The Development and Relationship of Emotion Recognition and Empathy in Children with Autism Spectrum Disorder

Angela Alegre Sanz

Contents

Abstract

Children who are able to recognize the emotions of others and react appropriately and empathetically are found to have better social skills and more positive social relationships. However, children with autism spectrum disorder (ASD) often show impairments in emotion recognition and empathy, which have a negative influence on their social life. So far it is not clear how emotion recognition and empathy develop over time in young children with ASD. Therefore, the present study examined the developmental trajectories of emotion recognition and empathy in children aged 29 to 83 months with ASD in comparison to typically developing peers (TD). Furthermore, empathy greatly relies on emotion recognition, as it involves a reaction to another person's emotion. Hence, we also examined the relationship between emotion recognition and empathy at different time points.

TD (N = 109) and ASD (N = 39) children (mean age = 55 months, SD = 13.01) were tested at three different time points. Behavioral tasks were administered to measure three components of emotion recognition, namely, discrimination, identification, and attribution, and three components of empathy, namely, affective empathy, cognitive empathy and prosocial behavior. The outcomes show that emotion recognition and empathy tend to improve over time in both TD children and children with ASD. However, children with ASD still lagged behind on emotion recognition and empathy skills as compared to TD peers, with cognitive empathy showing the greatest difference between groups. Furthermore, we found different patterns of associations between emotion recognition and empathy in children with ASD from those in TD children.

The findings suggest that children with ASD experience deficits in emotion recognition and empathy from early on, thus highlighting the importance of early intervention aimed at improving emotion recognition and empathy skills in children with ASD.

Keywords: emotion   recognition,   empathy,   autism   spectrum   disorder

The Development and Relationship of Emotion Recognition and Empathy in Children with Autism Spectrum Disorder

1. Introduction

Children with Autism Spectrum Disorder (ASD) are known for their difficulties in forming and maintaining meaningful social relationships (Russo-Ponsaran, Evans- Smith, Johnson, Russo, & McKown, 2016). Emotion recognition and empathy are two essential factors that facilitate adaptive social interactions (Rueda, Fernandez-Berrocal, & Baron-Cohen, 2015). Research has shown that children with ASD have difficulty understanding emotions (Dawson et al., 2002) and display less empathic behaviors when compared to typically developing (TD) children (Rogers et al., 2007). Although there is an abundance of research on emotion recognition and empathy in school-aged children with ASD, not many studies have examined the developmental trajectories of the two constructs in very young children. The present study focused on the developmental trajectory of emotion recognition and empathy in children aged 29 to 83 months with and without ASD. Besides, we examined the relationship between emotion recognition and empathy in these children, investigating whether the two constructs were related at different time points of development. Basic emotions were exclusively investigated (i.e., happiness, sadness, fear and anger) for this study.

1.1. Emotion recognition

Emotion recognition is conceptualized as having three different components: discrimination, identification and attribution (Wiefferink, Rieffe, Ketelaar, De Raeve & Frijns, 2013). Discrimination is the first component to develop and it refers to the ability to discriminate between different facial emotional expressions (Wiefferink et al., 2013). Around the first 4 to 9 months infants are able to discriminate basic facial expressions (Thomas, De Bellis, Graham, & LaBar, 2007). Tottenham, Hare and Casey (2011) studied children (5-12 years old), adolescents (13-18 years old), and adults (19-28 years old) and found an age-related linear increase in discrimination performance of basic emotions.

Nonetheless, emotion discrimination does not imply that a child acknowledges the differences between different emotional states (Wiefferink et al., 2013). Emotion identification, the second component of emotion recognition, allows children to link facial expressions with their corresponding emotions (Wiefferink et al., 2013). Identification of basic emotions appears to be fully developed by age 6 (Markham & Adams, 1992). Herba

and Phillips (2004) reviewed the development of emotion identification during childhood and adolescence and concluded that identification follows a gradual development, where children first learn to identify basic emotions and at a later stage mixed emotions.

Furthermore, recognizing people's emotions requires knowledge that a particular situation evokes a specific emotional state, i.e. attributing an emotion to a social context (Wiefferink et al., 2013). Four-year-olds have been found to accurately attribute basic emotions to a protagonist (Rieffe, Terwogt, & Cowan, 2005). As children grow older, they become more capable of collecting information from situational and facial expression cues, and they become more accurate in recognizing not only basic emotions but also more complicated emotions and mixed emotions (Herba & Phillips, 2004).

To summarize, the abilities to recognize basic emotions emerge from an early age and develop gradually over time. At an older age, children also learn to recognize more complex emotions and mixed emotions (Herba & Phillips, 2004; Tottenham, et al., 2011).

Emotion recognition in children with ASD

Emotion recognition develops through the interaction of neurobehavioral and environmental factors (as reviewed by McClure, 2000). The amygdala plays an important role in the recognition of faces very early in the developmental process, while cortical regions may become crucial in later stages of development (Wiefferink et al., 2013). Further, social interaction, like exposure to emotional situations and modeling from parents, is fundamental in the development of adequate understanding of facial expressions (McClure, 2000).

Children with ASD are exposed to both risks: they are found to have abnormal brain activities when processing emotional information (see review by Harms et al., 2010), and they seem to have low social interest and low interest in the mental states of others (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). For example, children with ASD look less at characters than at the background when shown social pictures (e.g. a group of friends chatting), and they fixate less on people, eyes, and faces than at background objects (Chevalier et al., 2012). This low social interest in children with ASD leads to their limited access to social learning, which is crucial for children's social-emotional development, including the ability to recognize emotions and understanding other people's emotional and mental states (Chevalier et al., 2012).

Many studies have examined emotion recognition in children with ASD, but the findings are mixed. Whereas some studies have reported emotion recognition deficiencies in children with ASD, others have not. For example, Evers, Steyaert, Noens, & Wagemans

(2015) studied labeling of basic emotions in children with ASD aged 6 to 14, and found that children with ASD performed worse than TD children. In contrast, Lacroix, Guidetti, Rogé, and Reilly (2014) investigated emotion discrimination, identification and labeling in 4 to 8 year-olds with ASD and found that children with ASD do not have a developmental delay in discrimination, identification or labeling when compared to TD children. Similarly, Castelli (2005) studied 12-year-olds with ASD and found that these children performed as well as controls in discrimination and identification tasks.

The review by Harms et al. (2010) suggests that the mixed findings of behavioral studies on facial emotion recognition in people with ASD could result from the heterogeneity of the demographic characteristics of the participants and the nature of the tasks. However, as pointed out by the review, eye-tracking studies and neuroimaging studies have mostly reported ASD deficits. Considering the core syndromes of ASD as well as their limited access to social learning, it is reasonable to assume that young children with ASD lag behind in all aspects of emotion recognition as compared to their TD peers.

1.2. Empathy

Empathy is conceptualized as consisting of three components: affective empathy, cognitive empathy, and prosocial behavior (Rieffe, Ketelaar, & Wiefferink, 2010). Affective empathy involves the ability to unconsciously mirror other's emotions, and does not require a differentiation between the self and the other (Ketelaar, Rieffe, Wiefferink, & Frijns, 2013). It is assumed that people are predisposed to automatically replicate the emotions of others, which implies that affective empathy may be an innate ability (Dacety & Jackson, 2004; Rieffe et al., 2010). Affective empathy can be observed very early on, e.g. when one infant cries, other infants might start crying as well (Rieffe et al., 2010). Rieffe and colleagues (2010) studied one- to five-year-olds and found that affective empathy was unrelated to age, suggesting that children do not show an increase in affective empathy, as they get older. Furthermore, Bensalah, Caillies, and Anduze, (2015) reported that children aged 4 to 6 years old showed emotional arousal in response to another person's emotion, and they found that affective empathy did not significantly change between the ages of 4 and 6 year.

Cognitive empathy is the capacity to understand others' feelings and perspectives by identifying people's mental states (Baron-Cohen, & Wheelwright, 2004; Rueda et al., 2015). It is closely related to the development of Theory of Mind (ToM), which refers to the ability to understand the intentions and motivations of others (Rueda et al.,

2015; Rogers, Dziobek, Hassenstab, Wolf, & Convit, 2007). Around two years old, when witnessing other people in distress, children begin to understand that their emotional reaction is due to another person's emotion, which allows them to relieve their own feelings of distress (Netten et al., 2015). Around the age of 5, children are capable of understanding that emotions are embedded in beliefs and desires, and that these beliefs and desires may differ from one person to another (Rieffe, Terwogt & Cowan, 2005).

