CrossMark

ORIGINAL PAPER

Reduced Recognition of Dynamic Facial Emotional Expressions and Emotion-Specific Response Bias in Children with an Autism Spectrum Disorder

Kris Evers · Jean Steyaert · Ilse Noens ·
Johan Wagemans

Abstract Emotion labelling was evaluated in two matched samples of 6–14-year old children with and without an autism spectrum disorder (ASD; $N = 45$ and $N = 50$, resp.), using six dynamic facial expressions. The Emotion Recognition Task proved to be valuable demonstrating subtle emotion recognition difficulties in ASD, as we showed a general poorer emotion recognition performance, in addition to some emotion-specific impairments in the ASD group. Participants' preference for selecting a certain emotion label, irrespective of the stimulus presented, played an important role in our results: response bias-corrected data still showed an overall decreased emotion recognition performance in ASD, but no emotion-specific impairments anymore. Moreover, ASD traits and empathy were correlated with emotion recognition performance.

Keywords Autism spectrum disorder (ASD) · Dynamic facial expressions · Emotion recognition · Face perception · Response bias · Theory of mind

K. Evers (✉) · J. Wagemans
Laboratory of Experimental Psychology, KU Leuven,
Tiensestraat 102, Box 3711, 3000 Leuven, Belgium
e-mail: kris.evers@ppw.kuleuven.be

K. Evers · J. Steyaert
Department of Child Psychiatry, UPC-KU Leuven, Leuven,
Belgium

K. Evers · J. Steyaert · I. Noens · J. Wagemans
Leuven Autism Research (LAuRes), KU Leuven, Leuven,
Belgium

J. Steyaert
Department of Clinical Genetics, University Hospital Maastricht,
Maastricht, The Netherlands

I. Noens
Parenting and Special Education Research Unit, KU Leuven,
Leuven, Belgium

I. Noens
Psychiatric and Neurodevelopmental Genetics Unit,
Massachusetts General Hospital, Boston, MA, USA

Introduction

Adequately recognizing emotional expressions is crucial for a successful participation in everyday social life. Social impairments are considered hallmark symptoms of autism spectrum disorders (ASD; American Psychiatric Association 2000, 2013) and early signals of ASD (Volkmar et al. 2005) suggest that social impairments are—at least partially—related to the recognition and interpretation of facial expressions.

Many researchers have studied the processing of facial emotional expressions in ASD, yielding inconclusive and highly mixed findings (for an overview of the literature, see Harms et al. 2010). Both demographic and task- and stimulus-related factors are considered to contribute to existing inconsistencies. Highly heterogeneous samples were examined with a broad variety of tasks and paradigms, ranging from rather implicit measures of emotional saliency, such as sorting or matching tasks, to explicit emotion recognition paradigms. Moreover, different emotional expressions were evaluated at different levels of emotional intensity: stimuli varied from well-controlled static isolated facial expressions, towards ecologically valid stimuli, such as interacting individuals in a social context. In spite of the large methodological differences

between studies, a recent meta-analysis provided evidence for an emotion recognition impairment in ASD (Uljarevic and Hamilton 2013), with neurotypical individuals mostly outperforming the ASD sample when recognizing fear, sadness, surprise, and disgust. Indeed, the recognition impairment in ASD seems to be emotion-specific, with mostly negative or lower-intensity expressions being affected in ASD (e.g., Bal et al. 2010; Humphreys et al. 2007; Philip et al. 2010; Smith et al. 2010; Wallace et al. 2011).

To account for the issue of small sample sizes (as mentioned by Uljarevic and Hamilton 2013), and in order to explicitly compare different emotional expressions at different levels of intensity, we applied the Emotion Recognition Task (Kessels et al. 2014; Montagne et al. 2007) to a large sample of 6–14-year old children with and without a clinical diagnosis of ASD. The Emotion Recognition Task is a widely-used and laboratory-controlled emotion labelling paradigm with six dynamic facial expressions, presented at different emotional intensities (ranging from subtle to full-blown expressions). The task already proved to be valuable in detecting subtle emotion recognition peculiarities in several clinical (e.g., Ammerlaan et al. 2008; Egger et al. 2013; Montagne et al. 2008; Verhoeven et al. 2008) and non-clinical samples (West et al. 2012). Furthermore, the paradigm demonstrated its usefulness in identifying facial emotion recognition impairments in adolescents with a clinical diagnosis of ASD (12–14-year olds in Kessels et al. 2010; 12-to-19-year olds in Smith et al. 2010) and in non-clinical adults with ASD traits (Poljac et al. 2013). Given that neurocognitive studies often examine school-aged children with ASD, a first aim of this study was to extend previous findings with the Emotion Recognition Task to a large sample of school-aged children with ASD, and a matched comparison sample.

Our second aim was to achieve a greater understanding in the underlying cognitive strategies, involved in emotion labelling tasks. Despite mixed evidence, it is nowadays generally accepted that individuals with ASD experience difficulties reading facial expressions, at least to some extent. However, little is known about their cognitive strategies or about the errors they make during Emotion Recognition Tasks. Only recently, researchers started to investigate the error patterns of individuals with ASD during Emotion Recognition Tasks. Most of these studies described remarkably similar error patterns in individuals with ASD and control samples, with confusions between anger (or surprise) and disgust, between sadness and anger, and between fear and surprise, being most frequently reported (Castelli 2005; Humphreys et al. 2007; Jones et al. 2011; Wallace et al. 2011). In addition, Eack et al. (2014) showed that adults with ASD tended to misinterpret neutral

expressions and assigned those a negative valence, more so than the comparison group did.

