

Implicit and explicit Theory of Mind reasoning in autism spectrum disorders: The impact of experience

Autism
2015, Vol. 19(4) 459–468
© The Author(s) 2014
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1362361314526004
aut.sagepub.com


Tobias Schuwerk, Maria Vuori and Beate Sodian

Abstract

This study aimed to investigate the relationship between explicit and implicit forms of Theory of Mind reasoning and to test the influence of experience on implicit Theory of Mind reasoning in individuals with autism spectrum disorders and in neurotypical adults. Results from two standard explicit Theory of Mind tasks are mixed: Individuals with autism spectrum disorders did not differ from neurotypical adults in their performance in the Strange Stories Test, but scored significantly lower on the Reading the Mind in the Eyes Test. Furthermore, in an implicit false-belief task, individuals with autism spectrum disorders differed from neurotypical adults in false belief–congruent anticipatory looking. However, this group difference disappeared by (1) providing participants with the outcome of a false belief–based action and (2) subsequently repeating this test trial. Although the tendency to fixate the false belief–congruent location significantly increased from the first to the second test trial in individuals with autism spectrum disorders, it differed in neither test trial from chance. These findings support the notion of an implicit Theory of Mind deficit in autism spectrum disorders, but give rise to the idea that anticipatory looking behaviors in autism spectrum disorders may be affected by experience. Additionally, the pattern of results from implicit and explicit Theory of Mind measures supports the theory of two independent Theory of Mind reasoning systems.

Keywords

anticipatory looking, autism spectrum disorders, Theory of Mind

Introduction

Theory of Mind (ToM), the ability to attribute mental states to others and oneself, is suggested to work in two ways (Apperly and Butterfill, 2009; Frith and Frith, 2008, 2012; Sodian, 2011): (1) An early developing implicit form of ToM reasoning allows for correct action anticipation, based on the ability to take an agent's mental state (e.g. a false belief) into account, without deliberately reflecting on this mental state. It acts fast, unconsciously and inflexibly. Gaze measures, such as anticipatory looking, have been employed as an implicit measure to assess this form of ToM reasoning (e.g. Clements and Perner, 1994; Southgate et al., 2007). (2) Explicit ToM reasoning, developing at around 4 years of age, allows for a deliberate consideration (e.g. a judgment) of others' mental states (Wellman et al., 2001; cf. Perner and Roessler, 2012). This type of ToM reasoning operates consciously, flexibly, and is cognitively demanding. Explicit tasks, such as the false belief task (Wimmer and Perner, 1983), are used to test this

form of ToM reasoning by directly asking for another's mental state.

Autism spectrum disorders (ASD) are characterized by social-cognitive deficits (see Frith, 2012). Moreover, individuals with ASD show a discrepancy between performance in implicit and explicit ToM tasks. Those of higher verbal ability can pass explicit ToM tasks (Bowler, 1992; Happé, 1995), presumably due to compensatory strategies (Frith, 2012; Senju, 2012). In contrast, implicit measures, such as anticipatory looking, indicate a persisting deficit in spontaneous, implicit ToM reasoning (Senju et al., 2009, 2010).

Ludwig-Maximilians-University Munich, Germany

Corresponding author:

Tobias Schuwerk, Department of Psychology, Ludwig-Maximilians-University Munich, Leopoldstr. 13, 80802 Munich, Germany.
Email: tobias.schuwerk@psy.lmu.de

Moreover, recent findings of studies employing gaze measures indicate that the impairment in spontaneous social cognition in individuals with ASD may be specific to false beliefs: Whereas spontaneous action anticipation based on another's false belief seems to be impaired in individuals with ASD (Senju et al., 2009, 2010), other implicit forms of social cognition, such as spontaneous agency perception and spontaneous visual perspective taking, may not be affected (Zwicker et al., 2011).

These findings raise the question of the pervasiveness of the implicit ToM deficit in ASD. In this study, we investigated whether and to what extent learning can modulate performance in implicit ToM tasks (cf. Senju, 2012)? The aim of this study was twofold:

1. To provide further evidence for the discrepancy between explicit and implicit ToM reasoning in ASD, we compared performance in a false belief-based anticipatory looking task (adapted from Senju et al., 2009; cf. Southgate et al., 2007) with performance in two standard ToM tasks: the Reading the Mind in the Eyes (RME) Test (Baron-Cohen et al., 2001), tapping the social-perceptual component of a ToM (cf. Grossman and Tager-Flusberg, 2008; Tager-Flusberg, 2007), and the Strange Stories Test (SST, cf. Happé, 1994), tapping advanced, explicit knowledge about mental states and social situations. Investigating the relationship between these measures helps to better understand different facets of ToM reasoning, their specific impairment in ASD, and can inform recently proposed two-systems accounts of ToM reasoning.
2. We further aimed at testing potential effects of experience on anticipatory looking behavior, assessed in an implicit ToM task. To this end, we changed two aspects of the implicit ToM task, previously employed by Senju et al. (2009): First, in contrast to Senju et al., we presented the false belief-based action and its outcome to the participants, to provide them with information about the contingency between the character's knowledge about prior events and her subsequent action. Second, we repeated this test trial to check for resulting changes in gaze behavior.