Prosocial behaviors include comforting and concern about the other, helping, and sharing (Netten et al., 2015; Zhan-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Two-year-olds show concern for others in distress, and provide comfort and help as they develop conscience and concern for the wellbeing of others (Dacety & Jackson, 2004; Zahn-Waxler et al., 1992). However, younger children often offer help from their own perspective, e.g. comforting an adult who is in distress by giving him his/her favorite toy. As children grow up and develop ToM and cognitive empathy, the prosocial behaviors become more appropriate and in accordance to the other person's need. Russell and colleagues (2012) found that prosocial behavior increases between the ages of 4 and 7, and remains relatively stable until the age of 13.

In summary, in TD children affective empathy is assumed to be an innate ability, which does not change extensively over time. In contrast, cognitive empathy and prosocial behaviors appear to develop over time.

Empathy in Children with ASD

Like emotion recognition, the development of empathy involves neurological and experiential factors. Affective empathy is related to areas largely developed at birth such as the mirror neuron system (Schulte-Rüther et al., 2013), and it has been argued to emerge independent from children's social experience (Netten et al., 2015). Studies have found that the mirror neuron system is not impaired in children with ASD and that these children display the same amount of affective empathy when witnessing others in distress (Hadjikhani et al., 2014; Pouw, Rieffe, Oosterveld, Huskens, & Stockman, 2013). Schwenck and colleagues (2012) asked 6- to 17-year-olds with ASD to report how affected they felt after watching a series of videos, and found that ASD performed as well as TD children. Pouw and colleagues (2013) administered the Empathy Questionnaire to children aged 10 to 12 and found no impairments in affective empathy in children with ASD.

Cognitive empathy is related to brain areas that undergo refinements from childhood to adulthood, such as the medial prefrontal cortex, superior temporal sulcus, and

temporal lobes (Schulte-Rüther et al., 2013). Besides, social learning is a key factor in the development of cognitive empathy and prosocial behavior (Netten et al., 2015). For example, a child observing a mother comforting her son after losing his favorite ball helps the child understand both the boy's emotional state (i.e. cognitive empathy) and identify an adequate response (i.e. prosocial behavior) (Netten et al., 2015).

Children with ASD have limited access to social learning (Bird & Viding, 2014), and find the emotions of others confusing and unpredictable which prevents them from showing empathic behaviors (Pouw et al., 2013). Regarding the development of cognitive empathy and prosocial behaviors, studies have found that children with ASD lag behind their TD peers. For example, Pouw and colleagues (2013) using the Empathy Questionnaire found that children with ASD aged 10 to 12 years old performed worse than TD children in cognitive empathy. Using the Strengths and Difficulties Questionnaire (SDQ) Izuka and colleagues (2010) reported prosocial behavior impairments in 6 to 12-year-olds with ASD. However, although children with ASD have deficits in these areas, research has shown that they improve over time. Schwenck and colleagues (2012) asked 6 to 17-year-olds with ASD to take the perspective of another individual, and their results showed that children with ASD had cognitive empathy impairments, and that older children with ASD performed better than younger ones. Employing the SDQ, Russell and colleagues (2012) showed that children with ASD showed less prosocial behavior at every developmental stage, but followed the same trajectory as TD children. Namely, there was an increase between the ages of 4 and 7, and it remained relatively stable until the age of 13.

To summarize, previous research suggests that children with ASD do not have impairments in affective empathy, but deficits are observed in cognitive empathy and prosocial behavior (Mazza et al., 2014; Rueda et al., 2015; Schwenck et al., 2012). Moreover, studies show that children with ASD improve over time in cognitive empathy and prosocial behaviors (Schwenck et al., 2012; Russell et al., 2012).

1.3. Relationship between Emotion Recognition and Empathy

Empathy greatly relies on emotion recognition, as by definition it involves a reaction to another person's emotion. Indeed, emotion recognition has been considered a prerequisite of empathy, as one should first be able to identify what others are feeling in order to act accordingly (Clark, Winkielman, & McIntosh, 2008; Rueda et al., 2015). For instance, a child who is able to recognize that a peer feels sad because he is not included in a game will be able to behave empathically towards the peer, for example by comforting the peer or by

inviting him or her to the game (Trentacosta & Fine, 2010). Rieffe et al., (2010) found that in one- to five-year-olds emotion understanding and recognition of the emotions of others were positively related to all empathy components, with cognitive empathy and prosocial behaviors having the strongest relationships. Furthermore, Camodeca and Coppola (2016) showed that in 6-year-old TD children, accurate identification and attribution of emotional expressions play an important role in the prosocial motivation to help victims of bullying in the classroom.

Children with ASD exhibit emotion recognition impairments, which can prevent them from learning about another person's feelings and behaviors, and thus hinder them from behaving empathetically (Dawson et al., 2002; Uljarevic, & Hamilton, 2013). We could not find research that directly studied the relation between emotion recognition and empathy in children with ASD, but we found some studies that investigated the link between emotion recognition and ToM, which is closely related to cognitive empathy. For example, Rice, Wall, Fogel and Shic (2015) reported a positive correlation between emotion discrimination for basic emotions and ToM in 5 to 11-year-olds with ASD. Moreover, Trevisan and Birmingham (2016) carried out a meta-analysis on the relation between emotion recognition and social function in 5 to 40-yearold people with ASD. They reported a strong positive relationship between identification of emotional facial expressions and ToM throughout the lifespan. Similarly, Dyck and colleagues (2006) found a strong positive correlation between emotion identification and attribution on the one hand, and ToM on the other hand in 8-year-olds with ASD.

1.4. Current Study

We examined the development of emotion recognition and empathy and their relation in children between the ages of 29 to 83 months with and without ASD. We chose this age group for two main reasons. First, according to the literature, the development of emotion recognition and empathy begins very early on, but not many studies have examined the emotion functioning in the young age group. Second, few studies have examined developmental trajectories of emotion recognition and empathy in this age group, as most studies are cross-sectional.

Our first hypothesis was that emotion discrimination, identification, and attribution follow a path of gradual increase with age in children both with and without ASD (Herba & Phillips, 2004; Tottenham, et al., 2011). Regarding group differences, we expected children with ASD were less able to discriminate, identify, and attribute emotions at all

developmental stages compared to TD children (Chevallier et al., 2012; Tager-Flusberg & Sullivan, 1995; Chevallier et al., 2012).

Our second hypothesis was that in children both with and without ASD, affective empathy remains stable over time (Bensalah et al., 2015; Rieffe et al., 2010), and cognitive empathy and prosocial behaviors increase with age (Russell et al., 2012; Schulte-Rüther et al., 2013). Regarding group differences, for affective empathy we expected no group difference between children with and without ASD (Pouw et al., 2013). For cognitive empathy and prosocial behaviors, we expected children with ASD to show lower level of cognitive empathy and less prosocial behaviors than TD peers at all developmental stages (Pouw et al., 2013; Mazza et al., 2014; Rueda et al., 2015; Iizuka et al., 2010; Russell et al., 2012).

Our third hypothesis was that in TD children, all empathy components are positively related to emotion identification and attribution (Rieffe et al., 2010). It is unclear how emotion discrimination relates to empathy components, and thus we made no assumption about their relation. Regarding children with ASD, we expected emotion discrimination, identification and attribution to be positively correlated to cognitive empathy (Dyck et al., 2006; Trevisan & Birmingham, 2016; Wall et al., 2015). It is unclear how affective empathy and prosocial behaviors relate to emotion recognition in children with ASD. Hence no assumptions were stated in this regard.

2. Method

2.1. Participants and Procedure

The sample included a total of 148 children: 39 children with ASD and 109 TD children. The children were tested on three different time points with a 10 to 20 month interval. Descriptive characteristics of participants at the three time points can be found in Table 1.

The mean age of ASD and TD children at each time point did not significantly differ. Furthermore, IQ scores were obtained during the first time point for both ASD and TD children. TD children had IQ scores than children with ASD. Moreover, there were significantly more males than females per time point in both the TD and the ASD group, and the male-female ratio significantly differed between the groups. That is, the ASD group had higher percentage of boys than the TD group.

Children with ASD were recruited via the Center for Autism in Leiden, the Netherlands, which specializes in the diagnosis of ASD in children and adolescents.

