Emotion recognition across visual and auditory modalities in autism spectrum disorder: A systematic review and meta-analysis

Florence Yik Nam Leung[a], Jacqueline Sin[b], Caitlin Dawson[a], Jia Hoong Ong[a], Chen Zhao[a], Anamarija Veić[a], & Fang Liu[a,*]

[a]School of Psychology and Clinical Language Sciences, University of Reading, Earley Gate, Reading, RG6 6AL, UK

[b]School of Health Sciences, City, University of London, London EC1V 0HB

Running title: Emotion recognition in autism spectrum disorder

*Corresponding author at:

School of Psychology and Clinical Language Sciences
University of Reading
Earley Gate

Reading RG6 6AL

UK

E-mail address: f.liu@reading.ac.uk (F. Liu)

Highlights

- We examined emotion recognition in ASD across visual and auditory domains.
- ASD is associated with general emotion recognition deficits for human faces.
- Emotion-specific deficits were observed for music and speech prosody in ASD.
- Autistic adults showed more pronounced deficits in emotion recognition than children.
- Deficits were consistently observed for verbal but not nonverbal tasks in ASD.

Abstract

An expanding literature has investigated emotion recognition across visual and auditory modalities in autism spectrum disorder (ASD). Findings, however, have been highly variable. The present work systematically reviewed and quantitatively synthesised a large body of literature, in order to determine whether autistic individuals differ from their neurotypical counterparts in emotion recognition across human face, nonhuman face, speech, and music domains. To identify eligible studies, the literature was searched using Embase, Medline, PubMed, Web of Science, and Google Scholar. Synthesising data from 72 papers, results showed a general difficulty with emotion recognition accuracy in ASD, while autistic individuals also showed longer response times than their neurotypical (NT) counterparts for a subset of emotions (i.e., anger, fear, sadness, and the six-emotion composite). These impairments were shown to be robust as they were not driven by differences in stimulus presentation time restriction and IQ matching, though the severity of impairments was less pronounced for a subset of emotions when full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust) had been undertaken. The heterogeneity among studies arose from a combination of sample characteristics (i.e., age but not IQ) and experimental design (i.e., stimulus domain and task demand) parameters. Specifically, we show that (i) impairments were more pronounced in autistic adults; (ii) full-scale, verbal, and nonverbal IQ did not moderate impairments; (iii) emotion-general impairments were found for human faces but emotion-specific impairments were observed for speech prosody (i.e., anger, happiness, and disgust) and music (i.e., fear and sadness), while no impairment was observed for nonhuman faces; (iv) impairments were found across emotions for verbal but not nonverbal tasks. Importantly, further research on the recognition of prosodic, musical, and nonhuman facial emotions is warranted, as the current findings are disproportionately influenced by studies on human faces. Future studies should also continue to explore the different emotion processing strategies employed by autistic individuals, which could be fundamental to promoting fulfilling emotional experiences in real life.

Keywords: Autism; Emotion recognition; Faces; Speech prosody; Music; Systematic review; Meta-analysis

