 1	Foil 2	Foil 3
Q1	79.5	6.1	2	12.3	Q19	74.3	7.6	15.5	2.6
Q2	53.5	24	12.3	10.2	Q20	67.3	10.2	3,2	19.3
Q3	**48.5**	**3.2**	**12.3**	**36.5**	Q21	51.8	17.5	7	23.7
Q4	**71.3**	**26.3**	**5.8**	**7**	Q22	67	6.1	7	19.9
Q5	66.7	1.5	9.4	22.5	Q23	54.4	10.5	24	11.1
Q6	**47**	**21.8**	**18.7**	**12.9**	Q24	64.3	3.2	10.8	21.6
Q7	62	17.5	8.5	12	Q25	50	17	2.3	20.8
Q8	65.2	7.9	11.4	15.5	**Q26**	**70**	**21.2**	**1.8**	**7**
Q9	**45.6**	**9.6**	**2.3**	**42.4**	**Q27**	**62**	**25.5**	**4.2**	**7.3**
Q10	60	8.9	6.6	24.7	Q28	57.9	24	9.4	8.8
Q11	69	17.3	5.3	8.5	Q29	50	12.3	4.6	23
Q12	56	23.3	14.8	6.2	Q30	80.8	6	9.1	4.1
Q13	73.4	14	3.5	9.1	Q31	58.5	8.2	8.8	24.5
Q14	**45.6**	**12**	**24.3**	**18.1**	**Q32**	**24.6**	**23.7**	**24.3**	**23.7**
Q15	65	8.1	24.2	5.9	Q33	67	1.8	14.9	16.4
Q16	**27.8**	**25.4**	**24.3**	**26**	Q34	66	16.3	11.7	6
Q17	78.4	2	12	7.6	**Q35**	**55**	**14.3**	**26**	**4.7**
Q18	80.4	2.6	15	1.8	Q36	52.3	2.9	25	19.6

Note. The highlighted items have failed to fulfill either criteria (1) or (2)

Questions 3, 4, 6, 9, 14, 16, 26, 27, 32, 35 failed to meet the criteria for distractor analysis (Baron-Cohen, Wheelwright, Hill, et al., 2001); questions 3, 4, 8, 9, 16, 26, 27, 32, 35 were found to have weak discrimination coefficients (see Table 2); questions 3, 4, 8, 9, 16, 27, 32 had weak discrimination index (Di) (see Table 3); and questions 8,11, 26, 35 were determined to be too easy or difficult and, therefore, inappropriate. A total of 9 questions (3, 4, 8, 9, 16, 26, 27, 32, 35) emerged as inappropriate based on all of the above criteria and eliminated. Of the final set of 27 questions considered appropriate, 16 were related to identifying emotions (set A) and 11 to identifying intentions (set B). The difficulty index of these questions ranged between 30% and 70% with an average of 60.45% (see Table 3).

Table 2. Percentage of Participants in the Upper and Lower 27th Percentile Choosing the Target in each Item

Item	Upper 27%	Lower 27%	Item	Upper 27%	Lower 27%
Q1**	100	38.1	Q19**	93.5	38.1
Q2**	77.4	33.3	Q20**	83.9	38.1
Q3	61.3	57.1	Q21*	77.4	47.6
Q4	64.5	42.9	Q22**	74.2	33.3
Q5**	77.4	28.6	Q23**	83.9	14.3
Q6**	71	23.8	Q24*	77.4	47.6
Q7**	80.6	33.3	**Q25****	74.2	19
Q8	71	57.1	Q26	80.6	67.1
Q9	51.4	42.9	Q27	77.4	52.4
Q10**	80.6	33.3	Q28**	74.2	38.1
Q11**	90.3	47.6	Q29**	74.2	23.8
Q12*	67.7	33.3	Q30**	90.3	52.4
Q13*	77.4	47.6	Q31**	90.3	33.3
Q14**	77.4	28.6	Q32	35.5	19
Q15**	77.4	33.3	Q33**	77.4	38.1
Q16	41.9	19	Q34**	80.6	42.9
Q17**	93.5	42.9	Q35	50.6	43.3
Q18*	90.3	61.9	Q36**	80.6	14.3

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Table 3. Item Analysis of MRF Test: Difficulty Index (Pi)a and Index of Discrimination (Di)b for Each Item Expressed as Percentages