Only the children who had received a diagnosis of ASD and had no other comorbidities were included in the sample. A certified psychologist or psychiatrist carried out the diagnosis following the DSM-IV-TR criteria. All ASD participants were high functioning, having IQ above 80. TD children were recruited from day care centers and primary schools in the Netherlands. The parents and/or teachers of these children indicated they did not suffer from any clinical problem. Permission to conduct the study was granted both by the Ethics Committee of Leiden University and the Center for Autism. In addition, written consent from all parents was obtained before testing.

Children were tested either at home, school or at the Center for Autism individually and in a quiet room. Children with ASD were tested by professionals with experience in working with the clinical group. TD children were tested by trained psychology students. This study was part of a large-scale longitudinal research study on children's social-emotional development, which administered more tasks than reported here.

Table 1. Descriptive characteristics of the participants

	First time point		Second time point		Third time point	
	ASD (n = 39)	TD (n = 109)	ASD (n = 39)	TD (n = 109)	ASD (n = 39)	TD (n =109)
Age, mean (SD) months	56.08 (10.55)	54.98 (13.82)	68.90 (10.44)	67.51 (14.18)	81.18 (10.63)	79.37 (14.08)
Age range, months	31-71	29-83	45-85	41-95	56-97	53-107
IQ, mean (SD)**	102.11 (15.98)	112.98 (14.75)	-	-	-	-
Gender, no. (%)**						
Male	37 (94.9)	59 (54.1)	37 (94.9)	59 (54.1)	37 (94.9)	59 (54.1)
Female	2 (2)	50 (45.9)	2 (2)	50 (45.9)	2 (2)	50 (45.9)
ASD subtype no. (%)						
Autistic disorder	22 (56.4)	-	22 (56.4)	-	22 (56.4)	-
PDD-NOS	17 (43.6)	-	17 (43.6)	-	17 (43.6)	-

IQ scores were missing for 3 children with ASD and 10 TD children
** Significant at $p < .01$.

2.2. Materials

Participants completed a set of tasks while their ability to recognize emotions and their empathetic reactions were observed and coded. Descriptive characteristics of the tasks per participant group and time point can be found in Table 2.

Emotion recognition tasks

Discrimination. The Emotion-Discrimination task was administered to measure children's ability to discriminate between facial expressions (Wiefferink et al., 2013). This task consisted of a neutral condition and a facial expression condition. First, the neutral condition was administered to assure that children understood how to sort different cards. This condition consisted of making a distinction between cars and flowers (task 1) and neutral faces with hats in contrast to neutral faces with glasses (task 2). Children who did not perform well enough on the neutral condition did not move on to the facial expression task, as it was assumed that sorting objects and neutral faces was easier than sorting emotional facial expressions. The facial expression condition was then administered and it consisted of two levels: 1. making a distinction between different facial expressions of different valences (task 3: happy versus sad) and 2. distinguishing emotional faces of the same valence (task 4: angry versus sad). The emotions used in this task were happy, sad and angry (see Figure 1).

In the discrimination task, children were given a sheet, which included sample pictures of one valence on one corner (e.g., happy face) and a sample picture of the different valence on the other corner (e.g., sad face). Children were then given six cards one by one (three cards for each valence in a random order), and were instructed to place the card on the same side as the card of the corresponding valence. The facial expression drawings used in this task were black and white computer generated images, based on photos of 3 and 4-year-olds drawn from a larger database of photos of different emotional expressions. Correctly placed cards were counted, and children could gain maximally 3 points per valence (Wiefferink et al., 2013).

Figure 1. Examples of the facial emotional expression drawings corresponding to the emotions "happy" (right), "sad" (middle), and "angry" (left) for the Emotion-Discrimination task (adapted from Wiefferink et al., 2013).

Identification. The Emotion-Identification task was administered to measure children's ability to associate emotion words with their corresponding facial expressions (Wiefferink et al., 2013). The four basic emotions were tested: happiness, sadness, fear and anger. The task consisted of eight drawings representing emotional expressions, two for each emotion. Two sheets were shown to the children, each containing four facial emotion expression drawings. The researcher asked the children: "Who looks happy?" Children were supposed to point to the drawing with the happy expression. Afterwards, the children were asked: "Is there anybody else who is happy?" The same procedure was repeated for the remaining emotions, i.e., sadness, fear and anger. Figure 2 shows examples of the drawings corresponding to the four basic emotions. The amount of correctly identified emotions was counted, with two being the maximum per emotion.
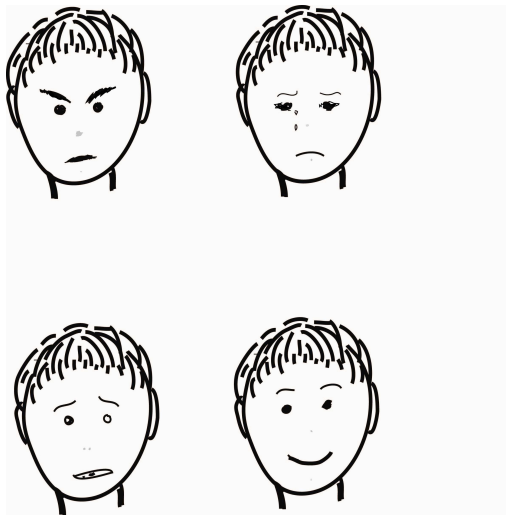

Figure 2. Examples of the facial emotional expression drawings corresponding to the emotions "anger" (top left), "sadness" (top right), "fear" (bottom left) and "happiness" (bottom right) used for the Emotion-Identification task (pictures adapted from Wiefferink et al., 2013).

Attribution. The Emotion-Attribution task was administered in order to evaluate children's ability to acknowledge that certain situations typically evoke particular emotions. The task consisted of eight sketches portraying typical emotion-eliciting situations, two sketches for each basic emotion (i.e., happiness, anger, fear, and sadness) (Wiefferink et al., 2013).

Children were shown images in addition to a spoken explanation, such as "boy sees crocodile". Then children were asked to first say how the boy would feel (verbal condition), and then to point to the image with the correct facial expression (visual condition). We employed both verbal and visual conditions, as it may be that children who do not know the word for an emotion can identify its corresponding emotional expression. It is common for

a situation to be potentially associated with different emotions of the same valence (e.g., a child falling off a bike can feel sad as he might think he will never learn to bike, or angry if someone pushed him off the bike). Thus, we scored an answer as correct when children attributed an emotion within the intended valence. Proportions of correct responses were then calculated to obtain children's scores.

Empathy task

An observational empathy task was carried out in order to observe the extent of the empathetic reactions expressed by the children (Rieffe et al., 2010). The task involved an experimenter acting out three different scenarios corresponding to either a happy, sad or angry facial expression. The scenarios were modified per time point to accommodate to the age of the children at the time of testing.

For example, in one of the happy scenarios the experimenter held a pen and clicked it several times while laughing and displaying a happy facial expression. In one of the sad scenarios, the experimenter pretended to get a paper cut when passing the page of notebook, and reacted by putting his finger in his mouth, and saying "ow" while displaying a painful facial expression. Lastly, in one of the anger scenarios, the experimenter pretended to write with a pen that did not properly write, and reacted by saying "silly pen" while displaying an angry facial expression.

Children's behavioral responses throughout the different scenarios were observed and scored on a checklist (0 = no, 1 = a bit, 2 = a lot). Further, responses were grouped into the three empathy components: affective empathy (e.g., the child shows similar facial expression as the experimenter), cognitive empathy (e.g., the child looks at the experimenter), and prosocial behavior (e.g., the child tries to comfort the experimenter).