In an attempt to achieve a greater understanding of the poorer emotion recognition skills in children with ASD, and in the underlying cognitive processes, we wanted to investigate another aspect of the error patterns in ASD. In particular, we investigated whether there was an atypical response bias in ASD. It could be that the ASD sample selected certain emotion labels more frequently, and others almost never, irrespective of the presented emotional stimulus. These response bias data provided us with a baseline to take into account in our further analysis, hereby enabling us to evaluate (1) the impact of the response bias on emotion recognition performances in general, and, more specifically, (2) the contribution of response biases in atypical emotion recognition in individuals with ASD.

Thirdly, we evaluated the relationship between emotion recognition performance and demographic variables, to provide insight (1) in the significance of experimentally measured emotion recognition skills for daily life social behaviour, and (2) in the importance of variables such as ASD traits, empathy, age and intelligence for emotion recognition. In addition, recent demonstrations of minor emotion recognition problems in first-degree relatives of individuals with ASD (Bolte and Poustka 2003; Losh et al. 2009; Palermo et al. 2006; Wallace et al. 2010), spurred on evidence about the presence of ASD symptoms in the general population. Indeed, a non-clinical sample of adults with high ASD traits exhibited poorer emotion recognition in comparison to individuals with low ASD traits (Poljac et al. 2013). Moreover, small yet significant correlations between empathy scores and emotion recognition performances were found in individuals with ASD, their first-degree relatives, and in a control sample (Sucksmith et al. 2013). Only a few studies have directly evaluated the development of emotion recognition skills with age, in ASD, mostly demonstrating atypical developmental trajectories (e.g., Gepner et al. 2001; Kuusikko et al. 2009; O'Connor et al. 2005; Rump et al. 2009). Nevertheless, a recent meta-analysis (Uljarevic and Hamilton 2013) did not reveal an overall influence of demographic factors such as age and intelligence on emotion recognition problems.

In summary, we compared emotion recognition performance in a large sample of school-aged children with and without ASD using the Emotion Recognition Task. This study had a three-fold aim. First, we wanted to evaluate emotion labelling skills of six dynamic facial expressions, presented at high and intermediate emotional intensity. Based on the previous studies using this paradigm, we expected the ASD children to perform worse than the typically developing (TD) children, at least on some emotional expressions (e.g., anger, disgust, fear or surprise), and especially at the more subtle level of emotional

intensity. Second, we wanted to evaluate the contribution of response biases to the previously reported difficulties: what is the impact of participants' preference for certain emotion labels over others? Third, we evaluated the association between emotion recognition skills and demographic factors, such as age, intelligence and ASD traits.

Methods

Participants

Two groups of 6–14-year old children without intellectual disability (IQ ≥ 70) participated in this study: a sample of children with ASD and a control group, consisting of TD children. None of these participants was using neuroleptics and all of them were Dutch-speaking and reported normal or corrected-to-normal vision. Intelligence was assessed using an abbreviated version (Sattler 2001) of the Wechsler Scales for Intelligence (WISC-III-NL; Wechsler 1992), comprising the following subtests: Block Design, Similarities, Picture Completion, and Vocabulary.

The ASD group consisted of 50 children diagnosed in a multidisciplinary team or by a child psychiatrist, according to DSM-IV-TR criteria (American Psychiatric Association 2000). They were recruited via a special needs school for children with ASD (n = 17), or via the Autism Expertise Centre at the university hospital (n = 33). These clinical diagnoses were confirmed within the experimental protocol using the Autism Diagnostic Observation Scale (ADOS; Gotham et al. 2007) in 46 out of 50 cases, with an average severity score of 5.30.[1] The TD group comprised 68 children, recruited via mainstream schools. None of these TD children suffered from a known psychiatric disorder or language disorder, nor did any of them had a first-degree family member with ASD (information gathered from the parents). ASD traits were measured with the Dutch version of the Social responsiveness Scales (SRS; Roeyers et al. 2012), with an average raw score far below T70 in the TD group (T70 refers to a norm score of 2 SD above the mean, corresponding to a raw total score of 60 for boys and of 55 for girls). In addition, a Dutch translation of the Empathizing-Systemizing questionnaire (Auyeung et al. 2008) was administered to all parents. The Dutch translation was created by the researchers using a procedure of translation and back-translation. Out of the initial 118 participants, two participant groups were selected to create two groupwise matched participant groups, which were matched based on age, intelligence level and had an equal gender

ratio (for detailed participants' characteristics, we refer to Table 1).

Stimuli

To obtain high experimental control, the emotional stimuli consisted of computer-generated morphs of coloured static pictures, with a build-up of emotional expressions appearing very similar to natural expressions. More details about the stimulus construction can be found in Montagne et al. (2007). For this study, we used dynamic facial expressions of four Caucasian actors (two males, two females), each showing six 'basic' emotions (Ekman and Cordaro 2011), namely anger, disgust, fear, happiness, sadness, and surprise. Based on piloting and previous work in adults (Poljac et al. 2013), two levels of emotional intensity were used: an intermediate intensity level, from a neutral face to 50 % emotional intensity (0–50 %) and a high intensity level, from neutral faces towards a full-blown emotional expression (0–100 %; depicted in Fig. 1). The duration of the animation was approximately 1 or 2 s (intermediate and high intensity, resp.).

Procedure: The Emotion Recognition Task

For pragmatic reasons, this study was incorporated within a larger, randomly ordered series of experiments on visual and emotional processing, with breaks in-between two experiments. These studies were approved by the ethical committee of the university hospital. With written informed consent of the parents and the participants themselves (if older than 12 years), participants were individually tested in a quiet room. Afterwards, participants received a small present and parents got mileage allowance, if applicable. The procedure was similar in both participant groups.