Item	Di	Pi	Item	Di	Pi
Q1	44.23	69.23	Q19	40.38	67.30
Q2	32.69	59.61	Q20	34.61	65.38
Q3	**13.46**	59.61	Q21	26.92	65.38
Q4	**19.23**	55.77	Q22	30.76	57.69
Q5	34.61	57.69	Q23	44.23	55.76
Q6	32.69	51.96	Q24	26.92	65.38
Q7	34.61	61.53	Q25	36.53	49.32
Q8	**19.23**	**76.92**	Q26	25	**71.15**
Q9	**13.46**	48.08	Q27	**17.30**	67.30
Q10	34.61	61.53	Q28	28.84	59.61
Q11	34.61	**73.07**	Q29	34.61	53.84
Q12	26.92	53.84	Q30	32.69	69.23
Q13	26.92	65.38	Q31	39.46	65.38
Q14	34.61	57.69	Q32	**13.46**	**28.84**
Q15	32.69	59.61	Q33	30.76	61.53
Q16	**17.30**	32.69	Q34	30.76	65.38
Q17	39.46	67.30	Q35	24.61	**76.92**
Q18	30.76	65.38	Q36	42.3	53.84

a Very difficult and improper (p<30); Average and proper (30 <p<70); Very easy and improper (p>70)
b D>39 = Excellent; 39>D>30 = Good; 29>D>20 = Mediocre; 19>D>0 = Poor; -1>D =Worst

Gender Differences

The results showed that on average women performed better on both Eyes tasks (Eyes Test) and on film tasks (Films Task). Scores of females (M=22.62, SD=3.7) on the Eyes test was significantly higher than that of males (M=21.27, SD=3.6), t=3.41, p=0.001. Also, in the Films Task, females (M=17.52, SD=3.0) scored significantly higher than males (M=16.44, SD=3.5), t=3.01, p=0.003 .The chi-square test was used to determine any significant gender differences in responses to each of the questions. Women performed better than men on six questions of which, three (A5, A10, A26) related to emotion (sadness, frustration, and fear) recognition and three (B4, B24, B27) related to identifying intentions (blaming, sarcasm, and decisiveness).

Scatter Plot and Correlations

The scatter plot shows the linear regression relationships between the Films Task and Eyes Test in the general population (R Sq Linear=0.32). Pearson's correlation coefficient between the Films Task and Eyes Test scores was used to determine the convergent validity of the Films Task. A significant correlation was found r=0.56, p <0.01). The mean score for the 27 films in 342 participants was

equal to 16.90 (*SD*=3.35, *Min*=6, *Max*=24). Skewness (*sk*=-0.431) and Kurtosis (*ku*=0.19) were calculated and the percentage of the correct answers to the 27-item questionnaire ranged from 45.6% to 80.8%, with a mean of 63%. The percentage of the correct answers of these individuals can be seen in Table 1.

Experiment 2

Participants

The sample included patients clinically diagnosed with autism and healthy controls with no diagnosis of a mental disorder. Twenty patients with ASC (17 males, 3 females) were selected using a purposive sampling strategy with the following inclusion criteria 1) a clinical diagnosis of an autism spectrum condition (ASC) according to DSM-5 (autism spectrum disorder) at a recognized specialist clinic by a psychiatrist or clinical psychologist 2) age ≥ 18 years 3, 3) verbal intelligence above 70 based on WAIS-IV 4) educational status of at least a high school diploma; 5) absence of psychiatric disorders (schizophrenia, personality disorder, substance use disorder) 6) absence of neurological disorder (brain damage, ADHD). Non-clinical control subjects were recruited locally by flyers, which did not specify the purpose of the study or the psychological traits or syndrome under investigation.

Control group participants were included if they had no diagnoses of ASC, and no first-degree relatives with ASC. Participants were excluded if they reported a diagnosis of any psychological disorder. From among the 113 volunteers twenty (17 males, 3 females) were selected after being matched with the ASC group in terms of age (ASC group:22.3 ± 3.1 years; control group: 23.6 ± 3.9 years, t=0.384, p =.703), gender and years of education (ASC group: 14.8 ± 1.64; control group: 15 ± 1.65, t=0.77, p =.442). The AQ was then administered to both groups, the clinically diagnosed individuals with autism and the healthy controls. The ASC group AQ scores ranging from 28 to 38 with a mean of 32.45 (*SD*=3.05) differed significantly from the control group AQ scores which ranged between 11 and 28 with a mean of 20.35(*SD*=3.66), t=7.93, p <.001.

