Discriminating autism spectrum disorders from schizophrenia by investigation of mental state attribution on an on-line mentalizing task: A review and meta-analysis

Vibeke Bliksted a,b, Shiho Ubukata c, Katja Koelkebeck d,*  

a Aarhus University Hospital Risskov, Psychosis Research Unit, Slavangervej 2, 8240 Risskov, Denmark  
b Aarhus University, Interactive Minds Centre, Jens Chr. Skoog Vej 4, 8000 Aarhus C, Denmark  
c Department of Neuropsychiatry, Graduate School of Medicine, Kyoto University, 5 Shogoin-Kawaharacho, Kyoto, 606-8507, Japan  
d Department of Psychiatry and Psychotherapy, School of Medicine, University of Muenster, Albert-Schweitzer-Campus 1, Building A9, 48149 Muenster, Germany

A R T I C L E   I N F O

Article history:
Received 2 September 2015  
Received in revised form 4 January 2016  
Accepted 12 January 2016  
Available online xxxx

Keywords:
Theory of mind  
Schizophrenia  
Autism  
ASD  
FMRI  
Differential diagnosis

A B S T R A C T

In recent years, theories of how humans form a “theory of mind” of others (“mentalizing”) have increasingly been called upon to explain impairments in social interaction in mental disorders, such as autism spectrum disorders (ASD) and schizophrenia. However, it remains unclear whether tasks that assess impairments in mentalizing can also contribute to determining differential deficits across disorders, which may be important for early identification and treatment. Paradigms that challenge mentalizing abilities in an on-line, real-life fashion have been considered helpful in detecting disease-specific deficits. In this review, we are therefore summarizing results of studies that assess the attribution of mental states using an animated triangles task. Behavioral as well as brain imaging studies in ASD and schizophrenia have been taken into account. While for neuroimaging methods, data are sparse and investigation methods inconsistent, we performed a meta-analysis of behavioral data to directly investigate performance deficits across disorders. Here, more impaired abilities in the appropriate description of interactions were found in ASD patients than in patients with schizophrenia. Moreover, an analysis of first-episode (FES) versus longer lasting (LLS) schizophrenia showed that usage of mental state terms was reduced in the LLS group. In our review and meta-analysis, we identified performance differences between ASD and schizophrenia that seem helpful in targeting differential deficits, taking into account different stages of schizophrenia. However, to tackle the deficits in more detail, studies are needed that directly compare patients with ASD and schizophrenia using behavioral or neuroimaging methods with more standardized task versions.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The concept of having a “theory of mind” (ToM), i.e. the ability to explain one’s own and the actions of others in terms of beliefs, desires and goals (“mentalizing”) (Blakemore et al., 2003), has been made use of in recent years to account for the development of certain symptomatology in mental disorders. Empathy (Blair, 2005) and mentalizing (Chung et al., 2014) deficits have repeatedly been described in autism spectrum disorders (ASD) and have been identified as core features to explain a lack of reciprocity (Kleinman et al., 2001). In schizophrenia, prominent psychotic symptoms, such as paranoid delusions, ideas of reference (e.g. Frith, 2004) and autistic features (Koelkebeck et al., 2010; Lugnegard et al., 2014), have been directly linked to a lack of ability to mentalize. A solid body of research has shown that patients with ASD (e.g. Frith, 1996) and schizophrenia (e.g. Brüne, 2005) have dysfunctional ToM abilities, which can discriminate these patients from other mental disorders (Murphy, 2006).

Moreover, these two neurodevelopmental disorders, i.e. ASD and schizophrenia, share abnormalities in neural systems that have been identified to form the cerebral “mentalizing network” (Voellm et al., 2006). Within this network, brain regions (such as the medial prefrontal cortex (mPFC) and the superior temporal sulcus (STS)) have been shown to be abnormally activated (Bliksted et al., 2014; Brüne et al., 2008; Castelli et al., 2002; Pedersen et al., 2012) and structurally altered in the disorders (Benedetti et al., 2009; Brieber et al., 2007; Hirao et al., 2008; Koelkebeck et al., 2013; Waiter et al., 2004). In addition, changes in functional and/or structural connectivity have been reported in both patient populations (Das et al., 2012a; Eack et al., 2013; Li et al., 2014). However, it is not yet clear whether abnormal performance on and neural activation to ToM tasks could identify differential deficits (see e.g. Bora et al. (2009) for an account of the specificity of ToM deficits in schizophrenia and bipolar disorder and Chung et al. (2014) for a meta-analysis of findings on different mentalizing tasks in ASD and schizophrenia).

http://dx.doi.org/10.1016/j.schres.2016.01.037
0920-9964/© 2016 Elsevier B.V. All rights reserved.