We investigated emotion recognition with the Emotion Recognition Task (Montagne et al. 2007), which is a widely-used, laboratory-controlled instrument. This computer-based emotion labelling task was displayed on a laptop with a 15 inch screen and a viewing distance of approximately 57 cm. The step-by-step-practice protocol started with introducing the task and emotion labels. To ensure that all participants understood the emotion labels, they were asked to provide an example situation for the different emotions. If participants could not provide an appropriate example, a standard example was given, and they were asked to come up with a second one. Then, a paper version of the task was demonstrated and only if all questions were overcome, the experimental protocol began. Children were instructed to carefully watch the video clip of an emotional face, and to select the appropriate emotion label. In each trial, a dynamic facial emotional expression

---

[1] ADOS severity scores are calibrated scores ranging between one and ten, with scores of four and five considered representative of an ASD classification (Gotham et al. 2009).

Table 1 Participant characteristics of the ASD and typically developing (TD) group

	Whole groups			Matched groups		
	ASD ($n = 50$) 47 M:3 F M (SD)	TD ($n = 68$) 58 M:10 F M (SD)	p value	ASD ($n = 45$) 42 M:3 F M (SD)	TD ($n = 50$) 47 M:3 F M (SD)	p value
Age	129.82 (28.13)	122.31 (28.74)	.1596	126.78 (27.78)	123.28 (28.69)	.5484
FSIQ	100.78 (13.16)	110.07 (9.98)	<.0001	103.24 (11.36)	106.65 (8.87)	.1099
VIQ	101.62 (16.91)	112.10 (12.09)	.0003	104.82 (14.58)	107.46 (9.99)	.3121
PIQ	99.94 (14.74)	108.03 (12.38)	.0016	101.67 (14.19)	105.84 (11.95)	.1232
SRS	87.33 (23.61)	28.73 (18.19)	<.0001	89.46 (23.67)	30.93 (19.86)	<.0001
ADOS	5.31 (1.65)	N/A	N/A	5.30 (1.71)	N/A	N/A
EQ	18.89 (7.58)	34.20 (8.60)	<.0001	18.82 (7.82)	32.62 (8.85)	<.0001
SQ	25.37 (7.91)	24.47 (6.29)	.5334	26.00 (8.10)	24.16 (6.46)	.2636

Note Means (M), standard deviations (SD) and *p* values (two-tailed distribution) of a two-sample *t* test (homoscedasticity assumption checked with *F* test) for age (depicted in months), full-scale, verbal, and performance IQ (FSIQ, VIQ and PIQ, resp.), SRS raw total scores, ADOS severity scores, Empathy Quotients (EQ) and Systemizing Quotients (SQ), for both participant groups. SRS questionnaires of 13 children were either missing or incomplete (ASD: $n = 7$; TD: $n = 6$; ASD match: $n = 6$; TD match: $n = 4$) and 6 out of 62 TD children scoring above the T70 score. EQ-SQ questionnaires of 20 children were either missing or incomplete (ASD: $n = 12$; TD: $n = 8$; ASD match: $n = 8$; TD match: $n = 6$)

was presented, always starting from a neutral face, with an incrementally increasing emotional intensity. During the course of a trial, six emotion labels were presented on the left side of the display, with the dynamic emotional stimulus on the right side (see Fig. 2), and participants had an unlimited response time to select the appropriate label with the mouse. The last frame of the animated expression remained on the screen while the participant responded, and then a new trial was automatically initiated.

Design

Emotion recognition was evaluated in a group of children with and without ASD, using an emotion labelling task with six emotional expressions (anger, disgust, fear, happiness, sadness and surprise), shown at two different levels of emotional intensity (high and intermediate). In total, this study comprised 48 trials: 24 trials with intermediate emotional intensity (6 emotional expressions × 4 identities) and 24 trials with high emotional intensity (6 emotional expressions × 4 identities). Both intensity levels were presented in two blocks (high-intensity block always preceding the intermediate-intense one), with a small break in-between,

and a randomized trial order within one block. We decided to present the high-intensity levels first so that all participants could get a clear idea of the kinds of emotions shown.

Data Analysis

All analyses were performed with the statistical package SAS 9.3. Average performance scores were calculated per participant for each emotion × intensity combination, resulting in 12 performance scores per participant. No outliers were detected, defined as participants with a performance score of two SDs above or below the group's mean performance. Due to technical problems, seven participants completed only half of the trials (six participants completed only the medium intensity trials, and one participant completed only the high intensity trials). Therefore, a mixed models ANOVA was performed, which can handle empty cells. We included participant group (between-subjects factor), emotion and intensity (within-subjects factors) as fixed factors, and a random intercept (for each participant). The main effects, all two-way interactions and the three-way-interaction effect were included in the model. Post-hoc exploration contrasts were
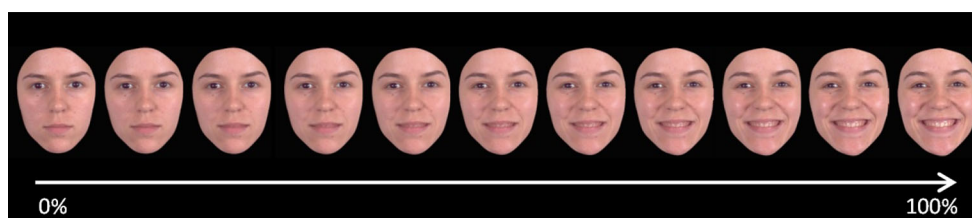
Fig. 1 Stimulus build-up, demonstrating the gradual transition from a neutral expression towards a highly intense facial emotional expression
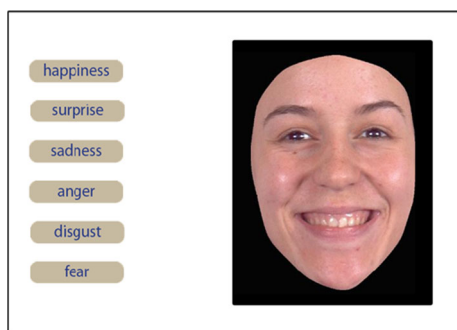
Fig. 2 Trial example. In the Emotion Recognition Task, participants had to label dynamic facial expressions (presented on the *right side*). Dutch labels were presented (on the *left side*), with the instructor reading the labels out loud in case participants were no fluent readers yet