Measures

Reading the Mind in the Eyes Test (Eyes Test) (Baron-Cohen et al., 2001)

The test has been described above. The Persian version of this test (Nejati et al., 2012) was used.

Mind Reading in Films Task (Films Task)

The final set of 27 film clips with a total duration of 242 seconds (*Min*=4, *Max*=15) were used. Of these, 16 included questions related to identifying emotions (set A) and 11 to identifying intentions (set B).

Empathy Quotient (EQ) (Baron-Cohen et al., 2004)

The EQ is a self-report measure assessing both cognitive and emotional empathy (Baron-Cohen et al., 2004). The EQ is a 60-item questionnaire, with 40 questions tapping empathy and 20 filler items. The instrument is scored on a scale of 0 (weak empathy) to 80 (strong empathy) (Baron-Cohen et al., 2004). Higher scores are indicative of increased levels of self-reported empathy. The EQ shows good test-retest reliability (r=0.97, p <0.001. Two studies have highlighted a three-dimensional structure including cognitive empathy(CE), emotional empathy (ER) and social skills (ES) (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004).

Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001)

The AQ is a forced choice self-report questionnaire, which measures the degree to which any adult of normal intelligence possesses traits related to the autistic spectrum. It consists of 50 questions, with 5 sets of 10 questions, each measuring domains of autistic traits, social skill, attention switching, attention to detail, communication and imagination. Scores range from 0 to 50, and the higher the score, the more the autistic traits a person possesses. Adequate reliability has been reported for this questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). Since the measure has been found to have acceptably high sensitivity and specificity at a cut-off score of 26, in the present study the score above 26 on the AQ was used as an indicator of autistic traits.

Procedure

In this phase, we evaluated the sensitivity and predictive power of the Films Task in discriminating the ASC group from healthy controls. The 27 film clips that emerged appropriate in experiment 1, totaling a duration of 242 seconds (*Min*=4, *Max*=15), were converted to a software and played for the ASC patients and matched healthy controls. Testing was computerized and performed individually. Each participant first filled out a demographic data sheet. Next, the stages of the experiment and instructions of how to respond to the questions were given to each individual along with an example. In this software, as soon as the start button is pressed, after 1 second the first question appears on the top of the screen. Simultaneously, the film and the options appear at the bottom of the page. After a 2 second interval, the film is automatically replayed and stops on the protagonist. If the participant cannot make a judgment and select the appropriate option after watching the film clip twice, he or she can click the repeat button on the right side of the screen, and after a second, the film will be displayed once more. Each short film could not be played more than three times and the participant could not proceed to watch the next film clip until one of the options was selected. Participants completed the Films Task, Eyes Test, Autism Spectrum Quotient (AQ), and Empathy Quotient (EQ) questionnaires all in counterbalanced order.

Statistical Analysis

The proportion of correct answers in both groups was first examined followed by a test of the validity and internal consistency of the Films Task. Scores of individuals with autism and those of their matched healthy counterparts on the Films Task, Eyes Test and EQ (cognitive and emotional empathy) were compared using independent samples t-tests. Pearson's correlation coefficients and scatter plot were obtained to determine any significant associations between the instruments. Finally, the ROC Curve analysis was used to determine the sensitivity of the Films Task in assessing complex emotion and mental state recognition ability and its discriminative capacity to assign participants to their respective groups.

Results

Repartition of Responses

To determine if there a significant relationship between proportion of correct responses to each question and group membership, a Chi-Square statistic was computed and the results are presented in Table 4. Although the healthy controls performed better than the ASC group on all items, the difference in performance was significant only for 19 questions, 11 of which pertained to emotion recognition or question set A and the remaining 8 pertained to mental state (intention) recognition or question set B.

Table 4. Percentage of Participants Choosing the Target in Each Item in Each Group

Item	Target	Autism	Normal	Item	Target	Autism	Normal
A1**	Hesitant	30	75	B15*	Sneer	25	60
A2*	Bitter	15	45	A16**	Fear	35	85
A3*	Angry	60	90	A17**	Confused	20	75
B4**	Resolute	25	95	B18	Helpless	55	65
A5	Hurried	40	55	B19**	Sympathetic	25	70
A6	Bethink	35	50	A20**	taken aback	20	70
B7**	Sneaky	10	70	A21	Ashamed	45	55
B8**	Bold	25	80	A22	Humiliated	45	60
A9**	Hate	40	70	B23	Sarcastic	60	80
A10*	Sad	30	75	B24**	Blaming	5	45
A11	Sulky	45	60	A25*	Surprise	25	60
A12**	Terrified	40	80	A26*	Horrified	35	75
B13**	Conceited	10	70	B27**	Taunting	40	100
B14	Annoyed	55	75				