Table 2. Descriptive statistics of the tasks

	First Time Point		Second Time Point		Third Time Point	
	ASD	TD	ASD	TD	ASD	TD
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Emotion recognition						
Answer range: 0-3						
Discrimination	(n = 39)	(n = 109)	(n = 39)	(n = 109)	(n = 39)	(n = 109)
	1.88 (.90)	2.02 (.74)	2.15 (.75)	2.37 (.72)	2.34[**] (.73)	2.77[**] (.40)

Answer range: 0-2 Identification	(n = 39)	(n = 107)	(n = 39)	(n = 109)	(n = 39)	(n = 109)
	.56 (.43)	.78 (.28)	.85 (.28)	.90 (.13)	.82 (.29)	.86 (.18)
Answer range: 0-1 Attribution	(n = 39)	(n = 107)	(n = 39)	(n = 109)	(n = 39)	(n = 109)
	.49* (.43)	.76* (.29)	.83 (.28)	.88 (.14)	.80 (.28)	.86 (.18)
Empathy						
Answer range: 0-2 Affective	(n = 39)	(n = 109)	(n = 39)	(n = 109)	(n = 39)	(n = 109)
	1.0 (.58)	1.11(.54)	1.38 (1.35)	1.60 (1.47)	.76* (.55)	1.03* (.66)
Answer range: 0-2 Cognitive	(n = 39)	(n = 108)	(n = 39)	(n = 107)	(n = 39)	(n = 109)
	.75** (.56)	1.15** (.38)	1.14* (.49)	1.33* (.31)	1.04** (.37)	1.25** (.24)
Answer range: 0-2 Prosocial	(n = 39)	(n = 108)	(n = 39)	(n = 107)	(n = 39)	(n = 109)
	.25 (.31)	.30 (.30)	.43 (.43)	.49 (.49)	.18 (.26)	.22 (.30)

* Significant at $p < .05$
** Significant at $p < .001$

2.3. Statistical analysis

Statistical analyses were carried out using the program SPSS version 21.0. Missing values were observed on certain measurements, but removing those participants did not change the results, thus they were not excluded. Moreover, the male-female ratio significantly differed between the TD and the ASD groups. However, removing female participants did not alter the mean age and IQ of the groups. Hence, both males and females were included in all further analyses. First, Mann-Whitney U test was used to detect group differences in each variable per time point. This nonparametric test was used since preprocessing showed that the data were not normally distributed.

Secondly, in order to compare the developmental trajectories of emotion recognition in ASD and TD children, we carried out three Repeated Measures Analysis of Variance (R-ANOVA) for discrimination, identification and attribution of emotional expressions separately. Preprocessing of this data showed that the assumption of normality was violated. No non-parametric test of mixed design is available at the moment. Thus, to correct for this violation, reciprocal transformations of the data were performed, but they were unsuccessful in improving normality. To solve this problem, we opted for performing a R-ANOVA on the

original data and on the transformed data separately. The rationale behind this approach is that if both analyses show a similar result, the effect they concur upon is likely there. However, if the analyses differ, no definite conclusion could be drawn, given the unsuitability of each analysis on its own. The same method was applied to analyze the development trajectories of empathy in ASD and TD children, as the data were also not normally distributed.

Finally, to study the relationship between emotion recognition and empathy in ASD and TD children, Spearman non-parametric correlation was employed because preprocessing showed that the data were not normally distributed. First, correlations for emotion recognition and empathy components per time point were calculated separately for the TD and ASD groups. Then we compared group differences in these correlations by performing Fisher r-to-z transformations.

3. Results

3.1. Group differences in emotion recognition and empathy per time point

First, independent-samples t-tests were used to analyze differences between ASD and TD children in the different tasks per time point (see Table 2).

As for the discrimination task, at the first time point the scores of groups did not differ ($U = 2030.0$, $p = .68$). Similarly, at the second time point, discrimination scores between groups did not differ ($U = 1711.5$, $p = .06$). However, at the third time point, children with ASD scored significantly lower than TD children on the discrimination task ($U = 1382$, $p < .001$).

As for the identification task, at the first time point the groups marginally differed ($U = 1668.5$, $p = .056$). At the second ($U = 2058.5$, $p = .75$) and third ($U = 2079$, $p = .83$) time point identification scores of ASD and TD children did not differ.

As for the attribution task, at the first time point, children with ASD scored significantly lower than TD children ($U = 1481.5$, $p = .006$). However, on the second ($U = 2064$, $p = .77$) and third ($U = 2041.5$, $p = .7$) time point attribution scores of ASD and TD children did not differ.

As for the affective empathy task, children with ASD did not differ from TD children at the first ($U= 1805$, $p = .23$) and second time point ($U = 1805$, $p = .16$). However, at the third time point children with ASD scored significantly lower than TD children on the affective empathy task ($U = 1644$, $p = .03$).

As for the cognitive empathy task, children with ASD scored significantly lower than TD children at the first ($U = 1147.5$, $p < .001$), second time point ($U = 1620$, $p = .03$), and third time point ($U = 1379.5$, $p < .001$).

In the prosocial behavior task, at the first ($U = 1885.5$, $p = .30$), second ($U = 1973.5$, $p = .60$) and third ($U = 2026$, $p = .63$) time point identification scores of ASD and TD children did not differ.

3.2. Development of emotion recognition in ASD and TD children

Discrimination

The development of the discrimination aspect of emotion recognition in ASD and TD children was examined with a 3 (time points 1, 2 and 3 on discrimination) x 2 (ASD and TD) mixed R-ANOVA design. See Figure 3.

Original data. Mauchly's test indicated that the assumption of sphericity was not violated (X^2 (2) = .998, $p = .836$). The analysis revealed a significant main effect of time (F (2, 292) = 38.26, $p < .001$) and diagnosis (F (1, 146) = 6.92, $p = .009$). Bonferroni corrected comparisons showed that children's scores in the discrimination task at time point one were significantly lower than time point two (F (1, 146) = 18.67, $p < .001$) and time point three (F (1, 146) = 77.36, $p < .001$) across both groups. Moreover, discrimination scores at time point two were significantly lower than at time point three (F (1, 146) = 19.43, $p < .001$) across groups. No interaction effect was found between time and diagnosis (F (2, 292) = 2.43, $p = .089$). However, the post hoc comparisons revealed that the increase in discrimination ability from the first to the third time point was significantly slower for ASD than for TD children (F (2, 292) = 4.63, $p = .033$).

Reciprocal transformed data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .442, $p < .001$); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.64$). The analysis revealed a significant main effect of time (F (2, 292) = 5134 $p < .001$) and diagnosis (F (1, 146) = 12.43, $p = .001$). Bonferroni corrected comparisons showed that children's scores on the discrimination task at time point one were significantly lower than time point two (F (1, 146) = 21.68, $p < .001$) and time point three (F (1, 146) = 5734, $p < .001$) across both groups. Moreover, discrimination scores at time point two were significantly lower than at time point three (F (1, 146) = 5475, $p < .001$) across groups. Further, there was an interaction between time and diagnosis (F (2, 292) = 13.62, $p < .001$). Post hoc comparisons revealed that the increase in discrimination ability in ASD was of smaller

magnitude than that in TD children from the first to the third time point (F (1, 146) = 16.62, $p < .001$) and from the second to the third time point (F (1, 146) = 12.86, $p < .001$). However, the increase in discrimination ability was of similar magnitude for both TD and children with ASD from the first to the second time point (F (1, 146) = 1.38, $p = .24$).
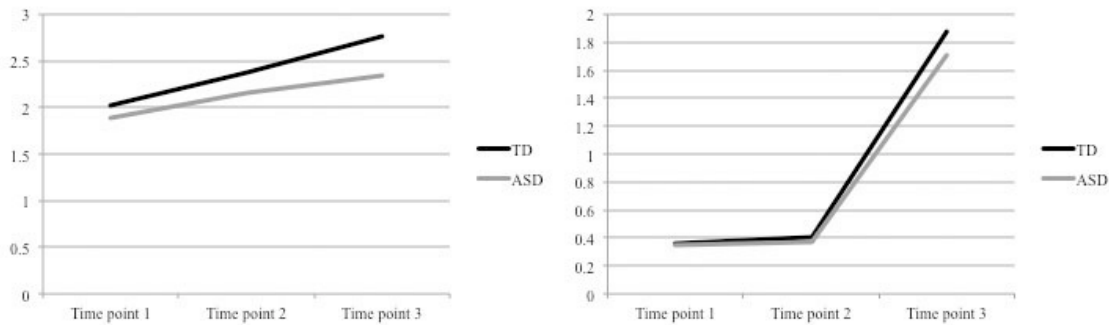

Figure 3. Development of emotion discrimination across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

Identification

The development of the identification component of emotion recognition on ASD and TD children was examined with a 3 (time points 1, 2 and 3 on identification) x 2 (ASD and TD children) mixed R-ANOVA design. See Figure 4.