Tukey–Kramer-corrected, and significant interaction effects were explored using simple means analyses.

To correct for the individual's response bias (see Table 2), we have also analysed corrected performance scores. For each participant, we calculated the response bias for the different emotion labels, indicating how often they selected a certain label, irrespective of the display shown. This resulted in a score between 0 and 100 for each emotion label, such that the response biases for all labels together added to 100 for each participant. Then, we divided a participant's performance score (for each emotional expression x intensity level) by the response bias of the participant for the specific emotional expression. In that way, equal performance scores combined with different response biases resulted in different corrected performances (for a concrete example, see footnote[2]).

To test the reliability of our results, all analyses were repeated within several (sub)groups of the participants, such as the whole group, matched groups, excluding TD children with a SRS score above T70, or excluding those participants with incomplete datasets. As these analyses all yielded essentially the same results, we will only report results for the matched participant samples here. Interested readers are referred to Supplementary Material for more details.

A significance level of $p < .05$ (two-sided) was adopted, and effects with a p value between .05 and .10 were indicated as marginally significant.

---

[2] An example: Participant X and Y have both labelled all of the high intenseangry trials correctly (performance = 100 %). However, participant X alwaysresponded 'anger', independent of the stimulus presented (response bias = 100 %), and participant Y only responded 'anger' when an angry face waspresented (response bias = 16.67 %). Consequently, the corrected performance ofparticipant X (100/100 = 1.00) will be lower than the corrected performance ofparticipant Y (100/16.67 = 6.00).

Results

Emotion Recognition Abilities in Children With and Without ASD

Given a chance level of 17 % (one out of six), overall performances were quite high, with the comparison group (M = .56, SD = .08) generally outperforming the ASD sample [M = .53, SD = .09, Cohen's $d = .36$; $F(1, 93) = 4.75$, $p = .0318$]. As expected, full-blown expressions were labelled more correctly than expressions presented at medium intensity level [$F(1, 88) = 31.96$, $p < .0001$]. A main effect of emotion was found [$F(5, 465) = 154.56$, $p < .0001$; depicted in Fig. 3], indicating that happiness and anger were labelled most accurately, whereas fear and sadness were more difficult to label (the only non-significant pair-wise contrast was between fear and sadness; Tukey–Kramer adjusted $p = .9993$). Moreover, the interaction between intensity and emotion was significant [$F(5, 440) = 3.04$, $p = .0103$]. Post-hoc analyses revealed that intensity level only affected performance when labelling anger, sadness and surprise (simple main effect analysis, $p \le .0033$).

In addition to the generally poorer emotion recognition in the ASD sample, we also observed a significant interaction between participant group and emotion [$F(5, 465) = 3.24$, $p = .0070$; see Fig. 4]. Post-hoc analyses revealed that the TD group particularly outperformed the ASD group when labelling sadness ($p = .0112$) and tended to perform better when labelling disgust ($p = .0856$) and surprise ($p = .0866$; simple main analyses). For the angry and happy emotional expressions, no group differences were found ($p = .1953$ and $p = .2344$, resp.). However, the ASD group performed better than the TD group when recognising fear ($p = .0158$; simple main effect analysis). The two-way interaction between intensity and participant group nor the three-way-interaction between intensity, emotion and participant group were significant (both $F < 1$). Including age or FSIQ in the mixed analysis did not change these results.

Error Patterns and Response Bias in Both Participant Groups

Despite remarkably similar confusion matrices (see Table 2), we found evidence for a difference in response bias between both participant groups [Rao-Scott corrected $\chi^2(5) = 28.04$, $p < .0001$, see Table 3]. Especially the label 'fear' was selected more often in the ASD group than in the TD group, which might explain the worse performance of the TD group at fearful faces (see section *Emotion Recognition Abilities in Children with and without ASD*).

Table 2 Confusion tables for the matched ASD and TD sample

	ASD sample						TD sample					
	Emotional expression (correct answer)											
	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Participants' response												
Anger	**275**	*157*	7	1	32	1	**314**	*153*	9	1	40	0
Disgust	*49*	**175**	46	5	*89*	17	*41*	**214**	47	6	*100*	19
Fear	9	8	**140**	1	*93*	35	7	7	**115**	1	54	15
Happiness	2	0	11	**302**	5	*165*	1	0	16	**343**	4	*175*
Sadness	10	12	11	1	**103**	1	11	10	8	0	**147**	3
Surprise	11	4	*141*	46	34	**137**	10	0	*189*	33	39	**172**

Values represented the number of responses. Correct responses were shown bold-faced on the diagonal, and most common errors were in italics

Fig. 3 Mean performances for the different emotional expressions at two levels of emotional intensity (intermediate and high), indicated that happiness and surprise were recognised most correctly, and that fear and sadness were the most difficult expressions. (*Error bars* represent 95 % confidence intervals of the means)


Fig. 4 In addition to the generally poorer emotion recognition performance in the ASD sample, this emotion recognition problem seemed to be emotion-specific, with the ASD especially being outperformed by the TD group when labelling sadness, disgust and surprise. Contrary to our expectations, the ASD group performed better in comparison to the TD group when recognizing fear. (*Error bars* represent 95 % confidence intervals of the means.)