* $p<0.05$; ** $p<0.01$ A= Emotion Recognition, B= Intent Recognition

Reliability Analysis

The internal consistency of the entire test, Cronbach's alpha, was 0.85 with a 95% *CI* from 0.769 to 0.907, while that of the 16 emotion recognition questions (set A) was 0.69 and that of the 11 mental state recognition questions (set B) was 0.78. Test-retest reliability was also assessed by comparing the responses of 20 individuals from the control group who had mean scores of 18.95 (*SD*=2.78) at the time of initial testing (T1) and 18.75 (*SD*=2.45) 6 weeks later at retest. Spearman's correlation coefficient revealed a positive and significant correlation (*rs*=0.89, *p*<.001) between the two sets of scores. A Wilcoxon Signed-ranks test indicated no significant difference between the performance of individuals at test (*Mdn* = 18.02) and retest (*Mdn* = 17.96), *Z* = 1.08, *p* >.05, *r* = .78

Group Differences

The differences between the mean scores of the two groups on the Films Task, Eyes Test and EQ are shown in Table 5. There were very significant differences of substantial effect size seen between the mean scores of the groups on the Films Task, Eyes Test and cognitive empathy (EQ-CE) subscale of the EQ, with lower scores observed in the autism group. Although a similar but moderately significant group difference was observed on the emotional empathy (EQ-ER) subscale, the effect size was small. The groups did not differ with regard to social skills (EQ-ES) component of the EQ and response time to the Films Task.

Table 5. Differences between Mean Scores of the Two Groups on Tasks MRFT, RMET, EQ

	Autism (20)		Control (20)		Effect Size	
	M (SD)	Range	*M (SD)*	Range	*t* (38)	Cohen's d
MRFT	8.95 (3.6)	4-16	18.95 (2.8)	15-24	9.78**	3.10
RMET	9.25 (2.7)	5-15	20.05 (3.4)	13-26	11.05**	3.52
EQ						
EQ-CE	3.05 (1.6)	1-7	16.45 (3.1)	10-22	17.15**	5.43
EQ-ER	8.15 (1.5)	5-10	9.15 (1.4)	7-13	2.16*	.689
EQ-ES	5.35 (1.4)	3-8	6.10 (1.4)	4-9	1.67	.536
Response Time	616.50 (160.02)	384-985	675.25 (118.65)	434-838	1.32	.417

* *p*<0.05; ** *p*<0.01 *d*=.2 'small' effect size, *d*=.5 'medium' effect size *d*>.8 'large' effect size

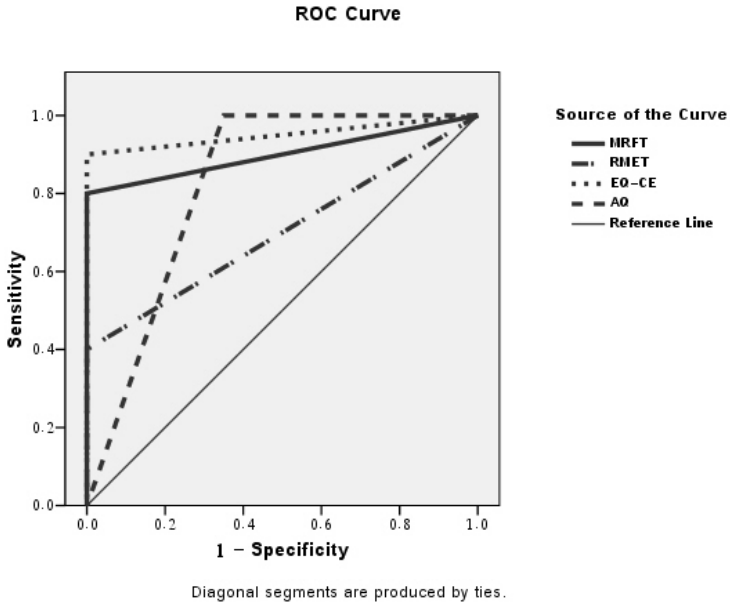
Correlation

Pearson's correlation coefficient between mind reading scores on Films Task and Eyes Test is significant and positive *(r*=0.75, *p*<.001), indicating good convergent validity of the Films Task. The correlation coefficients between the autism scores obtained by the AQ, and mind reading scores obtained by the Films Task (*r*=-0.63, *p*=0.000) and the Eyes Test (*r*=-0.57, *p*=0.000) are negative and significant, implying that the greater the extent of clinically significant autism