Original data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .567, $p < .001$); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.7$). The analysis revealed a significant main effect of time (F (2, 288) = 26.98, $p < .001$) and diagnosis (F (1, 144) = 9.56, $p = .002$). Bonferroni corrected comparisons showed that children's scores on the identification task at time point one were significantly lower than at time point two (F (1, 146) = 40.35, $p < .001$) and time point three (F (1, 144) = 21.71, $p < .001$) across both groups. Comparisons also revealed that children's scores on the identification task at time point two (F (1, 144) = 4.52, $p = .035$) were significantly higher than at time point three across both groups. Furthermore, there was an interaction between time and diagnosis (F (2, 288) = 6.19, $p = .007$). Post hoc comparisons revealed that the increase in identification ability in children with ASD was steeper than that in TD children from the first to the second (F (1, 144) = 7.35, $p = .008$) and third time point (F (1, 144) = 6.75, $p = .01$). However, ASD and TD children did not differ in their decrease in identification from the second to the third time point (F (1, 144) = .101, $p = .752$).

Reciprocal transformation. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .659, $p < .001$); therefore, the degrees of freedom were

corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.75$). The analysis revealed a significant main effect of time (F (2, 288) = 23.96, $p < .001$) and diagnosis (F (1, 144) = 6.55, $p = .011$). Bonferroni corrected comparisons showed that children's scores on the identification task at time point one were significantly lower than at time point two (F (1, 146) = 37.51, $p < .001$) and time point three (F (1, 144) = 18.96, $p < .001$) across both groups. Comparisons also revealed that children's scores on the identification task at time point two (F (1, 144) = 4.44, $p = .037$) were significantly higher than at time point three across both groups. Furthermore, there was an interaction between time and diagnosis (F (2, 288) = 5.07, $p = .013$). Post hoc comparisons revealed that the increase in identification ability in ASD was steeper than that in TD children from the first to the second (F (1, 144) = 6.21, $p = .014$) and third time point (F (1, 144) = 5.7, $p = .018$). However, ASD and TD children did not differ in their decrease in identification from the second to the third time point (F (1, 144) = .063, $p = .802$).
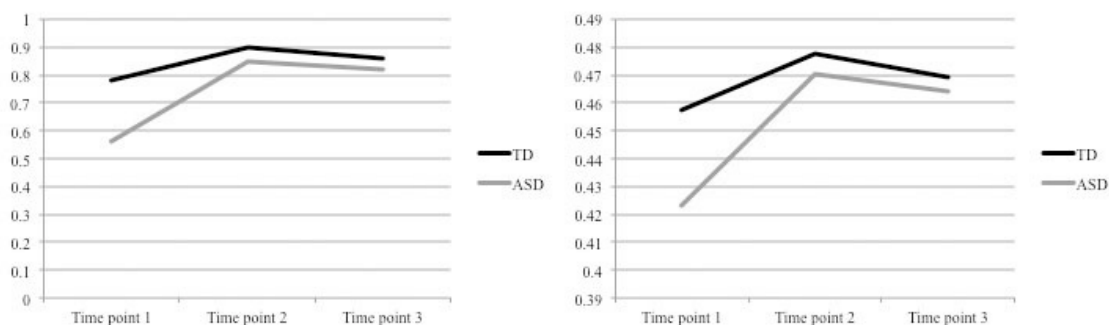

Figure 4. Development of identification across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

Attribution

The development of the attribution component of emotion recognition on ASD and TD children was examined with a 3 (time points 1, 2 and 3 on attribution) x 2 (ASD and TD children) mixed R-ANOVA design. See Figure 5.

Original data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .595, $p < .001$); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.71$). The analysis revealed a significant main effect of time (F (2, 288) = 37.64, $p < .001$) and diagnosis (F (1, 144) =12.25, $p < .001$). Bonferroni corrected comparisons showed that children's scores on the attribution task at time point one were significantly lower than at time point two (F (1,146) = 55.21, $p < .001$) and time point three (F (1, 144) = 33.45, $p < .001$) across both groups. Comparisons also revealed that children's scores on the identification task at time point two

(F (1, 144) = 2.20, p = .14) were significantly higher than at time point three across both groups. Further, there was an interaction between time and diagnosis (F (2, 288) = 8.31, p = .001). Post hoc comparisons revealed that the increase in attribution ability in ASD was steeper than that in TD children from the first to the second (F (1, 144) = 10.62, p = .001) and third time point (F (1, 144) = 8.73, p = .004). However, ASD and TD children did not differ in their increase in identification from the second to the third time point (F (1, 144) = .025, p = .875).

Reciprocal transformation. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .699, p < .001); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = 0.77). The analysis revealed a significant main effect of time (F (2, 288) = 33.05, p < .001) and diagnosis (F (1, 144) = 10.93, p = .001). Bonferroni corrected post hoc comparisons showed that children's scores on the attribution task at time point one were significantly lower than at time point two (F (1, 144) = 51.64, p < .001) and time point three (F (1, 144) = 29.47, p < .001) across both groups. Comparisons also revealed that children's scores on the attribution task at time point two (F (1, 144) = 2.09, p = .15) did not differ from time point three across both groups. Further, there was an interaction between time and diagnosis (F (2, 288) = 6.63, p = .004). Post hoc comparisons revealed that the increase in attribution ability in ASD was steeper than that in TD children from the first to the second (F (1, 144) = 8.99, p = .003) and third time point (F (1, 144) = 7.1, p = .009). However, ASD and TD children did not differ in their increase in identification from the second to the third time point (F (1, 144) = .006, p = .94).
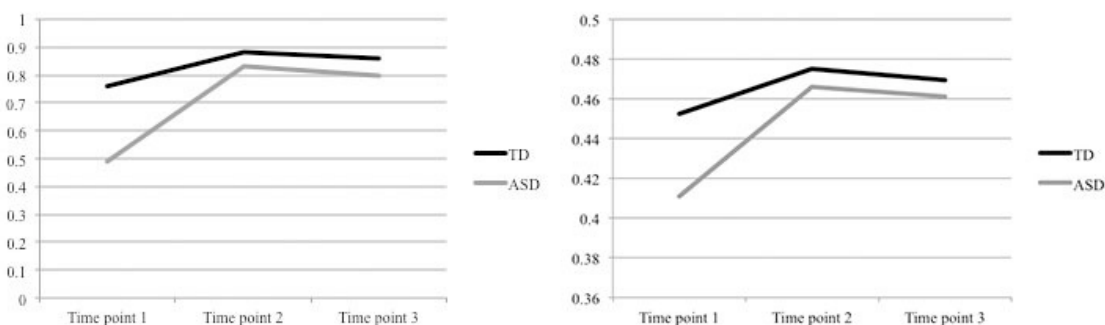

Figure 5. Development of attribution across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

3.3. Development of empathy in ASD and TD children

Affective empathy

The development of the affective component of empathy in ASD and TD children was examined with a 3 (time points 1, 2 and 3 on affective empathy) x 2 (ASD and TD children) mixed R-ANOVA design. See Figure 6.

Original data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .608, *p* < .001); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = 0.72). The analysis revealed a significant main effect of time (*F*(2, 290) = 10.4, *p* < .001) and diagnosis (*F*(1, 145) = 3.91, *p* =.05). Bonferroni corrected post hoc comparisons showed that children's scores on the affective empathy task at time point one were significantly lower than at time point two (*F* (1, 145) = 9.15, *p* = .003), but did not differ scores at time point three (*F* (1, 145) = 3.32, *p* = .07) across groups. Moreover, children's scores on the affective empathy task at time point two (*F* (1, 145) = 13.35, *p* < .001) were significantly higher than at time point three across groups. Further, no interaction was found between time and diagnosis (*F* (2, 290) = .71, *p* = .84).