Please cite this article as: Bliksted, V., et al., Discriminating autism spectrum disorders from schizophrenia by investigation of mental state attribution on an on-line mentalizing..., Schizophr. Res. (2016), http://dx.doi.org/10.1016/j.schres.2016.01.037
As regards the validity of tasks that assess ToM deficits, it has been agreed that some tasks that are valid in children or severely compromised patients with ASD do not sufficiently target subtler deficits in adults with high-functioning ASD (Abell et al., 2000; Mathersul et al., 2013). It has been shown that ASD patients with higher functional levels pass false-belief tasks more easily than those with stronger autistic traits (Abell et al., 2000; Castelli et al., 2002). Similar assumptions hold true for patients with schizophrenia (Horan et al., 2009).

A task that provokes mental state attribution through interacting geometrical shapes may be particularly useful to detect deficits in ToM performance in high-functioning ASD patients and patients with schizophrenia. This task, here referred to as the “animated triangles task”, requires an on-line interpretation of social information, providing the task with properties close to real-life, and demands both implicit (earlier, spontaneous and related to biological motion identification) and explicit (reasoning about the mental states of others) mentalizing abilities (Koelkebeck et al., 2013; Wolf et al., 2010). The task consists of a set of short animated movies that depict two triangles. The triangles either move randomly, not interacting (e.g. bouncing off the walls; random movement (RAN)), interact in a goal-directed manner (e.g. fighting; goal-directed movement (GD)), or interact as if they read each other’s minds (e.g. by mocking the other; ToM). After watching each movie, participants are asked to interpret the sequences freely. The answers are scored by criteria that subsume three dimensions: intentionality (usage of mental state terms), appropriateness of descriptions and meta-analysis of behavioral performance data. For neuroimaging data, it was also performed. The reference lists of all included articles were searched for additional publications and research groups that have published in the field were contacted for supplementary data. Articles written in English and published in peer-reviewed journals were included. Moreover, articles under submission and unpublished data the authors had notice of at the time of submission of this manuscript were included. All included studies were case–control studies. Single-case studies were excluded. In total, 51 studies using the animated triangles task were identified (see Fig. 1 for a flow diagram regarding the inclusion/exclusion of studies). The studies comprised investigations of healthy controls as well as clinical samples, using a broad methodology ranging from behavioral over neuroimaging to eye-tracking methods. Eleven studies of patients with ASD and ten studies of patients with schizophrenia as primary diagnosis were found (for a list of all studies see Supplementary material). In order to avoid duplicate study selection, all articles were evaluated and data extracted by the three authors. In cases of disagreement, the authors discussed the matter until consensus was achieved.

![Flow diagram](image-url)
2.2. Meta-analyses

Meta-analyses of data extracted from ten studies that were available to the authors were performed. Of the 21 papers that met the inclusion criteria, 11 data sets could not be obtained or data were duplicated. Nine studies included behavioral performance data of both patients with ASD (N = 98; four studies) and schizophrenia (N = 206; six studies) (one study provided data from both ASD and schizophrenia). The meta-analyses were performed using StataIC 13 (Sterne, 2009). Studies that did not use the scoring system originally proposed by Abell et al. (2000) or Castelli et al. (2002) were excluded. The means and standard deviations for each comparison were used to calculate the effect sizes. The effect sizes were calculated by looking at the differences between intentionality and appropriateness scores on the ToM and RAN animations in each individual study and the differences regarding the ToM animations were subtracted from the mean scores of the ToM animations. The mean scores of the ToM animations were estimated using Hedges’ g (Hedges and Olkin, 1985). In Hedges’ method, the differences in means are divided by an estimate of the standard deviation which is obtained from pooling the standard deviations of the scores from both types of animations (ToM and RAN) (Rosenthal, 1994). Furthermore, a small sample bias correction factor is incorporated in the calculations of Hedges’ g. The mean scores of the ToM animations were subtracted from the mean scores of the RAN animations in each individual study and the differences regarding the intentionality and the appropriateness scores in each patient group were analyzed. It was not possible to check for publication bias due to the small number of studies available.