spectrum disorder traits, the lower the mind reading ability as measured by the Films Task and Eyes Test. Films Task scores also correlated positively and significantly with cognitive empathy, EQ-CE, (r=0.77, p<.001) and with affective empathy, EQ-ER (r=0.52, p=0.001). Significant positive correlations were also found between the emotion recognition questions (set A) and affective empathy, EQ-ER, (r=0.36, p=0.024) and between the mental state recognition questions (set B) and cognitive empathy, EQ-CE, (r=0.74, p<.001) However, Films Task scores failed to correlate significantly with communication skills scores, EQ-ES (r=0.23, p=0.155) or participants' age (r=0.25, p=0.122).

Cut off and ROC Curves

The ROC curves were constructed to compare the ability of each of the measures used to discriminate between those clinically diagnosed with autism from their matched healthy counterparts. Each instrument was coded based on its respective cut-off score and its sensitivity was assessed using the ROC curve based on the two groups of patients and healthy controls (diseased vs. nondiseased). For the Films Task, the cut-off point was determined as ≥17(based on the scores of 342 individuals). That is, scores below 17 on the Films Task were considered indicative of deficits in ToM. ROC analyses were carried out for EQ and AQ as well only for the purpose of comparison with the Films Task (Figure 2). In order to construct the AUC for the AQ, a cut-off score of 26 was used.



Note. MRFT=Mind Reading in Films Task (Films Task); RMET = Reading the Mind in Eyes Test (Eyes Test); EQ-CE = Emotion Quotient-Cognitive Empathy; AQ = Autism Quotient

Figure 2. ROC curves for each of social cognition tasks

The ROC curves indicating the ability of the Films Task and other scales to correctly classify participants into their respective groups are displayed in Figure 2. Only the tests on which there was a statistically significant difference between the two groups were used. The Area Under the Curve (AUC) values for each of the measures were as follows: Films Task (AUC=.90, $p<.01$), EQ-CE (AUC=.95, $p<.01$), AQ (AUC =.82, $p<.01$) and Eyes Test (AUC=.70, $p<.05$). The sensitivity of Films Task in correct assignment of individuals to their respective groups exceeded that of both the Eyes Test and the autism spectrum questionnaire (AQ). In some previous studies, Eyes Test sensitivity was reported to be 79% (Dziobek et al., 2006) and 71% (Murray et al., 2017).

Discussion

The current study utilized Films Task as a new social ecological task to assess individual differences in the ability for the recognition of emotions and complex mental states from short social scenes in films. Understanding non-verbal expressions of complex emotions and intentions requires advanced ToM. The purpose of the present study was to compare the Films Task with the Eyes Test, a widely accepted and valid measure of ToM, to determine its convergent validity and its sensitivity in correctly discriminating individuals with autism spectrum conditions (ASC) from healthy controls without ASC.

Overall, findings substantiate the discriminative capacity of the Films Task to distinguish individuals with ASC from healthy nonclinical counterparts. The fairly acceptable correlation between Films Task and Eyes Test in the general population and the relatively high correlation between the two tests in individuals with ASC, also implies the potential utility of the Films Task as a viable alternative to the Eyes Test in measuring individual differences in social cognitive ability.

In general, using film stimuli, the Films Task opens a window toward social understanding in ASC, which serves to expand and add to previous research. In the Films Task, high-functioning adults in the autism spectrum condition (ASC) fared significantly worse than their healthy counterparts, a significant difference observed in the results of previous studies using film tasks (Dziobek et al., 2006; Golan et al., 2006; Heavey et al., 2000; Lombardo et al., 2007; Mathersul et al., 2013; Murray et al., 2017; Wilson et al., 2014). In contrast to the instruments used in these prior studies, the Films Task included a wide range of emotions and complex mental states, and measured both the emotional and cognitive dimensions of theory of mind.

Scores on the Films Task correlated positively and significantly with the Eyes Test and EQ, confirming its high convergent validity. Results also showed a positive and significant association between cognitive empathy (EQ-CE) and Eyes Test, which is consistent with many studies (Grove et al., 2014; Lawrence et al., 2004; Olderbak et al., 2015); however, the emotional empathy and social skills

subscale scores failed to correlate with the Eyes Test scores. This finding could be interpreted as an implication that the deficit in complex emotions and mental states recognition which is characteristic of autism spectrum conditions is a facet of cognitive empathy (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003; Baron-Cohen & Wheelwright, 2004; Lawrence et al., 2004).