In the first test trial, consistent with previous findings (Senju et al., 2009), we expected to find a lack of spontaneous false belief-based action anticipation in the ASD group, compared to neurotypical (NT) adults. In the second test trial, we expected a persistent lack of spontaneous false belief-congruent action anticipation if individuals with ASD cannot learn perception-action contingencies (e.g. the agent will search where she last saw the object) even when presented with action outcomes. Note, however, that such learning can occur in a specific task based

on behavioral cues, without an awareness of the content of the agent's belief. If individuals with ASD are able to learn situation-specific cues that are relevant to belief-based action prediction, then gaze patterns should be affected by the repeated presentation.

We were further interested in which factors might drive a potential modulation of task performance. Therefore, we analyzed the attentiveness to social stimuli (the character's face) in both test trials. Prior research reported comparable attention to social stimuli in individuals with and without ASD (Fletcher-Watson et al., 2009; Rutherford and Towns, 2008). It could be that in the ASD group, any change in gaze behavior between the trials might be attributable to increased attention toward the character's face in the second compared to the first test trial. Additionally, we correlated inter-individual variability in performance on the implicit ToM task with measures of verbal and nonverbal IQ, autistic traits, and performance on explicit ToM tasks. As especially individuals with high-functioning autism or Asperger's syndrome show compensatory learning in explicit ToM tasks (Bowler, 1992; Happé, 1995; cf. Scheeren et al., 2013), performance on these measures might be associated also with changes in gaze behavior between the first and second trial in the current implicit ToM task.

Methods

Participants

A total of 18 adults with ASD (mean age = 24.1 years, SD = 7.0 years; six females) took part in this study. They had been diagnosed by a qualified clinical psychologist or psychiatrist and met International Classification of Diseases–10th Revision (ICD)-10 criteria for high-functioning autism ($n = 3$), Asperger's syndrome ($n = 13$), or atypical autism ($n = 2$). To support group assignment, all participants completed the German short version of the Autism-Spectrum Quotient (AQ-k; cut-off criterion: score ≥ 17 ; Freitag et al., 2007; adapted from Baron-Cohen et al., 2001), a brief self-assessment questionnaire examining autistic traits. Within the ASD group, the average AQ-k score was 18.8 (SD = 8.1). Eleven additional participants with ASD had to be excluded from the data analysis due to missing data in the critical test trials of the eye tracking task ($n = 7$) or not fulfilling inclusion criteria for familiarization trials in the eye tracking task ($n = 4$; for a detailed description, see section "Procedure and apparatus").

The control group of 19 NT adults (mean age = 25.3 years, SD = 3.8 years; six females) was matched by chronological age, verbal, and nonverbal intelligence (Table 1). The AQ-k was administered to support control group assignment. The average AQ-k score in the NT group was 6.0 (SD = 2.8). Seven additional participants from the control group also had to be excluded due to missing data in the critical test trials of the eye tracking

Table 1. Mean test scores for each group and t-test results for group comparisons.

	ASD ^a		NT ^a		Group difference		
	M	SD	M	SD		p value	Effect size
Chronological age	24.1	7.0	25.3	3.8	$t(35) = -0.69$	$p = 0.496$	Cohen's $d = -0.21$
AQ-k	18.8	8.1	6.0	2.8	$t(35) = 6.50$	$p < 0.001$	Cohen's $d = 2.11$
Verbal IQ (MWT-B ^b)	104.3	18.0	102.5	11.9	$t(35) = 0.35$	$p = 0.728$	Cohen's $d = 0.12$
Nonverbal IQ (CFT-20-R ^c)	91.4	21.7	98.3	18.3	$t(35) = -1.04$	$p = 0.304$	Cohen's $d = -0.34$
RME	19.8	5.9	25.5	3.2	$t(33) = -3.54$	$p = 0.001$	Cohen's $d = -1.20$
SST—mental ^d	68.8	16.3	72.9	11.1	$F(1, 32) = 0.19$	$p = 0.667$	$\eta_p^2 = 0.006$
SST—control ^d	84.6	20.5	86.2	16.6			

ASD: autism spectrum disorder group; NT: neurotypical group; M: mean; SD: standard deviation; AQ-k: Autism-Spectrum Quotient; RME: Reading the Mind in the Eyes Test; SST: Strange Stories Test.

Mental: mental condition; control: control condition; significant p values are highlighted in bold.

^aResults are reported for 18 participants with ASD and 19 NT controls. Due to a technical problem, data from one participant with ASD are missing for the SST. Reported results in the SST thus refer to a sample size of 17 participants with ASD. Additionally, two participants from the NT group refused to take part in the RME test, thus results are reported for 17 NT controls.

^bMehrfachwahl-Wortschatz-Intelligenztest (German multiple choice vocabulary test).

^cCulture-Fair Test 20-R.

^dScore: percentage of correct justifications.

task ($n = 5$), not fulfilling inclusion criteria for familiarization trials in the eye tracking task ($n = 2$). A German multiple choice vocabulary test (*Mehrfachwahl-Wortschatz-Intelligenztest* (MWT-B); Lehl, 2005) was employed to assess verbal IQ. The Culture-Fair Test 20-R (CFT-20-R; Weiß, 2006) was used as a measure of non-verbal IQ in terms of general mental capacity. Informed written consent was obtained before testing. The local ethics committee approved the study.

Tasks

RME Test. Participants completed a German paper and pencil version (Bölte, 2005) of the revised RME Test for adults by Baron-Cohen et al. (2001). It consisted of 36 pictures showing the eye region of a face, each expressing a distinguishable mental state. Every picture was accompanied by four words, describing mental states. The most suitable description of what the person is thinking or feeling had to be chosen. The performance score equals the number of correctly chosen mental state terms. The task was conducted without speed limit. Individuals with ASD show weaker performance on this test compared to NT controls (e.g. Baron-Cohen et al., 2001).