Table 3 The selected response labels, irrespective of the presented emotional expression, in children with and without ASD

Selected response label	Participant group					
	ASD		TD		*t* test	*p* value
	n	%	*n*	%		
Anger	473	22.14	517	22.44	.01	.9946
Disgust	381	17.84	427	18.53	−.59	.5554
Fear	286	13.39	199	8.64	3.73**	.0003
Happiness	485	22.71	539	23.39	−.51	.6119
Sadness	138	6.46	179	7.77	−1.43	.1567
Surprise	373	17.46	443	19.23	−1.14	.2590

Response label selection was determined for each participant, and both the number of times selected (*n*) and its corresponding proportion (%) were depicted. Then, a two-samples *t* test was calculated (homoscedasticity assumption checked with an *F* test), with ** indicating statistically significant effects, applying the Bonferroni correction for multiple comparisons (e.g., lowering the critical α to .0083). A completely random selection of response labels would result in a proportion of 16.67 %

To correct for these differences in response bias, we repeated our mixed ANOVA analysis for the corrected performance scores (see section *Data and Data Analysis* for more details about this measure). Most of our results were highly similar: the main effects of emotion [$F(5, 458) = 38.75$, $p < .0001$] and emotional intensity [$F(1, 88) = 70.56$, $p < .0001$] were significant, as was the interaction effect between emotional intensity and emotion [$F(5, 436) = 11.66$, $p < .0001$]. Again, the interaction effects between participant group and intensity ($F < 1$), and the three-way-interaction between participant group, emotion and intensity were not significant [$F(5, 436) = 1.63$, $p = .1519$]. Although the TD group still generally outperformed the ASD group after correcting for response bias [$F(1, 93) = 5.59$, $p = .0202$; see Fig. 5], the
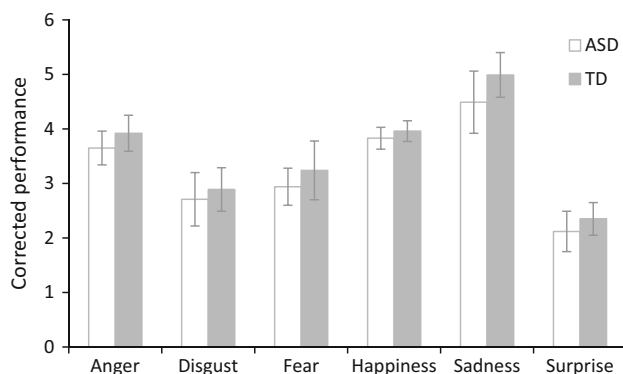
Fig. 5 After taking into account response bias, the ASD sample showed an overall poorer emotion recognition performance, but no evidence for an emotion-specific deficit was found. (*Error bars* represent 95 % confidence intervals of the means.)

interaction effect between participant group and emotion now disappeared ($F < 1$; see Fig. 5).

Correlation Between Emotion Recognition and Participant Characteristics

Pearson correlations between emotion recognition performance and demographic factors were reported across groups, although a similar pattern was found when analysing the ASD and TD group separately, unless explicitly mentioned. No correlation was found between overall emotion recognition performance and age ($r = .08$, $p = .4467$) or FSIQ ($r = .01$, $p = .9251$). Interestingly, emotion recognition performances were negatively correlated with SRS scores ($r = -.28$, $p = .0096$; depicted in Fig. 6), such that children with more ASD traits (i.e., higher SRS scores) had an overall lower performance on the Emotion Recognition Task. Within-group analyses yielded correlation patterns in the same direction, but none of them was significant (ASD group: $r = -.17$, $p = .2880$; TD group: $r = -.13$, $p = .3781$), probably due to the smaller sample sizes. In addition, Empathizing Questionnaire scores (EQ) seemed correlated with mean performances ($r = .24$, $p = .0362$), such that children with higher empathy performed more accurately at the Emotion Recognition Task. Within-group correlations, however, were not significant, and suggested that this correlation was driven by an association between empathy and emotion recognition in the comparison group ($r = .20$, $p = .1942$), but not in the ASD group ($r = -.06$, $p = .7287$).

Discussion

We evaluated emotion recognition performance in a large sample of school-aged children with ASD and a matched
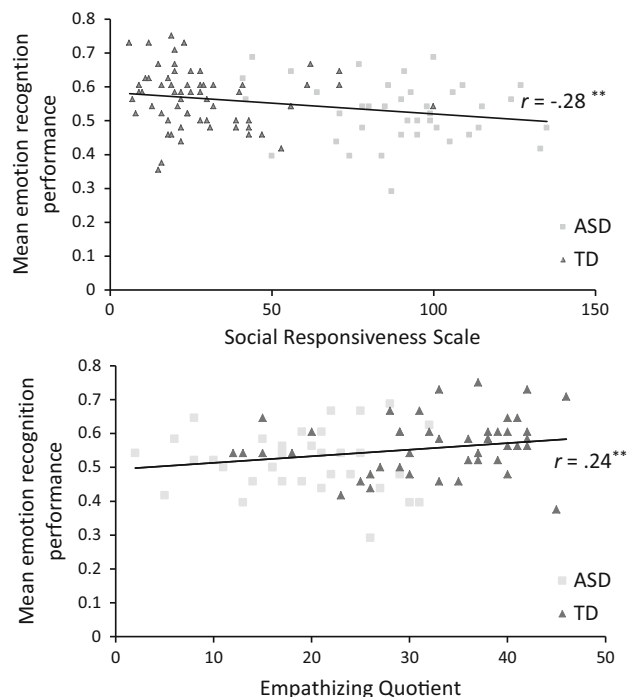
Fig. 6 Mean performance scores for both participant groups were plotted against raw SRS scores (*above*) and Empathizing Quotient scores (EQ; *below*), to show the relationship between emotion recognition performance and ASD traits. (The *full line* shows the linear trend line of the relationship between SRS scores or EQ scores and emotion recognition performance across both participant groups.)