Scores on the Films Task correlated with both types of empathy, cognitive and emotional. Previous research (Blair & Viding, 2008; Shamay-Tsoory & Aharon-Peretz, 2007; Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003) has endorsed an association between cognitive empathy and cognitive ToM, and between affective empathy and affective ToM. Therefore, the finding that the Films Task scores correlate significantly with both types of empathies, cognitive (EQ-CE) and emotional (EQ-ER), probably implies that the Films Task measures both cognitive ToM and affective ToM.

In the present study, the empathy (EQ) scores of patients with autism were significantly lower than those of the control group, which is consistent with other studies (Baron-Cohen et al., 2015; Baron-Cohen et al., 2014). Compared to the control group, subjects with ASC had a significantly lower score in cognitive empathy while no difference was observed in emotional empathy. These results were reported previously in clinical studies of adults and adolescents (Dziobek et al., 2008; Jones et al., 2010; Rogers, Dziobek, Hassenstab, Wolf, & Convit, 2007; Rueda, Fernández-Berrocal, & Baron-Cohen, 2015). And since this information was obtained through self-report, it indicates that high-functioning adults with ASC have some awareness of the difficulties they have with understanding the mental states of others. The results are also consistent with studies that have reported more difficulties in the area of cognitive empathy than emotional empathy in ASC individuals (Baron-Cohen, 2011; Mazza et al., 2014). Deficiencies in cognitive empathy have been reported as relatively characteristic of ASC (Baron-Cohen, 2000; Koehne, Hatri, Cacioppo, & Dziobek, 2016) including the broad autism phenotype (Sucksmith et al., 2013). Deficits in imitation (Williams, Whiten, Suddendorf, & Perrett, 2001) with an underlying malfunctioning of the mirror neuron system (Oberman & Ramachandran, 2007) and alexithymia (Aaron, Benson, & Park, 2015) have been suggested as contributing to deficient ToM in ASC. However, future research may shed light on whether the mirror neuron system dysfunction applies to empathy in general or to cognitive empathy in particular.

Eye reading has been mainly investigated in studies comparing ASC patients with non-ASC samples and have pointed to a significant inverse relationship between Eyes Test performance and social deficits in ASC patients (Baron-Cohen, Wheelwright, Hill, et al., 2001; Demurie et al., 2011; Kaland et al., 2008). The present study confirms those findings implying impaired ToM in ASC as revealed by Eyes Test performance.

In both Eyes Test and Films Task, the mean scores of women in the general population were significantly higher than those of men, replicating a gender-specific

difference observed in previous research (Baron-Cohen, 2012; Baron-Cohen et al., 2015; Hoekstra, Bartels, Cath, & Boomsma, 2008; Kirkland, Peterson, Baker, Miller, & Pulos, 2013; Rutherford, Baron-Cohen, & Wheelwright, 2002; Sucksmith et al., 2013; Warrier et al., 2017). Other findings of the current study that have been previously reported include a negative association between Eyes Test scores and AQ scores (Baron-Cohen, Wheelwright, Skinner, et al., 2001; Voracek & Dressler, 2006) and a positive association between the Eyes Test scores and EQ scores (Lawrence et al., 2004; Voracek & Dressler, 2006) The same pattern of association was found with the Films Task scores in the present study.

Some limitations of this study merit mention. Theory of mind is related to intelligence both verbal (Peterson & Miller, 2012) and practical (Baker, Peterson, Pulos, & Kirkland, 2014). However, intelligence was not measured or controlled in the present study and may have had a bearing on our results. In future studies, measures of intelligence and/or language abilities could be incorporated to better discriminate the ASC from non-clinical controls. Secondly, this study included only adults with high-functioning autism, those who voluntarily included themselves on a list of people affected by ASC and were able to complete the tasks on their own, but those ASC individuals with common comorbid psychiatric conditions were excluded. Therefore, we do not know whether the obtained findings can be generalized to subgroups of ASC with intellectual disabilities and those with significant comorbid psychiatric conditions. Thirdly, the sample studied was relatively young, and a wider age range deserves further exploration in future studies.