SST. This task comprised nine stories adapted from the original SST (Fletcher et al., 1995; Happé, 1994) and a modified German version (Vogeley et al., 2001). The mental condition consisted of five stories that described situations in which the protagonists' behavior can be explained by referring to their mental states. The stories contained a double bluff, a white lie, deception, misunderstanding (cf. Happé, 1994), and emotions (cf. Vogeley et al., 2001). They were presented verbally and in text form. Participants were

asked to justify the story protagonists' behavior in terms of their non-uttered mental states. In the control condition, four stories required inferences on causal physical relations (cf. Fletcher et al., 1995). All stories were controlled for number of words and semantic and syntactic complexity. Performance differences of individuals with ASD compared to NT controls were previously reported in the mental condition but not in the control condition (Jolliffe and Baron-Cohen, 1999; Kaland et al., 2005).

Implicit false belief task (false belief-based anticipatory looking task). This eye tracking version of the false belief task (cf. Southgate et al., 2007; Thøermer et al., 2012) was designed to tap false belief-based anticipatory looking (see Figure 1). The task consisted of four animated video clips: two familiarization trials (each lasting for 32 s) and two test trials (each lasting for 41 s). The laterality of events was counterbalanced. The familiarization trials depicted a character who watched a toy car moving self-propelled from one box to another. After that, the character disappeared behind a screen and two doors on this screen, each above one of the two boxes, were illuminated, accompanied by a chime. The last frame of this sequence was frozen for 3 s and served as anticipatory period to record eye movements. Subsequently, the door above the box in which the character had seen the car disappear opened and the character reached through it to retrieve the car. The familiarization trials demonstrated that the character's goal was to get the car by opening one of the two doors and that this action was preceded by the illumination of the doors and the chime.

The crucial difference between the test trials and the familiarization was that in the test trials a phone ring distracted the character. She did not witness the following

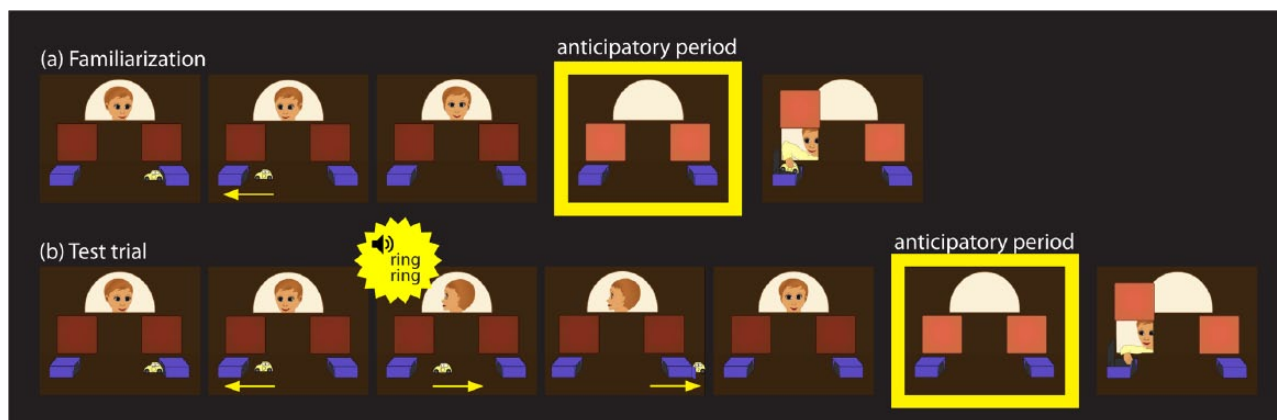


Figure 1. Implicit false belief task: still frames from (a) one familiarization trial and (b) one test trial. The bordered picture was displayed for 3 s and served as anticipatory period for recording of gaze data. The test trial ended with the presentation of the false belief-based action.

events: After arriving at the second box, the car went back into the first box, continued driving through it, and left the scene. By letting the car leave the scene, we aimed to avoid a possible “curse of knowledge” (Birch and Bloom, 2004). Subsequently, as in the familiarization, the character disappeared behind the screen, the doors were illuminated, the chime sounded, and the scene was frozen for 3 s to record fixation data. After that, the animation ended with the presentation of the false belief-based action and its outcome: the character opened the box, in which she believed the car was located, without finding it.

Procedure and apparatus

Participants attended the implicit false belief task first, followed by the RME Test, the SST, and the control measures (AQ-k, MWT-B, and CFT-20-R). For the assessment of eye movements, a Tobii T60 eye tracker (60 Hz sampling rate; Tobii Technology, Stockholm, Sweden) with an integrated 17-inch thin-film transistor (TFT) monitor (1280 × 1024 pixel) was used. Stimuli were presented with Tobii Studio 2.2 software (Tobii Technology). Participants were seated on a chair with a distance of approximately 60 cm away from the screen. A 5-point calibration procedure preceded the stimulus presentation.

Data analysis

SST. Participants’ answers to each story were audiotaped and transcribed. The given justifications were scored following prior described procedures (Happé, 1994; Kaland et al., 2005). In the mental condition, appropriate answers referred to the mental states of the protagonists. In the control condition, justifications that referred to, for example, physical events and outcomes were considered as correct. Answers could be scored as fully correct (2 points), partially correct (1 point), or incorrect (0 points). For each

condition, a total score for response accuracy (percentage of correct justifications) was assessed. A second rater independently judged a subset (35%) of randomly chosen answers (ASD: $n = 6$; NT: $n = 6$). Strong agreement was reached between the two raters (unadjusted intraclass correlation coefficients for the two conditions: SST-control = 0.93, SST-mental = 0.96).