3. Empirical findings

3.1. Autism spectrum disorders (ASD)

3.1.1. Behavioral findings

Out of 11 studies which used the animated triangles task in ASD, nine included the assessment of behavioral data. Table 1 depicts the characteristics of the selected studies. These 11 studies investigated, in total, 259 patients with ASD (75 adults and 184 children). All of the studies assessed differences in the task performance between individuals with ASD and healthy controls. The behavioral results of the studies using this task in ASD are substantially consistent. Individuals with ASD show difficulties in mentalizing (intentionality) compared with healthy participants in the ToM animations regardless of age (Castelli et al., 2002; Lugnegard et al., 2013; Marsh and Hamilton, 2011; White et al., 2011). These findings have been reported in all studies except two in which no significant differences in the intentionality scores on ToM animations between ASD patients and healthy participants were found (Salter et al., 2008; Zwickel et al., 2011). In each case, only one study found significant differences in intentionality scores between patients and healthy controls on the GD animations (Bal et al., 2013) and RAN animations (Zwickel et al., 2011). Several studies have shown that patients with ASD describe the ToM animations less appropriately than healthy controls (Bal et al., 2013; Castelli et al., 2002; Lugnegard et al., 2013; Salter et al., 2008; White et al., 2011). No differences in appropriateness were reported from the RAN animations (Bal et al., 2013; Castelli et al., 2002; White et al., 2011), but one study found that patients scored lower in appropriateness on the GD animations (Bal et al., 2013). Regarding the length of the participants’ responses, all but two studies (Castelli et al., 2002; Salter et al., 2008) found differences between patients and controls. In summary, on the animated triangles task children and adults with ASD showed a reduced use of mental state terms on the ToM animations in particular and were less able to describe the ToM interactions appropriately.

Two studies did not find significant differences between the groups. One reason might be a higher mean age of patients in Zwickel et al.’s study (2011), probably hinting at better ToM abilities in patients with ASD at a higher age.

3.1.2. Neuroimaging findings

On ASD, two neuroimaging studies using functional MRI and PET have been conducted. A total of 40 adult patients were assessed (for details, see Table 1). All of the patients were diagnosed as having a high-functioning ASD. In Castelli et al.’s (2002) study on patients with ASD, the authors determined abnormalities in the connectivity of brain regions. Investigating the activation patterns contrasting ToM and RAN animations, a weaker connectivity between the STS and the visual stream (extrastriate region/V3) was determined in patients with ASD. Moreover, the authors identified higher activation in the occipital cortex, which is related to motion identification and thus early visual processing (V3, magnocellular system, “What is it?”). In contrast, reduced activation as compared to healthy controls was identified in the STS, a part of posterior ToM regions. Kana et al. (2009) also investigated patients with ASD versus healthy controls. On the ToM versus RAN animations contrast, patients with ASD showed reduced activation in the...
frontal cerebral regions including the medial frontal gyrus, anterior paracingulate cortex, anterior cingulate cortex (ACC) and inferior orbital frontal gyrus (OFG) as compared to healthy controls. Moreover, a reduction in cortical network connectivity in patients with ASD between frontal ToM regions (medial frontal gyrus, anterior paracingulate cortex, OFG) and the posterior ToM regions (right middle and STS) was found during mentalizing. It is worth of note that patients with higher scores on a behavioral ToM task (Happé strange stories test; Happé, 1994) showed higher activation in the STS. Results of the two available studies on ASD patients indicated reduced activation in networks that form the mentalizing system, including the STS, as well as regions necessary for the understanding of shared social interaction. Results also suggested a reduced connectivity in these networks and insufficient transmission in visual pathways.