comparison group, using the Emotion Recognition Task, a labelling paradigm with dynamic facial expressions. Hereby, we extended previous research using this controlled paradigm to another age group of individuals with ASD. Secondly, we examined the impact of response bias on performance. And thirdly, we investigated the relationship between emotion recognition skills and demographic factors.

The 6-to-14-year old ASD group shows a poorer emotion recognition performance on the Emotion Recognition Task (Montagne et al. 2007). Complementing and extending previous findings on this emotion labelling task, we evaluated the value of this paradigm in detecting subtle performance differences in this clinical population. Hence, we investigated the usability of this instrument for neurocognitive studies, both in the context of clinical diagnostics and of scientific research. Indeed, neurocognitive research on emotion recognition mostly investigated school-aged children with ASD, and therefore it seemed important to apply the Emotion Recognition Task in that population too. Our findings were highly comparable to those of other studies using the Emotion Recognition Task in adolescents or adults with ASD, which all revealed emotion recognition impairments in ASD, some of them rather subtle, others more severe (Kessels et al. 2010; Poljac et al. 2013; Smith

et al. 2010). Furthermore, the emotion recognition disadvantage in ASD had a moderate effect size (.36) in our study, which at first sight seemed much smaller than the effect size reported in a recent meta-analysis (Uljarevic and Hamilton 2013). However, after applying a correction for publication bias, the reported effect size in that meta-analysis decreased to .41 (Uljarevic and Hamilton 2013), which corresponds with the size of our effect. We tested a rather large sample (matched groups: ASD: $N = 45$, TD: $N = 50$), and nevertheless, recognition problems in ASD appear to be rather small and subtle. Note, however, that affect perception in daily life implies the multimodal integration of different sources of—sometimes noisy or incongruent—information, e.g., auditory and visual information about facial expressions and body postures (Aviezer et al. 2012). Therefore, emotion recognition problems, seemingly minor and negligible within the context of a well-controlled experiment, could have a large impact on social life and on daily life functioning.

Unsurprisingly, and not the prominent focus of this study, some emotions such as e.g., happiness and anger, were generally identified more accurately than others. This is in line with findings about the developmental trajectories of facial emotion recognition, which are known to be emotion-dependent: the explicit labelling of happiness occurs most accurately and earliest in development, and surprise and fear latest (Herba et al. 2006; Thomas et al. 2007). The poorer performance of the ASD group in our study also appeared emotion-specific, with mostly the recognition of sadness, disgust and surprise being impaired, paralleling findings showing that especially the identification of negative expressions were affected (e.g., Ashwin et al. 2006; and also see review: Harms et al. 2010). Previous applications of the same emotion recognition paradigm to other age groups of individuals with (sub-)clinical ASD have revealed emotion-specific impairments too. However, there is some disagreement about the specific kinds of expressions affected in ASD, with only the identification of disgust generally found impaired (Kessels et al. 2010; Poljac et al. 2013; Smith et al. 2010), but a poorer identification of anger (Poljac et al. 2013; Smith et al. 2010), surprise (Poljac et al. 2013; Smith et al. 2010), and sometimes fear (Kessels et al. 2010) was demonstrated too. Remarkably, and contrary to these previous findings, our ASD sample recognised fear more accurately than the comparison group.

High intensity expressions had an overall recognition advantage in comparison to those presented at medium intensity. Moreover, the intensity effect was stronger for certain emotions, such as anger, sadness and surprise (which has been demonstrated in previous research, e.g., Herba et al. 2006). The emotion recognition problem in ASD was not especially profound during intermediate intensity trials. This is not in line with previous findings, showing that individuals with ASD have a decreased sensitivity for emotional expressions or are particularly impaired in recognizing more subtle emotional expressions, in comparison to full-blown ones (Humphreys et al. 2007; Wallace et al. 2011; even when using the Emotion Recognition Task: Smith et al. 2010). However, we want to emphasize that this of course does not exclude individuals with ASD experiencing more difficulties recognizing subtle emotional expressions. Poorer emotion recognition in laboratory conditions is often only found under highly ecologically valid circumstances, especially in persons with a higher intelligence, which have access to cognitively-mediated compensation mechanisms. Future research should evaluate more systematically the impact of more subtle levels of emotional intensity on recognition performance, in both participant groups, using well-controlled paradigms such as the Emotion Recognition Task, as well as more complex materials.