Reciprocal transformation. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .95, *p* = .03); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = 0.96). The analysis revealed a significant main effect of time (*F* (2, 290) = 4.1, *p* = .019) and diagnosis (*F* (1, 145) = 5.91, *p* = .016). Bonferroni corrected post hoc comparisons showed that children's scores on the affective empathy task at time point one did not differ from time point two across groups (*F* (1, 145) = .12, *p* = .73). However, children's scores in the affective empathy task at time point one (*F* (1, 145) = 6.06, *p* = .015) and time point two (*F* (1, 145) = 5.66, *p* = .019) were significantly lower than at time point three across groups. Further, no interaction was found between time and diagnosis (*F* (2, 290) = .27, *p* = .76).
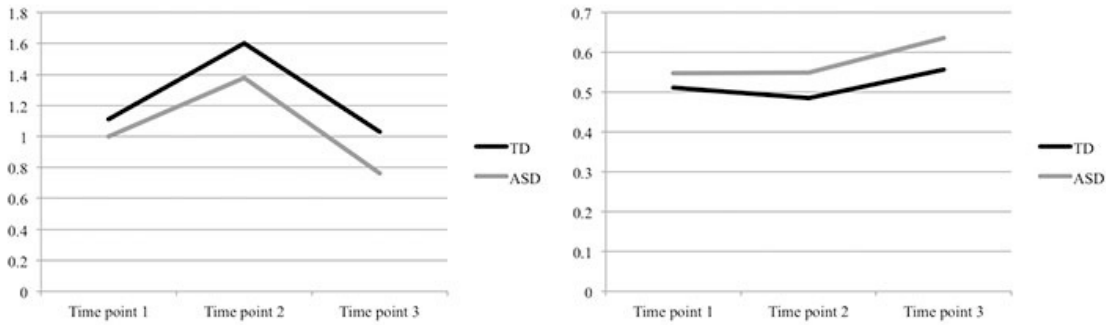
Figure 6. Development of affective empathy across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

Cognitive empathy

The development of the cognitive component of empathy in ASD and TD children was examined with a 3 (time points 1, 2 and 3 on cognitive empathy) x 2 (ASD and TD children) mixed R-ANOVA design. See Figure 7.

Original data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .934, p = .008); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = .94). The analysis revealed a significant main effect of time (F (2, 286) = 2.07, p < .001) and diagnosis (F (1, 143) = 38.22, p < .001). Bonferroni corrected post hoc comparisons showed that children's scores on the cognitive empathy task at time point one were significantly lower than at time point two F (1, 143) = 36.89, p < .001) and time point three (F (1, 143) = 15.38, p

< .001) across both groups. Comparisons also revealed that children's scores on the cognitive empathy task at time point two were significantly higher (F (1, 143) = 4.63, p = .03) than at time point three across both groups. Further, there was an interaction between time and diagnosis (F (2, 286) = 3.38, p = .039). Post hoc comparisons revealed that the increase in cognitive empathy in ASD was significantly steeper from the first to the second time point (F (1, 143) = 5.47, p = .021) and marginally steeper from the first to the third time point (F (1, 143) = 3.364, p = .058) than that in TD children. However, ASD and TD children did not differ in their increase in cognitive empathy from the second to the third time point (F (1, 143) = .107, p = .74).

Reciprocal transformation. Mauchly's test indicated that the assumption of sphericity was not violated (X^2 (2) = .988, p = .43); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = 0.94). The analysis revealed a significant main effect of time (F (2, 286) = 14.97, p < .001) and diagnosis (F (1, 143) = 26.28, p < .001). Bonferroni corrected post hoc comparisons showed that children's scores on the cognitive empathy task at time point one were significantly

lower than at time point two (F (1, 143) = 28.37, p < .001) and at time point three (F (1, 143) = 7.99, p = .005) across groups. Comparisons also showed that children's scores in the cognitive empathy task at time point two were significantly higher than at time point three (F (1, 143) = 7.37, p = .07) across both groups. Further, no interaction effect was found between time and diagnosis (F (2, 286) = 1.68, p = .188).
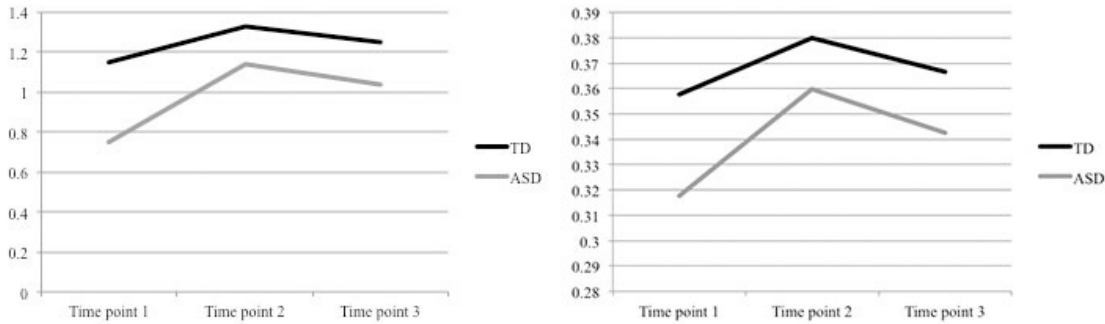

Figure 7. Development of cognitive empathy across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

Prosocial Behavior

The development of the prosocial component of empathy in ASD and TD children was examined with a 3 (time points 1, 2 and 3 on prosocial behavior) x 2 (ASD and TD children) mixed R-ANOVA design. See Figure 8.

Original data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .86, p < .001); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = .88). The analysis revealed a significant main effect of time (F (2, 286) = 16.87, p < .001), however not of diagnosis (F (1, 143) = 1.41, p = .24). Bonferroni corrected post hoc comparisons showed that children's scores in the prosocial behavior task at time point one were significantly lower than at time point two (F (1, 143) = 17.59, p < .001), but did not differ from time point three (F (1, 143) = 3.54, p = .06) across both groups. Comparisons also revealed that children's scores in the prosocial behavior task at time point two were significantly higher (F (1, 143) = 23.46, p < .001) than at time point three across both groups. Further, there was no interaction between time and diagnosis (F (2, 286) = .64, p = .92).

Transformed data. Mauchly's test indicated that the assumption of sphericity was violated (X^2 (2) = .72, p < .001); therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity (ε = .78). The analysis revealed a significant main effect of time (F (2, 286) = 16.22, p < .001), however not of diagnosis (F (1, 143) = 1.54, p = .22). Bonferroni corrected post hoc comparisons showed that children's scores on the prosocial behavior task at time point one were significantly

lower than at time point two (F (2, 143) = 16.66, p < .001), but did not differ from time point three (F (1, 143) = 3.11, p = .08) across both groups. Comparisons also revealed that children's scores in the prosocial behavior task at time point two were significantly higher (F (1, 143) = 20.68, p < .001) than at time point three across groups. Further, there was no interaction between time and diagnosis (F (2, 286) = .21, p = .81).
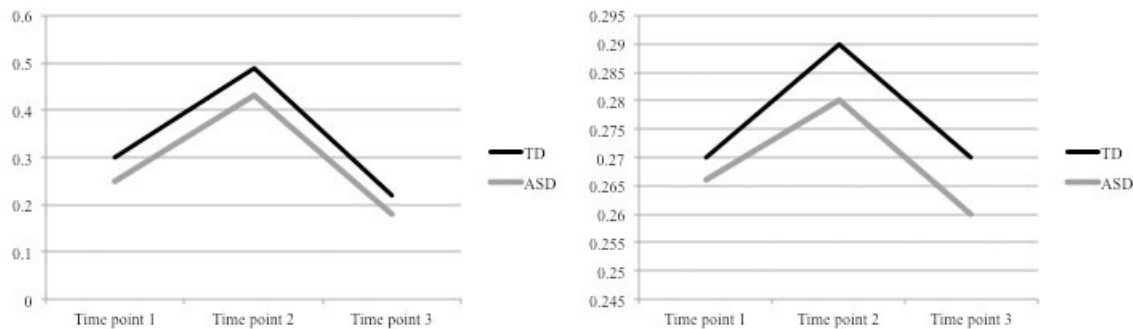

Figure 8. Development of prosocial behavior across three time points of TD and children with ASD using the original data (left) and the transformed data (right)

3.4. Relationship between emotion recognition and empathy

Spearman correlations were calculated to analyze the relation between emotion recognition and empathy components between time points and groups. Using Fisher's r- to-z transformation we compared the correlations between the TD and the ASD group. On the first time point, there was a significant difference between TD and ASD children in the correlation between emotion discrimination and cognitive empathy: whereas TD children showed a significant weak negative correlation (r (108) = -.203, p = .035), children with ASD showed a non-significant positive correlation. The groups also differed in their correlation between emotion identification and cognitive empathy, i.e., whereas TD children showed a significant weak negative correlation (r (106) = -.217, p = .026), children with ASD showed a non-significant positive correlation. Moreover, there was a significant difference between TD and ASD children's correlation between identification and prosocial behavior, i.e., whereas TD children showed a non-significant positive correlation, children with ASD showed a significant moderate positive correlation (r (39) = .434, p = .006). See Table 3.