Implicit false belief task. Participant’s eye movements were recorded with Tobii Studio 2.2 software. Further analysis of raw data was conducted using customized scripts. Fixations were defined by a velocity-based fixation filter (Salvucci and Goldberg, 2000) with a velocity threshold of 0.05°/ms. An additional temporal threshold was set to exclude fixations that lasted less than 80 ms.

Fixations on the doors during the 3-s anticipatory period were of particular interest, and so, the two doors served as areas of interest (AOIs; approximately 2.8° × 2.8° in width and height) for data analysis. The door the character opened after the anticipatory period is referred to as “correct door,” whereas the other door is defined as “incorrect door.” A differential looking score (DLS) as described by Senju et al. (2009) was calculated. The total duration of fixations on the incorrect door was subtracted from the total duration of fixations on the correct door, divided by the sum of the total duration of fixations on the correct and incorrect door. The DLS ranges from 1 (fixating the correct door most of the time) to −1 (fixating the incorrect door most of the time). A value around 0 indicates no preference for one of the two doors. Only participants who looked longer to the correct door than to the incorrect door in at least one of the two familiarization trials were included in the further analysis. Four individuals with ASD and two NT controls did not show this belief-congruent anticipatory looking behavior and were excluded. We further analyzed first looks in the anticipatory period. We coded whether the first saccade after the

illumination of the doors was directed toward the correct or incorrect door.

To check whether any change in looking behavior of the ASD group between the test trials would be related to verbal and nonverbal IQ, autistic traits, or to performance in explicit ToM tasks, we calculated a difference score (DLS in the second test trial minus DLS in the first test trial; first look in the second test trial minus first look in the first test trial) and correlated it with verbal and nonverbal IQ, as well as with performance on the AQ-k, SST, and RME Test. To further explore the attentiveness to social stimuli in the ASD and the NT group in the scene prior to the critical test period, we analyzed total fixation durations within an AOI that covered the character's face.

Results

All statistical analyses were conducted using IBM SPSS Statistics 20 (SPSS Inc., Chicago, IL, USA). The ASD and NT group did not differ in chronological age ($t(35) = -0.69, p = 0.496$, Cohen's $d = -0.21$), verbal IQ ($t(35) = 0.35, p = 0.728$, Cohen's $d = 0.12$), or nonverbal IQ ($t(35) = -1.04, p = 0.304$, Cohen's $d = -0.34$). AQ-k mean scores differed significantly between groups ($t(35) = 6.50, p < 0.001$, Cohen's $d = 2.11$), scores of the ASD group were significantly higher than that in the NT group (see Table 1).

Standard ToM tests

Two participants from the NT group refused to take part in the RME Test, thus results are reported for 17 NT controls. Additionally, due to a technical problem, the audio recording from one participant with ASD was missing in the SST. Reported results in this task thus refer to a sample size of 17 in the ASD group. Although these participants' data were missing, the two groups remained matched for the control measures. The NT group's average score on the RME Test was 25.5 (SD = 3.2). With a mean score of 19.8 (SD = 5.9), the ASD group showed significantly weaker performance in this test ($t(33) = -3.54, p = 0.001$, Cohen's $d = -1.20$). The SST scores represent percentage of correct justifications. An analysis of variance (ANOVA) with the factors group (ASD vs NT) and condition (mental vs control) showed a significant effect of condition, $F(1, 32) = 18.85, p < 0.001, \eta_p^2 = 0.38$. Accuracy rates for justifications given in the mental condition ($M = 70.8, SD = 14.0$) were lower than those given in the control condition ($M = 84.6, SD = 18.5$). The ASD group ($M = 76.7, SD = 16.5$) and the NT group ($M = 78.8, SD = 10.5$) did not differ in their performance on the SST, $F(1, 32) = 0.19, p = 0.667, \eta_p^2 = 0.006$. Also no significant interaction between group and condition was observed, $F(1, 32) = 0.43, p = 0.518, \eta_p^2 = 0.01$.

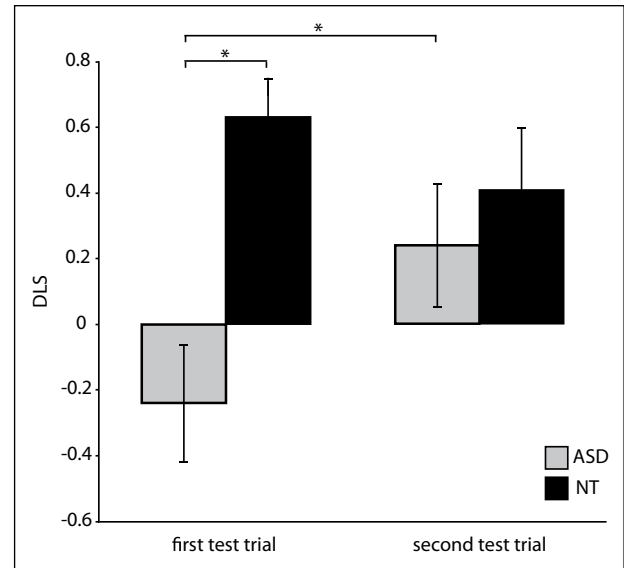


Figure 2. Mean differential looking scores (DLS; \pm SEM) from the autism spectrum disorders (ASD) group and the neurotypical (NT) control group in the implicit false belief task. Whereas in the first test trial, the ASD group showed a significantly reduced looking bias toward the false belief-congruent location compared to the NT group, this difference disappeared in the presentation of a second test trial. In the ASD group, the tendency to fixate the false belief-congruent location increased significantly, however, without leading to above chance performance in the second test trial. Symbols indicate significance level.