3.2. Schizophrenia

3.2.1. Behavioral findings

Several studies involving a total of 287 patients have shown that patients with schizophrenia are less able to use mental state terms and to identify intentional behavior (intentionality) when describing the ToM animations (Bliksted et al., 2014; Bliksted et al., submitted for publication; Das et al., 2012a; Horan et al., 2009; Koelkebeck et al., 2010, 2013; Lugnegard et al., 2013; Pedersen et al., 2012; Russell et al., 2006) (for details, see Table 2). However, in the same studies no differences were evident in intentionality scores regarding the RAN animations. Six studies used the GD animations in addition to ToM and RAN animations. Among these, four studies found significant differences in intentionality scores where healthy controls scored higher than patients with schizophrenia (Horan et al., 2009; Koelkebeck et al., 2010, 2013; Russell et al., 2006). Two studies found no difference in intentionality between patients and healthy controls on the GD animations (Lugnegard et al., 2013; Pedersen et al., 2012). There is evidence that patients with schizophrenia describe the ToM animations less appropriately than healthy controls (Bliksted et al., 2014; Bliksted et al., submitted for publication; Das et al., 2012a; Koelkebeck et al., 2010; Lugnegard et al., 2013). No differences in appropriateness scores were reported for the GD animations (Horan et al., 2009; Koelkebeck et al., 2010, 2013; Lugnegard et al., 2013; Pedersen et al., 2012), but one study found that patients scored lower in appropriateness on the RAN animations (Bliksted et al., submitted for publication). Regarding the length of the participants’ response, two studies found that patients used fewer words describing the animations than the controls on all three animation types (Koelkebeck et al., 2013; Pedersen et al., 2012). Two studies did not find any differences regarding the length of answers between patients and controls (Horan et al., 2009; Koelkebeck et al., 2010). Three studies also investigated subgroups based on positive and negative symptoms of schizophrenia (Bliksted et al., submitted for publication; Horan et al., 2009; Russell et al., 2006). One study implied a strong correlation between a range of negative symptoms and deficits regarding appropriateness of ToM and RAN animations as well as intentionality of the ToM animations (Bliksted et al., submitted for publication). Another study found that patients with high levels of specific negative symptoms (apathy and anhedonia) showed lower appropriateness scores but higher intentionality scores on the RAN animations than controls (Horan et al., 2009). In general, there seems to be a tendency of patients with a combination of low levels of both negative and positive symptoms to display fewer ToM deficits than other symptom subgroups (Bliksted et al., submitted for publication; Russell et al., 2006). Taken together, findings imply a reduced use of mental state terms and less appropriate descriptions, predominantly of the ToM animations, in patients with schizophrenia, with psychopathology impacting the level of alterations.

Table 2

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Mean age (S) (SD)</th>
<th>N (S)</th>
<th>N (HC)</th>
<th>Type of animations used in study</th>
<th>Behavioral data</th>
<th>MRI data</th>
<th>More than one rater a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russell et al. (2006)</td>
<td>*</td>
<td>61</td>
<td>22</td>
<td>X(4) X(3) X(2) X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Horan et al. (2009)</td>
<td>40.1 (10.8)</td>
<td>55</td>
<td>44</td>
<td>X(4) X(4) X(4) X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Koelkebeck et al. (2010) b</td>
<td>24.5 (5.6)</td>
<td>23</td>
<td>23</td>
<td>X(4) X(4) X(4) X</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Das et al. (2012a), Das et al. (2012b) c,d</td>
<td>34.5 (8.4)</td>
<td>20</td>
<td>19</td>
<td>X(4) – X(4) – X</td>
<td>-</td>
<td>X</td>
<td>No</td>
</tr>
<tr>
<td>Pedersen et al. (2012) e</td>
<td>29.0 (8.2)</td>
<td>15</td>
<td>14</td>
<td>X(3) X(3) X(3) X</td>
<td>X</td>
<td>X</td>
<td>No</td>
</tr>
<tr>
<td>Koelkebeck et al. (2013) e</td>
<td>34.5 (10.1)</td>
<td>18</td>
<td>30 (27)</td>
<td>X(3) X(3) X(3) X</td>
<td>X(3)</td>
<td>X</td>
<td>BVM No</td>
</tr>
<tr>
<td>Lugnegard et al. (2013)</td>
<td>28.8 (4.1)</td>
<td>36</td>
<td>50</td>
<td>X(4) X(4) X(4) X</td>
<td>-</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Bliksted et al. (2014)</td>
<td>22.7 (3.1)</td>
<td>36</td>
<td>36</td>
<td>X(4) – X(4) X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Bliksted et al. (submitted for publication) f</td>
<td>22.9 (3.5)</td>
<td>59</td>
<td>59</td>
<td>X(4) – X(4) X</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>

a Ratings of each animation done by at least two different persons.

b Used a shortened version of the animations.

c Only male subjects.

d Two articles published with different data from the same sample.

e Used (another) shortened version of the animations.

f Includes the participants from Bliksted et al. (2014).

* Used 4 patient subgroups with the following mean ages: 35.84 (8.82); 33.4 (9.52); 29.93 (8.03); 36.37 (11.59).