On the second time point, there was a significant difference between TD and ASD children's correlation between discrimination and cognitive empathy, i.e., whereas TD children showed a non-significant negative correlation, children with ASD showed a significant moderate positive correlation (r (39) = .32, p = .047). Moreover, the groups differed in their correlation between identification and prosocial behavior, i.e., whereas TD

children showed a non-significant negative correlation, children with ASD showed a significant moderate positive correlation (r (39) = .369, p = .021). TD and ASD children also differed in their correlation between attribution and prosocial behavior, i.e., whereas TD children showed a non-significant negative correlation, children with ASD showed a marginally significant moderate positive correlation (r (39) = .314, p = .052). See Table 4.

On the third time point, for both children with and without ASD, emotion recognition was not significantly related to empathy. See Table 5.

Table 3. First time point correlations between emotion recognition and empathy collapsed for groups.

	Discrimination	Identification	Attribution
Affective empathy	-.102	-.108	-.132
Cognitive empathy	-.203*/ .083	-.217*/ .295	-.024
Prosocial behavior	.018	.004/ .434 *	.139

Note. Correlations are provided separately for the TD (left) and ASD (right) group when they differ significantly for the two groups. Otherwise one value for the total group is given.
* Significant at $p < .05$ (2-tailed)

Table 4. Second time point correlations between emotion recognition and empathy collapsed for groups.

	Discrimination	Identification	Attribution
Affective empathy	.043	-.129	-.091
Cognitive empathy	-.142/ .32*	.071	.065
Prosocial behavior	.020	-.09/ .369*	-.081/ .314*

Note. Correlations are provided separately for the TD (left) and ASD (right) group when they differ significantly for the two groups. Otherwise one value for the total group is given.
* Significant at $p < .05$ (2-tailed)

Table 5. Third time point correlations between emotion recognition and empathy collapsed for groups.

	Discrimination	Identification	Attribution
Affective empathy	-.023	-.001	.006
Cognitive empathy	.124	-.013	-.026
Prosocial behavior	.045	-.059	-.127

Note. Fisher r-to-z transformation showed no group differences, therefore one value for the total group is given.

4. Discussion

The aim of this study was twofold. First, we investigated the developmental trajectories of emotion recognition including emotion discrimination, identification, and attribution, and

the developmental trajectories of empathy including affective empathy, cognitive empathy, and prosocial behaviors in children with ASD in comparison to the TD children. Besides, we studied the relationship between emotion recognition and empathy in these two groups. Overall, we found that emotion recognition and empathy improves from the first to the second time point across groups, and that children with ASD display difficulties in emotion recognition and showed lower level of empathy, especially in cognitive empathy. Furthermore, we found different associations between emotion recognition and empathy across groups.

4.1. Developmental trajectory

Regarding the developmental trajectory of emotion recognition and empathy, in general we found that children both with and without ASD improved from time point one to time point two, which conformed to our expectation. However, from time point 2 to 3 their abilities either stayed stable (emotion identification, emotion attribution) or actually showed a decrease (prosocial behaviors). We suggest that the stabilizing and decrease may be caused by an unwanted side effect of testing repeatedly. It is possible that after being tested with similar tasks twice, many children lost their motivation to be fully engaged in the tasks. In the following discussion, we will leave out the results of time point 3, and focus on the change of abilities from time point 1 to 2.

Emotion recognition

Regarding the development of emotion discrimination, as expected we found that both TD and ASD children improved over time. That is, children became consistently better at emotion discrimination from time point one until time point three, which conforms with previous studies that reported an age-related improvement in emotion differentiation (Thomas et al., 2007; Tottenham, et al., 2011).

Regarding the development of emotion identification and attribution, as expected we found that both TD and ASD improved from the first time point to the second time point. This observed increase was of greater magnitude in ASD than in TD children, which portrays a period of ample development for children with ASD in emotion identification and attribution.

Empathy

Regarding the development of affective empathy, no clear conclusions can be made based on our data due to the analyses of the original data and transformed data giving contradictory results for both TD and ASD children. Previous research has pointed out that affective empathy was an innate ability (Rieffe et al., 2010). Indeed, it seems logical to think that affective empathy does not fluctuate much as children grow older if it is an innate ability that is observable already in infants, and as a consequence not greatly influenced by the environment. Nonetheless, the environment can also influence affective empathy in a way that it decreases. For instance, as children grow older, they might learn how to distance themselves from others' distress and how to regulate their own emotional distress. Given the inconclusiveness of the data reported in this study on affective empathy, future research is needed to conclude whether affective empathy changes over time.

Regarding the development of cognitive empathy and prosocial behavior, as expected we found that both TD and ASD children improved from the first to the second time point. The observed increase from the first to the second time point in cognitive empathy and prosocial behaviors supports previous research and likely represents the critical period of development of these components (Baron-Cohen, Leslie, & Frith, 1985; Bensalah et al., 2015; Pouw et al., 2013; Rieffe, et al., 2005; Russell et al., 2012; Wellman 1990).

4.2. Group differences

Emotion recognition

Regarding group differences in emotion discrimination, unexpectedly, children with ASD started off at a similar level of performance at time point one and two as compared to TD children, but they performed at a lower level at time point three. Similarly, Lacroix and colleagues (2011) found no deficits in emotion discrimination in 4 to 8-year-olds with ASD. It is possible that no differences were found between the groups at time point one and two due to a floor effect, that is, children of both groups may have lacked the ability to differentiate emotions. Further, it may be that TD children made a greater improvement than children with ASD at time point three, thus showing a group difference. Future research should further investigate whether children with ASD display a delay in emotion discrimination.

Regarding group differences in emotion identification and attribution, children with ASD as expected performed worse than TD children at time point one. Surprisingly, at time

point two and three both groups performed similarly. These results show that although children with ASD may start off performing poorer, they are able to catch up to their TD counterparts. Our finding differs from Evers and colleagues (2015), who showed emotion identification difficulties in children with ASD aged 6 to 14 year old. An explanation is that Evers and colleagues (2015) used photos of real adults displaying emotions, which may be more challenging than identifying emotions from drawings. Furthermore, previous studies have found that high-functioning individuals with ASD might use their intact cognitive abilities to compensate their deficiency in recognizing emotions, and this may be the case for our participants with ASD at an older age (see review by Harms et al., 2010).

Empathy

Regarding group differences in affective empathy, as expected children with ASD started off at a similar level of performance at time points one and two as compared to TD children. These results support the notion of affective empathy being an innate ability, which do not change much over time. It should be noted that we measured affective arousal by observing the child's emotional arousal in response to different situations. However, in order to confirm that children with ASD do not differ from TD children in affective empathy, more sensitive and objective measures should be used.

Regarding group differences in cognitive empathy, as expected children with ASD showed a developmental delay throughout the three time points as compared to TD children. Our findings add supporting evidence to the literature that children with ASD have lower cognitive empathy than TD children, and they lag behind in understanding ToM (Broekhof et al., 2015; Mazza et al., 2014; Rueda et al., 2015; Schwenck et al., 2012).

Regarding group differences in prosocial behavior, unexpectedly children with ASD displayed prosocial behaviors comparable to those of TD children. This outcome contradicts previous studies that found impairments in prosocial behaviors in children with ASD (Iizuka et al., 2010; Russell et al., 2012). It should be noted that both groups showed minimal amount of prosocial behaviors. The lack of group difference can be due to a floor effect. That is, the task involved an experimenter who was a stranger to the child. It is possible that for children both with and without ASD, they did not really believe the acting of the experimenter, and/or they were not used to providing comfort or help to a stranger. In future studies, a design with higher ecological validity should be used to examine these prosocial behaviors in children with and without ASD.

4.3. Relationship between emotion recognition and empathy

Significant correlations between emotion recognition and empathy were found at time point 1 and 2, but the correlations disappeared at time point 3. Since we are not confident whether the data collected at time point 3 are reliable, in the following discussion we only focused on the correlations found at the previous time points.

For both children with and without ASD, unexpectedly affective empathy was not related to emotion recognition at time point one and two. It is possible that a person can become emotionally aroused as long as he or she detects that the other person is emotionally aroused, without the need to understand what that person is exactly feeling or why the person is feeling so.

We did not find any positive relation between empathy and emotion recognition in TD children at all time points, which comes as a surprise and contradicts previous findings like Rieffe et al. (2010). However, we did find some positive relations between empathy and emotion recognition in the ASD group. For children with ASD, we found that prosocial behaviors were positively related to emotion identification at time point one, and to both emotion identification and attribution at time point two. Besides, cognitive empathy was positively related to emotion discrimination at time point two. It is possible that the empathic ability of children with ASD relies more heavily on emotion recognition than in TD children. Only future research can shed light on this regard.