SEM: standard error of the mean.

* $p < 0.05$.

Implicit false belief task

In preliminary analyses, no effects of participant's sex or laterality of events (car starting either from left or right of the scene) were found, so these factors were disregarded in the following analyses.

DLS. DLS was analyzed using a 2×2 repeated-measures ANOVA. Mean scores are shown in Figure 2. The between-participants variable was group (ASD vs NT), and the within participant variable was test trial (first vs second). The ANOVA revealed a significant main effect for group, $F(1, 35) = 7.69, p = 0.009, \eta_p^2 = 0.18$. The NT group showed a greater looking bias toward the false belief-congruent location ($M = 0.52, SD = 0.62$) than the ASD group ($M = -0.004, SD = 0.53$). No effect of test trial, $F(1, 35) = 0.81, p = 0.373, \eta_p^2 = 0.02$, was observed. The interaction between group and test trial was significant, $F(1, 35) = 5.90, p = 0.020, \eta_p^2 = 0.14$. Follow-up t -tests revealed that in the first test trial, the NT group fixated significantly longer the correct location than the ASD group (NT: mean = 0.63, SD = 0.53; ASD: mean = -0.24, SD = 0.74), $t(35) = -4.16, p < 0.001$, Cohen's $d = -1.35$. Whereas the NT group showed a significant looking bias toward the correct

location ($t(18) = 5.26, p < 0.001$, Cohen's $d = 1.21$), the DLS revealed no preference for either location in the ASD group ($t(17) = -1.39, p = 0.182$, Cohen's $d = -0.33$). This pattern of results changed in the second test trial: the NT group and the ASD group did not differ in their looking bias toward the correct location (NT: mean = 0.41, SD = 0.81; ASD: mean = 0.24, SD = 0.79), $t(35) = -0.67, p = 0.506$, Cohen's $d = -0.21$. The NT group looked significantly longer to the correct location ($t(18) = 2.21, p = 0.041$, Cohen's $d = 0.51$). The DLS in the ASD group did not differ from zero ($t(17) = 1.28, p = 0.219$, Cohen's $d = 0.30$). Based on our a priori hypothesis that a looking bias toward the false belief-congruent location of participants with ASD might be affected by the repeated presentation of the test trial, we performed a one-tailed follow-up t -test within the ASD group between the DLS in the first and the second trial. A significant increase in looking toward the correct location from the first to the second trial was observed, $t(17) = -1.85, p = 0.041$, Cohen's $d = -0.63$.

First look. In the first test trial, 14 out of 19 participants of the NT group directed their first saccade toward the correct location (marginally significantly above chance, $p = 0.064$, binomial test). In the ASD group, only 6 out of 18 participants made their first fixation on the correct location, which did not differ from chance ($p = 0.238$, binomial test). The NT and ASD group significantly differed in the direction of the first look in the first test trial ($p = 0.022$, Fisher's exact test, two-tailed). This changed in the second test trial: 11 out of 19 NT adults focused their first look at the correct location (no difference from chance, $p = 0.648$, binomial test). In the ASD group, now 8 out of 18 individuals directed their first saccade toward the correct location, which still did not differ from chance ($p = 0.82$, binomial test). This change in directions of first looks led to an absence of a group difference in the second test trial, $p = 0.517$, Fisher's exact test, two-tailed. Comparison of first looks between the first and second trial showed no significant performance difference for the ASD group ($p = 0.377$, McNemar's Test, one-tailed). Figure 3 displays the ratio of correct first look for each group and test trial.

Fixation duration on character's face. A 2 (group: ASD vs NT) \times 2 (test trial: first vs second) repeated-measures ANOVA of the total duration of fixations on the character's face prior to the anticipatory period revealed a significant main effect for group, $F(1, 35) = 4.96, p = 0.032, \eta_p^2 = 0.12$. Overall, the NT group looked longer to the character's face in the prior scene ($M = 11.4, SD = 3.0$) than the ASD group ($M = 8.6, SD = 3.8$). Neither a significant main effect for test trial, $F(1, 35) = 0.03, p = 0.861, \eta_p^2 = 0.00$, nor a significant interaction between group and test trial, $F(1, 35) = 0.21, p = 0.647, \eta_p^2 = 0.01$, was observed. Because of the specific hypothesis that a change in attentiveness to the protagonist's face could explain

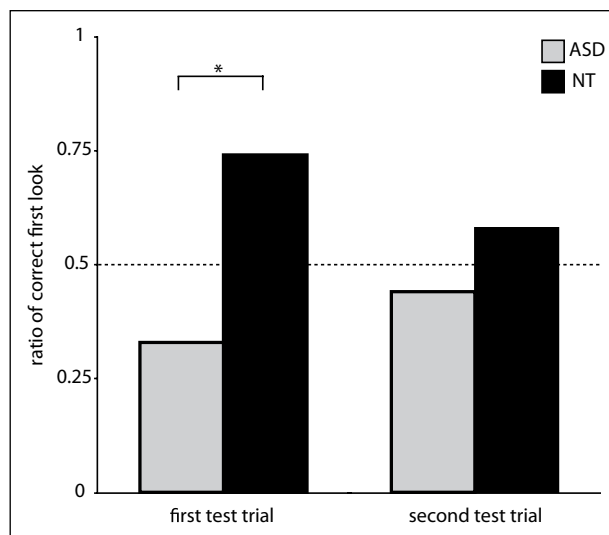


Figure 3. Ratio of correct first looks of individuals with autism spectrum disorders (ASD) and of neurotypical controls (NT). In the first test trial, a significant group difference was observed, which disappeared in the repeated presentation of this trial. Symbols indicate significance level.