1. Introduction

Autism spectrum disorder (ASD) is a complex, pervasive neurodevelopmental disorder with a prevalence of approximately 1 in 100 individuals globally (Alshaban et al., 2019; Baron-Cohen, 2009; Kim et al., 2011; Sun et al., 2019). According to DSM-5 and ICD-11, impairment in social communication and interaction is a core diagnostic feature of ASD (American Psychiatric Association, 2013; World Health Organisation, 2018); this lack of sophisticated understanding of nonverbal communicative functions has been reported as particularly central and persistent across developmental levels in ASD (Seltzer, Shattuck, Abbeduto, & Greenberg, 2004; Shattuck et al., 2007). Yet, empirical studies of emotion recognition in ASD have produced mixed results. Using meta-analyses, previous reviews suggest emotion recognition impairments in ASD in the visual modality, including human faces and body gestures (Lozier, Vanmeter, & Marsh, 2014; Uljarevic & Hamilton, 2013). However, there is a lack of consensus about the general emotion recognition ability in ASD in other communication domains, such as nonhuman faces (e.g., Brosnan, Johnson, Grawmeyer, Chapman, & Benton, 2015; Isomura, Ogawa, Yamada, Shibasaki, & Masataka, 2014b), speech prosody (e.g., Heikkinen et al., 2010; Schelinski & von Kriegstein, 2019), and music (e.g., Järvinen et al., 2016; Quintin, Bhatara, Poissant, Fombonne, & Levitin, 2011). In addition, questions about whether impairments vary by age, IQ, stimulus domain, task demand, and emotion have also been raised previously (Harms, Martin & Wallace, 2010; Nuske, Vivanti, & Dissanayake, 2013). Thus, in this review, we sought to provide comprehensive answers to (i) whether the ability to recognise emotions differs between autistic individuals and neurotypical (NT) individuals across different domains, and (ii) whether age, IQ, stimulus domain, and task demand moderate any identified impairments. Given that “autistic” and “neurotypical” are the preferred terms of the autistic community (Kenny et al., 2016), we will use these terms rather than “with ASD” or “typically developing” throughout the text.

1.1. Past findings of emotion processing in ASD

1.1.1. Stimulus domain

Research on emotion has predominantly focused on six basic emotions, including anger, fear, happiness, sadness, disgust, and surprise (Ekman & Friesen, 1976; Ekman, Friesen, & Ellsworth, 1972; Prinz, 2004). Emotions can be communicated nonverbally through different channels, such as human faces (Etcoff & Magee, 1992; Palermo & Rhodes, 2007; Russell, 2003), nonhuman faces (Donath, 2001; Rosset et al., 2008), speech prosody (Juslin & Laukka, 2003; Koolagudi & Rao, 2012), and music (Argstatter, 2016; Mohn, Argstatter, & Wilker, 2011). Within the visual modality, difficulty in identifying and understanding emotions through facial expressions is the most common social-cognitive impairment in ASD (Eack, Mazefsky, & Minshew, 2015; Griffiths et al., 2019; Pelphrey et al., 2002; see Harms et al., 2010 for a discussion), despite contradictory evidence suggesting otherwise (Jones et al., 2011; Tracy, Robins, Schriber, & Solomon, 2011). In contrast, the ability to recognise emotions from nonhuman facial stimuli such as cartoons, caricatures, and schematic faces appears intact in ASD (Brosnan, Johnson, Grawmeyer, Chapman, & Benton, 2015; Isomura et al., 2014b; Miyahara, Bray, Tsujii, Fujita, & Sugiyama, 2007; Rosset et al., 2008), possibly because of autistic individuals' greater or restricted interest in these stimuli (Anthony et al., 2013; Grelotti et al., 2005; Kuo, Orsmond, Coster, & Cohn, 2014; Rosset et al., 2008; Spiker, Lin, Van Dyke, & Wood, 2012).

Within the auditory modality, studies investigating emotional prosody in ASD have yielded mixed findings, with some reporting clear impairments (e.g., Doi et al., 2013; Schelinski & von Kriegstein, 2019; Taylor, Maybery, Grayndler, & Whitehouse, 2015) and others reporting no impairments (e.g., Baker, Montgomery, & Abramson, 2010; Heikkinen et al., 2010). These discrepancies may in part be explained by variations in the methodologies used across studies, such as presenting prosodic stimuli that elicit emotions (in)congruent to its verbal content or in neutral semantics (Wang & Tsao, 2015), implementing low-pass filtering methods that eliminate verbal content (Grossman, Bemis, Plesa Skwerer, & Tager-Flusberg, 2010), and presenting emotional prosody in the presence of contextual cues (Le Sourn-Bissaoui, Aguert, Girard, Chevreuil, & Laval, 2013). Although emotions in music and speech prosody are expressed using similar acoustic cues (Coutinho & Dibben, 2013; Juslin & Laukka, 2003), recognition of musical emotions (such as recognising happy or sad emotions from instrumental musical excerpts) seems unimpaired in ASD (Gebauer, Skewes, Westphael, Heaton, & Vuust, 2014; Heaton, Hermelin, Pring, 1999), especially when verbal IQ was statistically controlled for (Quintin et al., 2011). This is not surprising since music has been widely documented as a domain of preserved abilities in autistic individuals (Heaton, 2009; Mottron, Peretz, & Ménard, 2000; Quintin, Bhatara, Poissant, Fombonne, & Levitin, 2013). Taken together, the findings from different domains suggest that the emotion recognition ability of autistic individuals may be moderated by communication domain, a potential contributing factor to the mixed results in the literature.