Further research is needed on the strategies that adults with ASC use to decipher mental and emotional states of others in social situations. The proposed Films Task is based on a compulsory-selective paradigm that reveals the performance of individuals with ASC on ToM tasks fairly well. Developing a compulsory-selective paradigm that can be used online will facilitate the extension of research to include samples with diverse demographics.

References

Aaron, R. V., Benson, T. L., & Park, S. (2015). Investigating the role of alexithymia on the empathic deficits found in schizotypy and autism spectrum traits. *Personality and Individual Differences, 77*, 215-220.

Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and "Reading the Mind in the Eyes". *Intelligence, 44*, 78-92.

Baron-Cohen, S. (2000). Theory of mind and autism: A review. In *International review of research in mental retardation* (Vol. 23, pp. 169-184): Elsevier.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in cognitive sciences, 6*(6), 248-254.

Baron-Cohen, S. (2011). *Zero degrees of empathy: A new theory of human cruelty* (Vol. 30): Penguin UK.

Baron-Cohen, S. (2012). *The science of evil: On empathy and the origins of cruelty*: Basic books.

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.

Baron-Cohen, S., Cassidy, S., Auyeung, B., Allison, C., Achoukhi, M., Robertson, S., ... Lai, M.-C. (2014). Attenuation of typical sex differences in 800 adults with autism vs. 3,900 controls. *PloS one, 9*(7), e102251.

Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. (2004). Mind reading emotions library. *Computer software]. London, England: Jessica Kingsley Publishers*.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high–functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1430), 361-374.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders, 34*(2), 163-175.

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. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241-251.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders, 31*(1), 5-17.

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology, 7*(1), 104-122.

Blair, R. J., & Viding, E. (2008). Psychopathy. *Rutter's Child and Adolescent Psychiatry, Fifth Edition*, 852-863.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839-1849.

Chakrabarti, B., & Baron-Cohen, S. (2006). Empathizing: neurocognitive developmental mechanisms and individual differences. *Progress in brain research, 156*, 403-417.

Coricelli, G. (2005). Two-levels of mental states attribution: from automaticity to voluntariness. *Neuropsychologia, 43*(2), 294-300.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*: ERIC.

Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The scientific World journal, 6*, 1146-1163.

Demurie, E., De Corel, M., & Roeyers, H. (2011). Empathic accuracy in adolescents with autism spectrum disorders and adolescents with attention-deficit/hyperactivity disorder. *Research in Autism Spectrum Disorders, 5*(1), 126-134.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of autism and developmental disorders, 36*(5), 623-636.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of autism and developmental disorders, 38*(3), 464-473.

Fortier, J., Besnard, J., & Allain, P. (2018). Theory of mind, empathy and emotion perception in cortical and subcortical neurodegenerative diseases. *Revue neurologique*.

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The "reading the mind in films" task: complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience, 1*(2), 111-123.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. (2007). The 'reading the mind in the voice'test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of autism and developmental disorders, 37*(6), 1096-1106.

Grove, R., Baillie, A., Allison, C., Baron-Cohen, S., & Hoekstra, R. A. (2014). The latent structure of cognitive and emotional empathy in individuals with autism, first-degree relatives and typical individuals. *Molecular autism, 5*(1), 42.

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological psychiatry, 67*(7), 692-694.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and developmental disorders, 24*(2), 129-154.

Heavey, L., Phillips, W., Baron-Cohen, S., & Rutter, M. (2000). The Awkward Moments Test: A naturalistic measure of social understanding in autism. *Journal of autism and developmental disorders, 30*(3), 225-236.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *Journal of autism and developmental disorders, 38*(8), 1555-1566.

Jones, A. P., Happé, F. G., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 51*(11), 1188-1197.

Kaland, N., Callesen, K., Møller-Nielsen, A., Mortensen, E. L., & Smith, L. (2008). Performance of children and adolescents with Asperger syndrome or high-functioning autism on advanced theory of mind tasks. *Journal of autism and developmental disorders, 38*(6), 1112-1123.

Khorashad, B. S., Baron-Cohen, S., Roshan, G. M., Kazemian, M., Khazai, L., Aghili, Z., ... Afkhamizadeh, M. (2015). The "Reading the Mind in the Eyes" test: investigation of psychometric properties and test–retest reliability of the persian version. *Journal of autism and developmental disorders, 45*(9), 2651-2666.

Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., & Pulos, S. (2013). Meta-analysis Reveals Adult Female Superiority in" Reading the Mind in the Eyes Test". *North American Journal of Psychology, 15*(1), 121.