* $p < 0.05$.

different anticipatory looking patterns between the first and the second test trial, we performed follow-up t -tests, which revealed no significant difference of fixations on the protagonist's face between the first and second test trial, neither in the ASD group, $t(17) = 0.45, p = 0.659$, Cohen's $d = 0.11$, nor in the NT group, $t(18) = -0.20, p = 0.841$, Cohen's $d = -0.06$.

Analysis of within-group variability. No significant correlations between change of the DLS or first looks between the first and second test trial and verbal IQ, nonverbal IQ, performance in the AQ-k, SST, or RME Test were found in the ASD group (all $r_s = (-0.335; 0.227)$, $p \geq 0.19$). To assess the relationship among performance on implicit and explicit ToM tasks, as well as the other control measures, we correlated the DLS and first looks of the first test trial (as a measure of implicit ToM reasoning) with performance on verbal IQ, nonverbal IQ, performance in the AQ-k, the SST, and the RME Test, separated for each group. No significant correlations between the DLS in the first test trial and the above listed measures were observed either in the ASD group (all $r_s = (0.007; 0.330)$, $p \geq 0.18$), or the NT group (all $r_s = (-0.290; 0.352)$, $p \geq 0.14$). In the ASD group, no significant point-biserial correlations among first looks toward the false belief-congruent location in the first test trial and the verbal and nonverbal IQ, AQ-k, the SST, and the RME Test were found (all $r_{pb}s = (-0.015; 0.390)$, $p \geq 0.11$). For the NT group, the point-biserial correlations among false belief-congruent first looks and the RME Test were significant, $r_{pb} = 0.537, p = 0.03$. Furthermore, the point-biserial

correlations among false belief–congruent first looks and the AQ-k score were significant, $r_{pb} = 0.491$, $p = 0.03$. Due to the small variance of the AQ-k score in NT samples, we refrain from an interpretation of this finding. All other point-biserial correlations were not significant, $r_{pbs} = (-0.357; -0.279)$, $p \geq 0.13$.

Discussion

The performance of individuals with ASD and NT adults in the current set of implicit and explicit measures of ToM reasoning showed that individuals with ASD lack a spontaneous sensitivity to another's false belief in an implicit ToM task. This replicates findings from Senju et al. (2009) by using a different paradigm. However, anticipatory looking no longer differed from NT adults in the single repetition of the test trial, giving rise to the idea that individuals with ASD might be able modify their performance in an implicit ToM task based on experience. Furthermore, individuals with ASD exhibited equal performance in the SST, but weaker performance in the RME Test. This pattern of results supports the notion of two ToM reasoning systems: an explicit system to verbally explain behavior based on mental states, and an implicit system for spontaneous sensitivity to other's mental states.

Does experience affect performance in the implicit false belief task?

In the first test trial of the present implicit ToM task, gaze data revealed a lack of spontaneous false belief–based action anticipation in individuals with ASD. Whereas gaze patterns of NT adults indicated an anticipation of the false belief–based behavior of the story's protagonist, individuals with ASD did not show a looking bias toward this false belief–congruent location. Furthermore, the groups differed significantly in their looking bias.

However, in the repeated presentation of the test trial, individuals with ASD no longer differed from NT adults in their looking bias toward the false belief–congruent location. Furthermore, although within the ASD group the looking bias did not increase above chance level, the tendency to fixate the false belief–congruent location increased significantly from the first to the second test trial. Showing the false belief–based action and its outcome (failure in finding the car) and presenting the test trial one more time were sufficient to alternate the pattern of results observed in the first test trial and reported in the study by Senju et al. (2009). Based on the finding that no group difference between participants with ASD and controls was observed in the second trial, we argue that experience may modulate anticipatory looking of participants with ASD in an implicit ToM task. The increased tendency to fixate the false belief–congruent location in the second

test trial (although not resulting in above chance level performance) also points in that direction.

Intriguingly, a recent elegant study by Schneider et al. (2013), which addressed learning processes in a similar implicit ToM task, reported apparently contradictory results: Just like in the current study, Schneider et al. employed an adaptation of Senju et al.'s (2009) task to test the effect of the repeated presentation of test trials on belief–congruent anticipatory looking in individuals with and without ASD. Individuals with ASD showed no looking bias toward the location, where an actress falsely believed an object would be located, and in contrast to the current study, gaze behavior was not affected by the repeated presentation of the test trials, indicated by a lasting group effect between participants with ASD and NT controls over several trials. This led the authors to conclude that individuals with ASD have a sustained deficit of implicit ToM reasoning and do not show spontaneous learning in implicit ToM tasks. However, Schneider et al.'s (as well as Senju et al.'s) implicit ToM task differs in one crucial aspect from the current paradigm: The participants were not presented with a belief–corresponding action of the protagonist and its outcome. It is possible that not providing perception–action contingencies, that is, the false belief–based searching for the object in the empty box, prevented a modulation of gaze behavior in individuals with ASD. The results of the current study suggest that the single observation of the false belief–based action (he opens the box where he falsely believes the object is located) and its outcome (he does not find the object in there) might be a source for experience-based modulation of belief–congruent anticipatory looking.