4.4. Limitations

Needless to say, the current study has several limitations. First, using the same tasks repeatedly at different time points may have caused children to experience diminished motivation and thus influenced their performances. Future research could prevent this by spreading testing throughout a longer period of time, or by administering distinct tasks, which bear the same nature at the different testing sessions. Secondly, the TD group had a larger sample size than the ASD group, which may have hindered the results. Hence, future studies should aim at having equal sample sizes to assure more accurate results. Thirdly, the validity of outcomes reported in this study is based on the concurrence of the original data and the transformed data analyses. Therefore, we need to be cautious when drawing conclusions. A different statistical model such as Linear Mixed Model may be a better tool to analyze the data.

Lastly, we did not check for possible gender differences. Previous studies suggest that girls have better emotion recognition skills than boys (McClure, 2000) as well as showing more empathetic behaviors (Rueckert, 2011). In our study, the TD group had significantly more girls than the ASD group, which could influence the results. For example, the group differences may diminish or disappear if only boys were included in both groups. Future studies should check whether there is a gender effect in emotion recognition and empathy and whether there is an interaction effect of gender and diagnosis.

4.5. Conclusion

The outcomes of the present study show that like TD children, in children with ASD the ability of emotion recognition and empathy tend to improve with age, but yet children with ASD showed some lagging-behind in certain aspects of emotion recognition and empathy, especially in cognitive empathy. Besides, we found that for children with ASD their ability to recognize emotions were associated with their empathy skills. Our findings indicate that early interventions aiming at improving emotion recognition and empathy skills in children with ASD may benefit their overall development.

References

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., & 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.

Bensalah, L., Caillies, S., & Anduze, M. (2015). Links Between Cognitive Empathy, Theory of Mind, and Affective Perspective Taking by Young Children. The Journal of genetic psychology, (just-accepted), 00-00.

Bird, G., & Viding, E. (2014). The self to other model of empathy: providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520-532.

Broekhof, E., Ketelaar, L., Stockmann, L., Zijp, A., Bos, M. G., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of autism and developmental disorders*, *45*(7), 2035-2045.

Camodeca, M., & Coppola, G. (2016). Bullying, empathic concern, and internalization of rules among preschool children: The role of emotion understanding. *International Journal Of Behavioral Development*, *40*(5), 459-465. doi:10.1177/0165025415607086

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, *9*(4), 428-449. doi:10.1177/1362361305056082

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in cognitive sciences*, *16*(4), 231-239.

Clark, T. F., Winkielman, P., & McIntosh, D. N. (2008). Autism and the extraction of emotion from briefly presented facial expressions: stumbling at the first step of empathy. *Emotion*, *8*(6), 803.

Dacety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews*, *3*(2), 71-100.

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with

autism spectrum disorder, developmental delay, and typical development. *Child development*, *73*(3), 700-717.

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006). Are abilities abnormally interdependent in children with autism?. *Journal of Clinical Child and Adolescent Psychology*, *35*(1), 20-33

Evers, K., Steyaert, J., Noens, I., & Wagemans, J. (2015). Reduced recognition of dynamic facial emotional expressions and emotion-specific response bias in children with an autism spectrum disorder. *Journal Of Autism And Developmental Disorders*, *45*(6), 1774-1784. doi:10.1007/s10803-014-2337-x

Ketelaar, L., Rieffe, C., Wiefferink, C. H., & Frijns, J. H. (2013). Social competence and empathy in young children with cochlear implants and with normal hearing. *The Laryngoscope*, *123*(2), 518-523.

Hadjikhani, N., Zürcher, N. R., Rogier, O., Hippolyte, L., Lemonnier, E., Ruest, T., …& Helles, A. (2014). Emotional contagion for pain is intact in autism spectrum disorders. *Translational psychiatry*, *4*(1), e343.

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychology review*, *20*(3), 290-322.

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, *45*(7), 1185-1198.

Iizuka, C., Yamashita, Y., Nagamitsu, S., Yamashita, T., Araki, Y., Ohya, T., & Matsuishi, T. (2010). Comparison of the Strengths and Difficulties Questionnaire (SDQ) scores between children with high-functioning autism spectrum disorder (HFASD) and attention-deficit/hyperactivity disorder (AD/HD). *Brain & Development*, *32*(8), 609-612. doi:10.1016/j.braindev.2009.09.009

Lacroix, A., Guidetti, M., Rogé, B., & Reilly, J. (2014). Facial emotion recognition in 4- to 8-year-olds with autism spectrum disorder: A developmental trajectory approach. *Research In Autism Spectrum Disorders*, *8*(9), 1146-1154. doi:10.1016/j.rasd.2014.05.012

Markham, R., & Adams, K. (1992). The effect of type of task on children's identification of facial expressions. *Journal of Nonverbal Behavior*, *16*(1), 21-39.

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.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, *126*(3), 424.

Netten AP, Rieffe C, Theunissen SCPM, Soede W, Dirks E, Briaire JJ, et al. (2015) Low Empathy in Deaf and Hard of Hearing (Pre) Adolescents Compared to Normal Hearing Controls. PLoS ONE 10(4): e0124102. doi:10.1371/journal. pone.0124102.

Pouw, L. B., Rieffe, C., Oosterveld, P., Huskens, B., & Stockmann, L. (2013). Reactive/proactive aggression and affective/cognitive empathy in children with ASD. *Research in developmental disabilities*, *34*(4), 1256-1266.

Ridgeway, D., Waters, E., & Kuczaj, S. A., (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21,* 901-908.

Rieffe, C., Ketelaar, L., & Wiefferink, C.H. (2010). Assessing empathy in young children; construction and validation of an empathy questionnaire (EmQue). *Personality and Individual Differences, 49,* 362–367.

Rieffe, C., Terwogt, M. M., & Cowan, R. (2005). Children's understanding of mental states as causes of emotions. *Infant and Child Development*, *14*(3), 259-272.

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.

Rueckert, L. (2011). Gender differences in empathy. *Psychology of empathy*, 221-234.

Rueda, P., Fernandez-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.

Russell, G., Golding, J., Norwich, B., Emond, A., Ford, T., & Steer, C. (2012). Social and behavioural outcomes in children diagnosed with autism spectrum disorders: A longitudinal cohort study. *Journal Of Child Psychology And Psychiatry*, *53*(7), 735-744. doi:10.1111/j.1469-7610.2011.02490.x

Russo-Ponsaran, N. M., Evans-Smith, B., Johnson, J., Russo, J., & McKown, C. (2016). Efficacy of a facial emotion training program for children and adolescents with autism spectrum disorders. *Journal Of Nonverbal Behavior*, *40*(1), 13-38. doi:10.1007/s10919-015-0217-5

Schulte-Rüther, M., Greimel, E., Piefke, M., Kamp-Becker, I., Remschmidt, H., Fink, G. R., ... & Konrad, K. (2014). Age-dependent changes in the neural substrates of empathy in autism spectrum disorder. *Social cognitive and affective neuroscience*, *9*(8), 1118-1126.

Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., & Freitag, C. M. (2012). Empathy in children with autism and conduct disorder: Group-specific profiles and developmental aspects. *Journal Of Child Psychology And Psychiatry*, *53*(6), 651-659. doi:10.1111/j.1469-7610.2011.02499.x

Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, *16*(3), 241-256. doi:10.1017/S0142716400007281

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, *10*(5), 547-558. doi:10.1111/j.1467-7687.2007.00614.x

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in psychology*, *2*, 39.

Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, *19*(1), 1-29.

Trevisan, D. A., & Birmingham, E. (2016). Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis. *Research in Autism Spectrum Disorders*, *32*, 24-42.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: a formal meta-analysis. *Journal of autism and developmental disorders*, *43*(7), 1517-1526.

Wellman, H.M. (1990). The Child's Theory of Mind. MIT Press: Cambridge, MA. Wiefferink, C. H., Rieffe, C., Ketelaar, L., De Raeve, L., & Frijns, J. H. (2013).

Emotion understanding in deaf children with a cochlear implant. *Journal of deaf studies and deaf education, 18*(2), 175-186.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology, 28*(1), 126-136. doi:10.1037/0012-1649.28.1.126