Koehne, S., Hatri, A., Cacioppo, J. T., & Dziobek, I. (2016). Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. *Cognition, 146*, 8-15.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological medicine, 34*(5), 911-920.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One, 2*(9), e883.

Lombardo, M. V., & Baron-Cohen, S. (2011). The role of the self in mindblindness in autism. *Consciousness and cognition, 20*(1), 130-140.

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. *Development and psychopathology, 26*(4pt1), 933-945.

Mahalanobis, P. C. (1930). On test and measures of group divergence, Part I: Theoretical formulae.

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 35*(6), 655-668.

Mazza, M., Pino, M. C., Mariano, M., Tempesta, D., Ferrara, M., De Berardis, D., ... Valenti, M. (2014). Affective and cognitive empathy in adolescents with autism spectrum disorder. *Frontiers in human neuroscience, 8*, 791.

Murray, K., Johnston, K., Cunnane, H., Kerr, C., Spain, D., Gillan, N., ... Happé, F. (2017). A new test of advanced theory of mind: The "Strange Stories Film

Task" captures social processing differences in adults with autism spectrum disorders. *Autism Research, 10*(6), 1120-1132.

Nejati, V., Zabihzadeh, A., Maleki, G., & Tehranchi, A. (2012). Mind reading and mindfulness deficits in patients with major depression disorder. *Procedia-Social and Behavioral Sciences, 32*, 431-437.

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological bulletin, 133*(2), 310.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Frontiers in psychology, 6*, 1503.

Oshima, F., Iwasa, K., Nishinaka, H., & Shimizu, E. (2015). Early maladaptative schemas and autism spectrum disorders in adults. Journal of Evidence-Based Psychotherapies, 15(2), 191-205.

Peterson, E., & Miller, S. (2012). The eyes test as a measure of individual differences: how much of the variance reflects verbal IQ? *Frontiers in psychology, 3*, 220.

Petroni, A., Canales-Johnson, A., Urquina, H., Guex, R., Hurtado, E., Blenkmann, A., ... Ibanez, A. (2011). The cortical processing of facial emotional expression is associated with social cognition skills and executive functioning: a preliminary study. *Neuroscience letters, 505*(1), 41-46.

Pinkham, A. E. (2014). Social cognition in schizophrenia. *The Journal of clinical psychiatry, 75*, 14-19.

Prevost, M., Carrier, M.-E., Chowne, G., Zelkowitz, P., Joseph, L., & Gold, I. (2014). The Reading the Mind in the Eyes test: validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cognitive neuropsychiatry, 19*(3), 189-204.

Rogers, K., Dziobek, I., Hassenstab, J., Wolf, O. T., & Convit, A. (2007). Who cares? Revisiting empathy in Asperger syndrome. *Journal of autism and developmental disorders, 37*(4), 709-715.

Rueda, P., Fernández-Berrocal, P., & Baron-Cohen, S. (2015). Dissociation between cognitive and affective empathy in youth with Asperger Syndrome. *European Journal of Developmental Psychology, 12*(1), 85-98.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of autism and developmental disorders, 32*(3), 189-194.

Ruzich, E., Allison, C., Smith, P. *et al.* Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism* **6,** 2 (2015). https://doi.org/10.1186/2040-2392-6-2

Shamay-Tsoory, S., Tomer, R., Yaniv, S., & Aharon-Peretz, J. (2002). Empathy deficits in Asperger syndrome: A cognitive profile. *Neurocase, 8*(2), 245-252.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia, 45*(13), 3054-3067.

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: the role of the right ventromedial prefrontal cortex. *Journal of cognitive neuroscience, 15*(3), 324-337.

Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record, 56*(1), 3-21.

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia, 51*(1), 98-105.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: systematic review of psychometric properties and a validation study in Italy. *Cognitive neuropsychiatry, 18*(4), 326-354.

Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D: 4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41*(8), 1481-1491.

Warrier, V., Grasby, K., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., ... Hottenga, J. (2017). Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Molecular psychiatry*.

Williams, J. H., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience & Biobehavioral Reviews, 25*(4), 287-295.

Wilson, C. E., Happé, F., Wheelwright, S. J., Ecker, C., Lombardo, M. V., Johnston, P., ... Lai, M. C. (2014). The neuropsychology of male adults with high-functioning autism or Asperger syndrome. *Autism Research, 7*(5), 568-581.

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.