Please cite this article as: Bliksted, V., et al., Discriminating autism spectrum disorders from schizophrenia by investigation of mental state attribution on an on-line mentalizing., Schizophr. Res. (2016), http://dx.doi.org/10.1016/j.schres.2016.01.037
patients with schizophrenia in comparison to healthy controls in an fMRI study using the animated triangles task. They assessed not only the task-dependent activation of the ToM network using ToM, GD, and RAN animations (GD + RAN = nToM), but also the time course of activation. Authors observed in patients with schizophrenia, as compared to healthy controls, excessive activation in the right inferior and middle frontal gyri, the left STS, the precuneus and the left cerebellum on the contrast ToM versus nToM. In patients, however, this activation was identified only in the second half of the animation presentations (12 s), whereas controls showed similar activations during the first half of the animations. Summarizing the little available neuroimaging data, they suggest both reduced or exaggerated activation in regions related to the cerebral ToM network in schizophrenia, probably involving a delay of activation in specific brain regions and/or reduced coupling mechanisms of comprehensive networks, e.g. the DM and the mediodorofrontal networks.

3.3. Comparison between patients with ASD and schizophrenia

One study directly compared patients with ASD, patients with schizophrenia and healthy controls (Lugnegard et al., 2013). This study reported that both patients with ASD and with schizophrenia scored significantly lower on the behavioral ToM task than the healthy controls. Patients with schizophrenia displayed lower ToM intentional and appropriateness scores than patients with ASD, with male participants with schizophrenia showing the lowest task performance within all three groups.

4. Meta-analyses

We conducted several meta-analyses comparing the score means and SDs of the ToM and RAN animations (see Table 3). As most of the studies did not use the GD animations and results were inconsistent, we focused on the ToM and RAN animation intentionality and appropriateness scores. Intentionality refers to the ability to apply mental state terms to the animations. We expected more mental state terms reported on the ToM animations than on the RAN animations in both patient groups and healthy controls, which would result in positive SMD values (range 0 to 5, where 5 signifies no mental state attributions to RAN animations combined with maximum mental state attribution to the ToM animations). We expected SMD values below 5 in both patient groups due to ToM deficits. Appropriateness corresponds to the ability to accurately describe the proceedings of the animations. We expected ASD patients and patients with schizophrenia to make less appropriate descriptions on the ToM animations, which would lead to negative SMD. Healthy participants, who are able to make appropriate descriptions of both types of animations, would score close to 0 in SMD. With regard to the only available study that compared matched groups of ASD, schizophrenia and healthy controls, we expected the patients with schizophrenia to have a smaller SMD difference on the intentionality scores and a larger SMD difference on the appropriateness scores than the ASD patients.

We also performed subgroup meta-analyses regarding FES and LLS patients. The range of age for the FES patients was 22.9–24.5 years and 28.8–40.1 years for the LLS patients. The range of the duration of the illness was 0.35–0.36 years for the FES patients and 5.6–17.4 years for the LLS patients.

4.1. Intentionality

The differences between ToM and RAN animation scores for intentionality produced similar results for the ASD (SMD = 3.43; Chi² = 6.61 and P = 0.09; I² = 54.6%; Tau² = 0.32; SMD = 0 test: Z = 8.74 and P < 0.00) and schizophrenia patient sample (SDM = 3.18; Chi² = 50.14 and P < 0.00; I² = 90.0%; Tau² = 1.23; SMD = 0 test: Z = 6.51 and P < 0.00) (see Figs. 2 and 3). In the subgroup analyses of the schizophrenia sample, we found that the FES patients had higher intentionality scores (SMD = 4.19; Chi² = 0.42 and P = 0.52; I² = 0.00; Tau² = 0.00; SMD = 0 test: Z = 14.75 and P < 0.00) than the participants with LLS (SDM = 2.56; Chi² = 14.52 and P < 0.00; I² = 79.3%; Tau² = 0.50; SMD = 0 test: Z = 6.17 and P < 0.00) (see Figs. 4 and 5). Results indicate no differences between patients with ASD and schizophrenia regarding the use of mental state terms, but FES patients showed a better performance than patients with LLS.