In an attempt to provide a greater insight in the underlying cognitive strategies involved in emotion recognition difficulties in individuals with ASD, we have examined the importance of a response bias in the context of our emotion labelling paradigm. More concretely, we investigated how often the emotion labels were selected, irrespective of the presented expression. Our results provided evidence for a different response bias in both participant groups. Notably, the ASD group especially selected the label 'Fear' more frequently than the TD group did, which could explain the better performance of the ASD group at recognizing fearful expressions. We then corrected performance data using the response bias as *baseline*. Hereby we evaluated the contribution and impact of response biases on emotion recognition performance, and on the atypical pattern in the ASD sample. Having applied this correction, the overall poorer emotion recognition in ASD remained, but no evidence for an emotion-specific deficit was found anymore, indicating that this response bias substantially contributed to the emotion-specificity of the recognition disadvantage in the ASD group. Despite the vast number of studies focusing on emotion recognition in ASD, researchers have only recently started to look into their confusion patterns. Comparisons of the nature of errors in ASD and in neurotypicals have mostly reported rather similar mistakes in ASD and comparison groups, with more confusions between anger (or surprise) and disgust, between sadness and anger, and between fear and surprise (Castelli 2005; Humphreys et al. 2007; Jones et al. 2011; Wallace et al. 2011). However, incorporating neutral expressions in the stimulus set, adults with ASD tended to assign negative valences to neutral faces more often than controls (Eack et al. 2014). The impact of a response bias on emotion recognition in the neurotypical population has often been

neglected. To our knowledge, the contribution of response biases on the atypical performance in individuals with ASD has not been studied before. Based on our findings, we want to emphasize the possible impact of this bias, maybe even contributing to the inconsistencies in the literature (Harms et al. 2010), and certainly to the contradicting findings on the emotion-specificity of the emotion recognition difficulties in ASD. Indeed, when different studies use different types of emotions within their stimulus sets, the response biases may lead to different performance levels for different emotions.

In our study, participants' age and IQ were both not correlated with emotion recognition performance. Although some previous studies provided evidence for a different developmental trajectory of emotion recognition in ASD (e.g., Gepner et al. 2001; Kuusikko et al. 2009; O'Connor et al. 2005; Rump et al. 2009), a recent meta-analysis (Uljarevic and Hamilton 2013) did not reveal an atypical influence of moderating factors such as age and intelligence on emotion recognition performance. The lack of a relationship between age and emotion recognition in our study, seems to suggest that no significant maturation of emotion recognition abilities occurs between the age of 6 and 14 year, at least not for these basic emotional expressions. However, the absence of an effect of age or intelligence on emotion recognition does not exclude the impact of (language) maturation, and of the development of an emotion vocabulary on the identification of (subtle) emotional expressions (Harms et al. 2010). But probably the specific impact of the demographics is strongly influenced by specific characteristics of a study, e.g., the age range of the sample, the task demands and the complexity of the emotional stimuli. In addition, the absence of an association between age and emotion recognition in a cross-sectional study, does of course not exclude the impact of maturation on an individuals' performance.

Emotion recognition skills were correlated with SRS and EQ scores, indicating that participants with more ASD traits and with less empathy performed poorer at the Emotion Recognition Task. This small yet significant relationship between ASD traits and emotion recognition problems is in line with the more dimensional perspective on ASD. Besides, it parallels previous findings of the correlation between ASD traits and associated symptoms (Bolte and Poustka 2006; Losh et al. 2009; Palermo et al. 2006; Poljac et al. 2013; Wallace et al. 2010), or between empathy and emotion recognition performance (Sucksmith et al. 2013), both within and beyond the ASD population. We note, however, that effect sizes were rather small and that within-group correlations never reached the significance threshold. Whereas the association between ASD traits (SRS scores) and emotion recognition had the same direction and effect size in the ASD and TD samples, this

was not the case for the correlation between empathy and emotion recognition. There, the association was not present in the ASD sample, but the comparison group seemed to drive the across-groups correlation. Although the interpretation of the reported correlations warrants caution, these findings suggest that facial emotion recognition could be relevant in the search for neurocognitive endophenotypes (Rommelse et al. 2011), attempting to provide insight into the link between genetics and behaviour.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th edition text revision)*. Washington, D.C.: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, D.C.: Author.

Ammerlaan, E. J. G., Hendriks, M. P. H., Colon, A. J., & Kessels, R. P. C. (2008). Emotion perception and interpersonal behavior in epilepsy patients after unilateral amygdalohippocampectomy. *Acta Neurobiologiae Experimentalis, 68*(2), 214–218.

Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: A test of the amygdala theory. *Social Neuroscience, 1*(3–4), 349–363. doi:10.1080/17470910601040772.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The Autism spectrum Quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders, 38*(7), 1230–1240. doi:10.1007/s10803-007-0504-z.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225–1229. doi:10.1126/science.1224313.

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders, 40*(3), 358–370. doi:10.1007/s10803-009-0884-3.

Bolte, S., & Poustka, F. (2003). The recognition of facial affect in autistic and schizophrenic subjects and their first-degree relatives. *Psychological Medicine, 33*(5), 907–915. doi:10.1017/S0033291703007438.

Bolte, S., & Poustka, F. (2006). The broader cognitive phenotype of autism in parents: How specific is the tendency for local processing and executive dysfunction? *Journal of Child Psychology and Psychiatry, 47*(6), 639–645. doi:10.1111/j.1469-7610.2006.01603.x.

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

Eack, S. M., Mazefsky, C. A., & Minshew, N. J. (2014). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism* (Epub ahead print). doi:10.1177/1362361314520755.

Egger, J. I. M., Wingbermuhle, E., Verhoeven, W. M. A., Dijkman, M., Radke, S., de Bruijn, E. R. A., & Koolen, D. (2013). Hypersociability in the behavioral phenotype of 17q21.31 microdeletion syndrome. *American Journal of Medical Genetics Part A, 161A*(1), 21–26. doi:10.1002/ajmg.a.35652.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3*(4), 364–370. doi:10.1177/1754073911410740.