1.1.2. Specific emotions

Alongside studies showing general impairments across all emotions in ASD (Lindner & Rosén, 2006; Sawyer, Williamson, & Young, 2012; Song, Zhong, Jia, & Liang, 2020), other research suggests that emotion recognition impairments in ASD may be specific to certain emotions. In the face domain, specific impairments are reported in the recognition of negative emotions, particularly for fearful (Ashwin, Chapman, Colle, & Baron-Cohen, 2006; Pelphrey et al., 2002; Tell, Davidson, & Camras, 2014; Uono, Sato, & Toichi, 2011, 2013; Wallace, Coleman, & Bailey, 2008) and sad expressions (Ashwin et al., 2006; Boraston, Blakemore, Chilvers, & Skuse, 2007). According to the amygdala theory of autism, the poor recognition of fear and other negative emotions is a result of amygdala dysfunction in ASD (Ashwin et al., 2006; Baron-Cohen et al., 2000; Howard et al., 2000). However, in the speech prosody domain, emotion-specific impairments have been noted in the recognition of happiness, a positive emotion (Hubbard, Faso, Assmann, & Sasson, 2017; Wang & Tsao, 2015). Moreover, no general or specific impairments have been reported with nonhuman faces (Brosnan et al., 2015; Rosset et al., 2008) nor music (Järvinen et al., 2016; Quintin et al., 2011). Thus, the specific impairments for negative emotions in the face domain do not seem to generalise to the nonhuman face, speech prosody, and music domains, which might be attributed to the varying difficulty levels associated with the encoding and decoding of different expressions across domains. Indeed, certain emotions are more robustly expressed in one domain compared to the other in NT (Keltner, Tracy, Sauter, Cordaro, & McNeil, 2016). For instance, happiness is the most accurately recognised emotion from human faces, followed by anger and sadness, with fear being the least accurately recognised (Elfenbein & Ambady, 2002). Conversely, across the speech and music domains, anger and sadness are better recognised than fear and happiness (Chronaki, Wigelsworth, Pell, & Kotz, 2018; Elfenbein & Ambady, 2002; Juslin & Laukka, 2003). Thus, if emotion-specific difficulties were indeed present in ASD, it is plausible that they would manifest differently across domains. The current review will examine this possibility by analysing each emotion type separately across and within domains.

1.1.3. Age

The ability to explicitly label emotional expressions emerges during early childhood and develops through adolescence to adulthood (De Sonneville et al., 2002; Herba & Phillips, 2004). Several cross-sectional studies on human faces and speech prosody provide some evidence that emotion recognition improves less over time in ASD than in NT (Gepner, Deruelle, & Grynfeltt, 2001; Kuusikko et al., 2009; Van Lancker, Cornelius, & Krieman, 1989). Specifically, children in both ASD and NT groups show similar improvement and perform comparably on emotion recognition (Rump et al., 2009; Van Lancker et al., 1989). However, autistic adolescents do not seem to develop proficiency beyond those present by late childhood, while NT adolescents continue to refine emotion recognition skills (Greimel et al., 2014; Rump et al., 2009; Van Lancker et al., 1989). Consequently, autistic adults do not reach the level of proficiency demonstrated by NT adults (Rump et al., 2009). The different developmental trajectories between ASD and NT groups may, therefore, have contributed to the inconsistent group differences observed in the literature on human faces and speech prosody. However, studies investigating emotion recognition of nonhuman faces have indicated intact performance in both autistic children and adolescents (Brosnan et al., 2015; Miyahara et al., 2007; Rosset et al., 2008), and the same for music across autistic children, adolescents, and adults (Heaton et al., 1999; Järvinen et al., 2016; Quintin et al., 2011). These findings thus suggest that age does not moderate emotion recognition ability of autistic individuals in the nonhuman face and music domains. Given the contradictory findings across domains, this review will evaluate whether age is a moderating factor for the heterogeneity in the literature across different domains.