4.2. Appropriateness

We investigated appropriateness scores for both groups on ToM and RAN animations. We found that patients with ASD had a larger SMD (SMD = −2.37; Chi² = 18.14 and P < 0.001; I² = 83.5%; Tau² = 0.90; SMD = 0 test: Z = 4.44 and P < 0.001) compared to the LLS group (SMD = 1.06; Chi² = 0.00 and P = 1.00; I² = 0.00; Tau² = 0.00; SMD = 0 test: Z = 2.37 and P < 0.001) compared to the LLS group (SMD = 1.06; Chi² = 0.00 and P = 1.00; I² = 0.00; Tau² = 0.00) (see Figs. 6 and 7). Moreover, we performed subgroup analyses of patients with FES and LLS regarding appropriateness scores. The FES group had a rather similar SMD (SMD = −1.03; Chi² = 2.33 and P = 0.13; I² = 57.0%; Tau² = 0.10; SMD = 0 test: Z = 3.62 and P < 0.001) compared to the LLS group (SMD = −1.06; Chi² = 5.50 and P = 0.14; I² = 54.5%; SMD = 0 test: Z = 7.72 and P < 0.001) (see Figs. 8 and 9). Based on the mean appropriateness scores (see Table 3), the results imply that patients with ASD have more severe deficits than patients with schizophrenia when describing the ToM animations appropriately, while schizophrenia subgroups did not differ significantly.

Table 3

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>ASD (A)/schizophrenia (S)</th>
<th>N</th>
<th>Intentionality</th>
<th>Appropriateness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>ToM Mean SD</td>
<td>ToM Mean SD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horan et al. (2009)</td>
<td>S</td>
<td>55</td>
<td>3.10 0.90</td>
<td>1.80 0.50</td>
</tr>
<tr>
<td>Koekelrebeck et al. (2010)</td>
<td>S</td>
<td>23</td>
<td>2.70 0.80</td>
<td>1.79 0.56</td>
</tr>
<tr>
<td>Pedersen et al. (2012)</td>
<td>S</td>
<td>15</td>
<td>3.11 0.55</td>
<td>1.85 0.68</td>
</tr>
<tr>
<td>Koekelrebeck et al. (2013)</td>
<td>S</td>
<td>18</td>
<td>3.00 1.50</td>
<td>1.35 0.75</td>
</tr>
<tr>
<td>Bliksted et al. (submitted for publication)</td>
<td>S</td>
<td>59</td>
<td>3.49 0.86</td>
<td>2.24 0.63</td>
</tr>
<tr>
<td>Lugnegard et al. (2013)</td>
<td>S</td>
<td>36</td>
<td>2.85 1.10</td>
<td>1.60 0.53</td>
</tr>
<tr>
<td>Lugnegard et al. (2013)</td>
<td>A</td>
<td>54</td>
<td>3.53 0.98</td>
<td>2.03 0.55</td>
</tr>
<tr>
<td>Castelli et al. (2002)</td>
<td>A</td>
<td>10</td>
<td>2.90 0.60</td>
<td>0.50 0.20</td>
</tr>
<tr>
<td>Zwickel et al. (2011)</td>
<td>A</td>
<td>19</td>
<td>2.51 0.61</td>
<td>1.21 0.49</td>
</tr>
<tr>
<td>White et al. (2011)</td>
<td>A</td>
<td>16</td>
<td>3.02 0.49</td>
<td>1.03 0.42</td>
</tr>
</tbody>
</table>

* Used a 2-point appropriateness scale which was converted to a 3-point scale ((score/2) × 3).

† Used a 5-point appropriateness scale which was converted to a 3-point scale ((score/5) × 3).
5. Discussion

In our comprehensive review and meta-analysis, the main aim was to evaluate the use of the animated triangles task in the differential diagnosis of ASD and schizophrenia. The task is known for its properties as an on-line, real-world measure of ToM abilities (Castelli et al., 2002; Das et al., 2012a) with the potential as an endophenotype (Koelkebeck et al., 2010). While common genetic and neuronal mechanisms (Couture et al., 2010; Froese et al., 2013) as well as phenotypic similarities in deficient social functioning (Cheung et al., 2010) have been suggested as indicative for ASD and schizophrenia, different contributions to cognitive styles from both disease entities have also been discussed (Couture et al., 2010). Despite the potential benefits of the animated triangles task in differential diagnosis in this cognitive domain, we are aware of only one study that compared patients with ASD with schizophrenia patients employing this task (Lugnegard et al., 2013). Thus, we summarized a total of 21 papers that investigated patients with ASD or schizophrenia by means of behavioral and brain imaging.
methods. Moreover, we performed a meta-analysis comparing both patient groups with regard to behavioral findings and additionally investigated subgroups of patients with schizophrenia with shorter and longer duration of illness.