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders, 31*(1), 37–45. doi:10.1023/A:1005609629218.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS Scores for a Measure of Severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705. doi:10.1007/s10803-008-0674-3.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*(4), 613–627. doi:10.1007/s10803-006-0280-1.

Harms, M., 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. doi:10.1007/s11065-010-9138-6.

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry, 47*(11), 1098–1106. doi:10.1111/j.1469-7610.2006.01652.x.

Humphreys, K., Minshew, N., Leonard, G. L., & Behrmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia, 45*(4), 685–695. doi:10.1016/j.neuropsychologia.2006.08.003.

Jones, C. R. G., Pickles, A., Falcaro, M., Marsden, A. J. S., Happe, F., Scott, S. K., & Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 52*(3), 275–285. doi:10.1111/j.1469-7610.2010.02328.x.

Kessels, R. P. C., Montagne, B., Hendriks, A. W., Perrett, D. I., & de Haan, E. H. F. (2014). Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8-75. *Journal of Neuropsychology, 8*(1), 75–93. doi:10.1111/jnp.12009.

Kessels, R. P. C., Spee, P., & Hendriks, A. W. (2010). Perception of dynamic facial emotional expressions in adolescents with autism spectrum disorders (ASD). *Translational Neuroscience, 1*(3), 228–232. doi:10.2478/v10134-010-0033-8.

Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M.-L., Ebeling, H., & Moilanen, I. (2009). Emotion recognition in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(6), 938–945. doi:10.1007/s10803-009-0700-0.

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological profile of autism and the Broad Autism Phenotype. *Archives of General Psychiatry, 66*(5), 518–526.

Montagne, B., de Geus, F., Kessels, R. M., Denys, D., de Haan, E. H. F., & Westenberg, H. G. M. (2008). Perception of facial expressions in obsessive-compulsive disorder: A dimensional approach. *European Psychiatry, 23*(1), 26–28. doi:10.1016/j.eurpsy.2007.07.007.

Montagne, B., Kessels, R. P. C., De Haan, E. H. F., & Perrett, D. I. (2007). The Emotion Recognition Task: A paradigm to measure the perception of facial emotional expressions at different intensities. *Perceptual and Motor Skills, 104*(2), 589–598. doi:10.2466/Pms.104.2.589-598.

O'Connor, K., Hamm, J. P., & Kirk, I. J. (2005). The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. *Brain and Cognition, 59*(1), 82–95. doi:10.1016/j.bandc.2005.05.004.

Palermo, M. T., Pasqualetti, P., Barbati, G., Intelligente, F., & Rossini, P. M. (2006). Recognition of schematic facial displays of emotion in parents of children with autism. *Autism, 10*(4), 353–364. doi:10.1177/1362361306064431.

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine, 40*(11), 1919–1929. doi:10.1017/S0033291709992364.

Poljac, E., Poljac, E., & Wagemans, J. (2013). Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits. *Autism, 17*(6), 668–680. doi:10.1177/1362361312455703.

Roeyers, H., Thys, M., Druart, C., De Schryver, M., & Schittekatte, M. (2012). *SRS: Screeningslijst voor autismespectrum stoornissen.* Amsterdam: Hogrefe.

Rommelse, N. N. J., Geurts, H. M., Franke, B., Buitelaar, J. K., & Hartman, C. A. (2011). A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience and Biobehavioral Reviews, 35*(6), 1363–1396. doi:10.1016/j.neubiorev.2011.02.015.

Rump, K. M., Giovannelli, J. L., Minshew, N. J., & Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child Development, 80*(5), 1434–1447.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M Sattler Publisher Inc.

Smith, M. J. L., Montagne, B., Perrett, D. I., Gill, M., & Gallagher, L. (2010). Detecting subtle facial emotion recognition deficits in high-functioning autism using dynamic stimuli of varying intensities. *Neuropsychologia, 48*(9), 2777–2781. doi:10.1016/j.neuropsychologia.2010.03.008.

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia, 51*(1), 98–105. doi:10.1016/j.neuropsychologia.2012.11.013.

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.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders,.* doi:10.1007/s10803-012-1695-5.

Verhoeven, W., Wingbermuhle, E., Egger, J., Van der Burgt, I., & Tuinier, S. (2008). Noonan syndrome: Psychological and psychiatric aspects. *American Journal of Medical Genetics Part A, 146A*(2), 191–196. doi:10.1002/ajmg.a.32115.

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology, 56*, 315–336. doi:10.1146/annurev.psych.56.091103.070159.

Wallace, G. L., Case, L. K., Harms, M. B., Silvers, J. A., Kenworthy, L., & Martin, A. (2011). Diminished sensitivity to sad facial

expressions in high functioning autism spectrum disorders is associated with symptomatology and adaptive functioning. *Journal of Autism and Developmental Disorders, 41*(11), 1475–1486. doi:10.1007/s10803-010-1170-0.

Wallace, S., Sebastian, C., Pellicano, E., Parr, J., & Bailey, A. (2010). Face processing abilities in relatives of individuals with ASD. *Autism Research, 3*(6), 345–349. doi:10.1002/aur.161.

Wechsler, D. (1992). *Wechsler intelligence scale for children* (3rd ed.). London: The Psychological Corporation.

West, J. T., Horning, S. M., Klebe, K. J., Foster, S. M., Cornwell, R. E., Perrett, D., & Davis, H. P. (2012). Age effects on emotion recognition in facial displays: From 20 to 89 years of age. *Experimental Aging Research, 38*(2), 146–168. doi:10.1080/0361073X.2012.659997.