In the current study, no relationship between interindividual differences within the ASD group in looking behavior and IQ, autistic traits, or performance in the standard ToM tasks was found. Furthermore, we checked whether increased selective attention to social stimuli, in particular to the character's face, prior to the anticipatory phase in the second trial, might explain the pattern of results of the two test trials in the ASD group. This was not the case: attentiveness to the protagonist's face did not change between test trials. Thus, increased consideration of social stimuli as a compensatory mechanism cannot explain the current results. We propose that individuals with ASD might benefit from experience in the current implicit ToM task, but not specifically by increased attention to social stimuli.

The nature of such an adaptive mechanism remains to be investigated. Ruffman et al. (2001) suggested that eye gaze in an implicit ToM task indicates spontaneous core insights into social behavior. Alternatively, it can be argued that the present results indicate general behavioral learning, rather than reflecting spontaneous processes, specifically dedicated to social understanding. The observed outcome in the first test trial (actor opens false

belief-congruent door) could have caused learning from effect. Knowledge about the order of events in the first test trial might result in improved anticipation of the false belief-congruent action in the second test trial, however, without the need to actually consider the actor's belief. Note, however, that in order to learn from the sequence of events presented in the present task, participants had to minimally form an association between the agent's looking behavior and subsequent action. Individuals with ASD might be sensitive to behavioral cues that are relevant for action prediction.

It is important to note that although the looking bias toward the false belief-congruent location, assessed by means of the DLS, increased from the first to the second test trial in the ASD group, the DLS differed in neither test trial from chance. Thus, it cannot be concluded that experience has led to a false belief-congruent anticipatory looking bias (which would have required above chance task performance in the second test trial). The same holds also for the first looks, which differed in neither test trial from chance in the ASD group. Thus, at this point, it can only be speculated that the observed pattern of results show that anticipatory looking might be affected by experience.

To what extent deficits in spontaneous ToM reasoning can be modified by experience needs further research. In a future study, multiple trials showing false and true belief-based actions and their outcomes could be presented. This would make it possible to test if and when experience with this type of implicit ToM task is sufficient to elicit belief-congruent anticipatory looking. Such findings would be crucial to better understand if and how individuals with ASD can make use of experience to compensate deficits in implicit ToM reasoning. Such compensatory strategies are well known in explicit ToM (e.g. Bowler, 1992; Happé, 1995), but we currently lack profound evidence for compensatory learning in implicit ToM reasoning tasks.

The current findings may also help to explain apparently discrepant findings from eye tracking studies of spontaneous social cognition. While Senju et al. (2009) report a lack of spontaneous ToM reasoning, Zwickel et al. (2011) found no differences in spontaneous agency perception and spontaneous visual perspective taking. Notably, the two studies diverge in the time interval on which eye movement data analysis was based. While Senju et al. coded fixations from a 6-s-period in one test trial, Zwickel et al. extracted fixation data from 18-s-long video clips. Furthermore, Fletcher-Watson et al. (2009) reported deviant eye movement patterns of individuals with ASD only in the first fixations of their trials. The critical test phase of the current implicit ToM task lasted for 3 s. In this short time interval, individuals with ASD revealed deviant gaze patterns compared to controls. However, in the subsequent repetition of this test trial, this difference vanished. Together with prior findings, this suggests that differences in eye movements between individuals with

ASD and NT individuals are primarily apparent in very early fixations, which rely on spontaneous and fast processes.

Relationship between implicit and explicit ToM reasoning

In line with prior findings by Senju et al. (2009), no performance differences between individuals with ASD and NT controls were observed in the SST. Individuals with ASD were as good as the NT adults in providing behavior justifications, which required descriptions of other's mental states. In contrast to performance in this explicit ToM task, as in Senju et al.'s study, the ASD group performed weaker than the NT controls in the first test trial of the implicit ToM task, which supports notions of two distinct ToM reasoning systems, an explicit and an implicit one. The changed pattern of results in the second test trial of the implicit ToM task might indicate the use of explicit knowledge about the relation between an agent's access to information, mental state, and action, to guide action expectations in the repeated presentation of the present implicit ToM task.

Interestingly, also the RME Test revealed a group difference. It may be that the implicit ToM task and the RME Test tap common ToM reasoning competencies. Whereas the former assesses spontaneous sensitivity to a goal-directed action based on mental states, the latter requires decoding of mental states from facial expressions. Although the RME Test asks explicitly to name these mental states, the mere presentation of the eye region of faces may trigger fast and automatic spontaneous reasoning about the person's mental state. The spontaneous appreciation of mental states seems to be affected in ASD, independently of general cognitive functioning, and independently of advanced, explicit ToM reasoning. The finding that implicit ToM task performance (first looks in the first test trial) correlated with the RME Test performance in the NT group, but not in the ASD group, supports this notion: intact ToM reasoning (including both an intact implicit and explicit subcomponent) might underlie successful task performance of NT adults. However, individuals with ASD might adapt their performance task-specifically, resulting in task-specific compensatory strategies that vary in their effectiveness.

The results of this study support the notion of implicit and explicit ToM processing systems which are at least partly independent: Advanced explicit, verbal reasoning about others' mental states is possible in individuals with marked impairments in processing systems that are relevant for a fast and spontaneous appreciation of others' mental states, the processing of emotional faces, and the processing of goal-directed actions based on an agent's false belief. Future research has to address the precise characterization of implicit and explicit forms of social cognition (cf. Frith and Frith, 2012).