While the behavioral results across the reviewed studies indicate general difficulties in using mental state terms and limited appropriateness of descriptions in both patient groups, findings are inconsistent. These inconsistencies are most probably due to differences in stimuli, cueing, task versions and rating systems (see Tables 1 and 2). Moreover, biased results due to small patient groups, mainly in the studies with patients with ASD, must be taken into account. We therefore addressed the question of small sample sizes as a possible source of divergent results in the behavioral data sets. A meta-analysis of the studies that published behavioral data on patients with ASD (N = 98) and schizophrenia (N = 206) was performed. In the analysis, the ASD group displayed larger deficits in the correct description of the animations (appropriateness), mirroring stronger interaction comprehension abnormalities. Performance in patients with schizophrenia showed that they described both the ToM and the random animations only partly correct, irrespective of duration of illness (see Figs. 6 and 7). ASD and patients with schizophrenia had similar problems regarding detection of intentionality, with ASD patients performing slightly better.

![Fig. 4. Forest plot for studies with first-episode patients (FES) with schizophrenia (intentionality).](image)

![Fig. 5. Forest plot for studies of patients with schizophrenia with longer-lasting illness duration (LLS) (intentionality).](image)
This result could also be due to the tendency of patients with schizophrenia assigning more intentionality to the random animations (see Figs. 2 and 3).

It has been hypothesized that deficits in schizophrenia might be much more profound, generalized (Horan et al., 2009) and an indicator for underlying cognitive performance deficits (Bora et al., 2009) compared with ASD, displaying domain-specific deficits. The only available study on the animated triangles task that found worse performance of patients with schizophrenia than of ASD patients (Lugnegard et al., 2013) seems to underscore this assumption. However, results of comparison studies on other social cognition tasks (Chung et al., 2014; Couture et al., 2010) showed equally reduced performance of both groups in comparison to controls, or even stronger deficits in ASD patients (Bolte and Poustka, 2003). These findings might not only hint at domain-specific deficits that can be observed in ASD or schizophrenia over different tasks, but might also depend on sample size or task presentation modes.

ASD are developmental disorders prevalent from early childhood, while schizophrenic symptomatology usually develops in the early 20s. As ToM deficits have been found to be stable across different phases.

**Fig. 6.** Forest plot for studies with patients with ASD (appropriateness).

**Fig. 7.** Forest plot for studies with patients with schizophrenia (appropriateness).
of schizophrenia, and seen among first-degree relatives (Bora and Pantelis, 2013; Green et al., 2012; Horan et al., 2012), they have been discussed as an endophenotype in this disorder. The investigation in different stages of the disease, e.g. in FES and LLS, is of interest in this regard and might help understand development of social cognitive abnormalities in schizophrenia versus such in ASD. A decline in task-performance over the years might be observed in LLS, leading to a performance closer to that of ASD patients (Ozguven et al., 2010). Thus, we investigated patients with FES versus LLS. Our subgroup analysis showed that FES patients used more mental state terms (intentionality) than LLS (see Figs. 4 and 5) while appropriateness of description was not affected by the years of illness (see Figs. 8 and 9). This implies that duration of illness might contribute to a reduction in mentalizing abilities that is probably rooted in brain pathological changes (Nickl-Jockschat et al., 2011), specifically resulting in a reduced use of emotion-related language, which has previously been shown to correlate with psychopathology (Hong et al., 2015). This finding may also be explained by reduced social contact in patients with LLS as compared to FES, as language perception and production might be altered in patients with social withdrawal (Kumari et al., 2010). The appropriateness of descriptions, however, was not affected in LLS patients; an approximation to autistic symptomatology can, consequently, not be assumed.
A very limited number of cerebral imaging studies are available that employed the animated triangles task. They indicate, with one exception (Pedersen et al., 2012), reduced activation of networks that are task-relevant (MPFC, TPJ (STS) and ACC) in both groups (Castelli et al., 2002; Das et al., 2012a; Kana et al., 2009). Moreover, abnormal network activation and de-synchronization in both samples (Castelli et al., 2002; Das et al., 2012b; Kana et al., 2009) with a general under-activation of frontal networks was suggested, including a reduced coupling between task-positive (ToM-related) and task-negative (e.g. DM) networks (Das et al., 2012b). A recent meta-analysis of ten ToM tasks in fMRI investigated patients with ASD and schizophrenia as well as healthy participants (Sugranyes et al., 2011). Both patient groups showed medial prefrontal hypoactivation, which was more pronounced in ASD, while ventrolateral prefrontal activation reductions were seen mostly in patients with schizophrenia. Amygdala hypoactivation was observed in patients with ASD during more complex ToM tasks. Both disorders were associated with hypoactivation within the STS during ToM tasks. Only in schizophrenia the somatosensory cortex was activated more strongly, while the thalamus failed to activate, at all. In ASD, the somatosensory cortex was activated less. A neuroimaging study comparing ASD and psychotic disorders on a picture-sequence task also revealed reduced cortical activation to the experimental condition in ASD and aberrant activation in the control condition in schizophrenia (Ciaranidaro et al., 2015). Thus, results of neuroimaging studies are generally in favor of abnormal activation in ToM networks in both groups, with distinct activation patterns evident for each group. However, due to the limited number of studies available, no further statement regarding the use of imaging studies employing the animated triangles task in differential diagnosis can be made.

As already indicated, our data review and meta-analysis suffer from the fact that the animated triangles task uses a complex behavioral rating system which is liable to subjective bias, as answers are often rated by only one person, and allows multiple scoring approaches. Most of the studies featured different task designs, including varying lengths of animations and incomplete task presentation (less than 12 animations). Some studies used relatively small sample sizes, which may have reduced the effect sizes of the task. To address the issue of group-size, we created a large and uniform data-set. For this, we had to transform data and cannot eliminate the possibility that the data are biased through the limitations mentioned above. However, for our meta-analysis we were able to compare a relatively consistent data set with a large number of participants. The number of patients with FES and LLS is rather small. Nonetheless, as results are in line with what could have been expected, we present the data here. As intra-individual factors might be strong predictors of individual outcome (Borsboom et al., 2004), it is debatable how valid cognitive research is on the individual level. However, as the animated triangles task has been validated in the SCOPE project (Social Cognition Psychometric Evaluation) (Pinkham et al., 2014) as well as the Human Connectome Project, the validity of the task must be deemed adequate to justify the use on the individual level.

Based on the results of this review of the existing literature, we have identified several problems that should be addressed in future. First, the task requirements regarding behavioral investigations need to be unified to create comparable and reliable data sets. For now, the inconsistencies in methodology described earlier lead to difficulties in efficiently comparing the data. The Human Connectome Project has aimed at using standardized test batteries which should also be used in patient samples to achieve larger data sets. More standardized ways of rating as suggested by White et al. (2011) might also be helpful. Moreover, the animated triangles task consists of an implicit mentalizing aspect, which is a prerequisite of explicit mentalizing (Van Overwalle and Vanderkervchke, 2013). The question of when and if implicit and explicit mentalizing is involved in task performance has not been specifically addressed. These different mechanisms might contribute to divergent results and could be targeted by functional imaging studies. Other methods, e.g. eye-tracking, could be utilized to strengthen fMRI and behavioral findings, as this would achieve a more reliable measuring of differences in task performance, unbiased by language skills (Zwickel, 2009). Furthermore, patients with ASD are at risk of developing co-morbid schizophrenia or psychotic symptomatology (Mouridsen et al., 2008; Stahlberg et al., 2004). It would be interesting to assess patients with both symptom constellations to investigate overlap and severity of ToM deficits. It should be noted that most of the studies mentioned above investigated ASD patients with higher levels of functioning and without severe mental retardation. Thus, results of this review and meta-analysis cannot be extended to forms of autistic disorders with severer deficits.

In summary, we reviewed studies assessing ToM abilities in patients with ASD and schizophrenia on an online-mentalizing task and performed a meta-analysis on behavioral data to address the question of differential diagnosis in both patient groups. Results did not only imply reduced performance of both patient groups as compared to healthy controls, but also of patients with ASD in appropriateness of animation descriptions as compared to patients with schizophrenia, who performed worse on the use of mental state terms. Moreover, patients with FES showed an advantage on intentionality over LLS patients. Results indicate different performance levels of both patient groups, which are, in schizophrenia, dependent of the duration of illness. Differential task performance and underlying mechanism of ToM deficits might be assessed through neuroimaging studies, but results are too inconsistent to achieve sophisticated results yet.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2016.01.037.

Role of the funding source
None.

Contributors
All authors contributed to the design of the review, to the literature search, participated in the consensus process, extracted data from the selected articles, contacted authors for additional information, and wrote the manuscript. V.B. performed the statistical analyses. All authors contributed to, critically revised and have approved the final version of the manuscript.

Conflict of interest
All authors declare that they have no conflicts of interest.

Acknowledgments
We sincerely thank Uta and Christopher D. Frith for their supervision in the progress of this manuscript. We would also like to thank Vladimir Cherkassy, Antonia Hamilton, William Horan, Rajesh Kana, Annette Maria Klein and Jan Zwickel for their help in providing the data for the meta-analysis.

References