Acknowledgements

We thank all participants who took part in this study. We are grateful to Christine Schulz (AutKom München), Peter Josenhans (Maut), and Heike Vogel (BBW, Abensberg) for their support. Thanks are due to Iyad Aldaqre for his assistance in data analysis as well as Stefanie Sabot and Djulia Tucev for their help with data acquisition. We also thank Birgit Träuble and Nike Tsalas for helpful comments on earlier versions of the article. We finally thank our reviewers for their informed and helpful comments.

Funding

This research was funded by a grant from the Volkswagen Foundation (Research group “The social brain”).

References

- Apperly I and Butterfill SA (2009) Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116(4): 953–970.
- Baron-Cohen S, Wheelwright S, Hill J, et al. (2001) The “Reading the Mind in the Eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry* 42(2): 241–251.
- Birch S and Bloom P (2004) Understanding children’s and adults’ limitations in mental state reasoning. *Trends in Cognitive Sciences* 8(6): 255–260.
- Bölte S (2005) *Reading Mind in the Eyes Test für Erwachsene (dt. Fassung) von S. Baron-Cohen* [Reading the mind in the eyes test for adults (German version) by S. Baron-Cohen]. J. W. Goethe Universität Frankfurt/M: Eigendruck.
- Bowler DM (1992) “Theory of Mind” in Asperger’s syndrome. *Journal of Child Psychology and Psychiatry* 33(5): 877–893.
- Clements WA and Perner J (1994) Implicit understanding of belief. *Cognitive Development* 9(4): 377–395.
- Fletcher PC, Happé FGE, Frith U, et al. (1995) Other minds in the brain: a functional imaging study of “Theory of Mind” in story comprehension. *Cognition* 57(2): 109–128.
- Fletcher-Watson S, Leekam SR, Benson V, et al. (2009) Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia* 47(1): 248–257.
- Freitag CM, Retz-Junginger P, Retz W, et al. (2007) Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—die Kurzversion. *Zeitschrift für Klinische Psychologie und Psychotherapie* 36(4): 280–289.
- Frith C and Frith U (2008) Implicit and explicit processes in social cognition. *Neuron* 60(3): 503–510.
- Frith CD and Frith U (2012) Mechanisms of social cognition. *Annual Review of Psychology* 63: 287–313.
- Frith U (2012) Why we need cognitive explanations of autism. *Quarterly Journal of Experimental Psychology* 65(11): 2073–2092.
- Grossman RB and Tager-Flusberg H (2008) Reading faces for information about words and emotions in adolescents with autism. *Research in Autism Spectrum Disorders* 2: 681–695.
- Happé FGE (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.
- Happé FGE (1995) The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development* 66(3): 843–855.
- Jolliffe T and Baron-Cohen S (1999) The Strange Stories Test: a replication with high-functioning adults with autism or Asperger syndrome. *Journal of Autism and Developmental Disorders* 29(5): 395–406.
- Kaland N, Moller-Nielsen A, Smith L, et al. (2005) The Strange Stories Test—a replication study of children and adolescents with Asperger syndrome. *European Child & Adolescent Psychiatry* 14(2): 73–82.
- Lehrl S (2005) *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B*. Balingen: Spitta Verlag, p. 5 (unveränderte Aufl).
- Perner J and Roessler J (2012) From infants’ to children’s appreciation of belief. *Trends in Cognitive Sciences* 16(19): 519–525.
- Ruffman T, Garnham W and Rideout P (2001) Social understanding in autism: eye gaze as a measure of core insights. *Journal of Child Psychology and Psychiatry* 42(8): 1083–1094.
- Rutherford MD and Towns AM (2008) Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders* 38(7): 1371–1381.
- Salvucci DD and Goldberg JH (2000) Identifying fixations and saccades in eye-tracking protocols. In: Proceedings of the 2000 symposium on eye tracking research & applications, Palm Beach Gardens, FL, 6–8 November, pp.71–78. New York: ACM Press.
- Scheeren AM, de Rosnay M, Koot HM, et al. (2013) Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry* 54(6): 628–635.
- Schneider D, Slaughter VP, Bayliss AP, et al. (2013) A temporally sustained implicit theory of mind deficit in autism spectrum disorders. *Cognition* 129(2): 410–417.
- Senju A (2012) Spontaneous theory of mind and its absence in autism spectrum disorders. *Neuroscientist* 18(2): 108–113.
- Senju A, Southgate V, Miura Y, et al. (2010) Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology* 22(2): 353–360.
- Senju A, Southgate V, White S, et al. (2009) Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science* 325(5942): 883–885.
- Sodian B (2011) Theory of mind in infancy. *Child Development Perspectives* 5(1): 39–43.
- Southgate V, Senju A and Csibra G (2007) Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18(7): 587–592.
- Tager-Flusberg H (2007) Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science* 16(6): 311–315.
- Thoermer C, Sodian B, Vuori M, et al. (2012) Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology* 30(1): 172–187.

- Vogeley K, Bussfeld P, Newen A, et al. (2001) Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* 14(1): 170–181.
- Weiß RH (2006) *Grundintelligenztest Skala 2—Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Wellman HM, Cross D and Watson J (2001) Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development* 72(3): 655–684.
- Wimmer H and 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.
- Zwicker J, White SJ, Coniston D, et al. (2011) Exploring the building blocks of social cognition: spontaneous agency perception and visual perspective taking in autism. *Social Cognitive and Affective Neuroscience* 6(5): 564–571.