1.1.4. IQ

Emotion recognition encompasses both verbal and nonverbal intellectual functioning: while nonverbal ability enables the perception and integration of nonverbal characteristics that differentiate emotion expressions, verbal ability allows the interpretation and assignment of an appropriate emotional label to their meaning (Davitz & Beldoch, 1964). Indeed, intellectual ability has not only been identified as a significant correlate of emotion recognition ability but has also shown significant predictive power in NT (Khawar, Malik, Maqsood, Yasmin, & Habib, 2013). Previous studies have indicated that the correlation between IQ and emotion recognition also extends to ASD, with IQ being an important predictor of emotion recognition ability regardless of diagnosis (full-scale IQ: Jones et al., 2011; nonverbal IQ: Salomone, Bulgrelli, Thommen, Rossini, & Molina, 2019). However, it has also been shown that the association between IQ and emotion recognition is significantly more prominent in the ASD group compared to the NT group (verbal IQ: Dyck, Piek, Hay, Smith, & Hallmayer, 2006), or that this association is uniquely present in the ASD group only but not in the NT group (full-scale IQ: Tanaka et al., 2012; verbal IQ: Atkinson, 2009; Quintin et al., 2011; Wallace et al., 2008). These findings suggest that emotion recognition may involve higher-level analytical processes that autistic individuals could employ to aid performance, whereas NT individuals may opt for intuitive rather than analytical strategies (Grossman, Klin, Carter, & Volkman, 2000). In particular, verbal IQ may contribute to the intact performance of the ASD group in studies requiring explicit knowledge about emotional labels (see Trevisan & Birmingham, 2016 for a review). IQ may, therefore, constitute a compensatory mechanism for emotion recognition especially in autistic individuals with higher cognitive functioning (Harms et al., 2010), which may explain individual differences in emotion recognition within ASD (see Nuske et al., 2013 for a review). In this sense, autistic individuals with higher IQ may be relatively less impaired in emotion recognition when compared against their NT counterparts, i.e., higher IQ of the ASD group may be associated with smaller/no group differences.

Nonetheless, contradictory findings have also been reported in previous studies. Rommelse et al. (2015) examined three groups of ASD and NT participants with below average, average, or above average IQ, and found that in absolute terms, autistic individuals with below average IQ indeed performed worse than those with higher IQ on facial and prosodic emotion recognition tasks. However, in relative terms, a larger group difference in performance on these tasks was found between the ASD and NT groups with above average IQ than the groups with lower IQ (Rommelse et al., 2015). These findings, thus, suggest that autistic individuals with higher IQ may be more severely impaired in emotion recognition when compared against their NT counterparts, i.e., higher IQ of the ASD group may be associated with larger group differences. Furthermore, in other studies, IQ has been shown to be unrelated to autistic individuals' ability to recognise emotions (full-scale IQ: Heaton, 2012; Parron et al., 2008), discounting the potential moderating role of IQ on group differences in emotion recognition. Given the disparate findings regarding the role of IQ in moderating performance by autistic individuals relative to their NT counterparts, the present review sought to examine whether the different measures of IQ (namely full-scale IQ, verbal IQ, and nonverbal IQ) could account for the heterogeneity among results across studies.

1.1.5. Task demand

The majority of emotion recognition studies in ASD implemented a verbal task, involving identifying, labelling, discriminating, matching, or detecting different emotions using verbal cues. Such tasks are cognitively demanding as they place great reliance on emotion vocabulary necessary for labelling emotions following perception (Palermo, O'Connor, Davis, Irons, & McKone, 2013). Findings from studies employing a verbal task have been inconsistent. Implementing a forced-choice task that involves participants selecting their response to the target emotion from pre-generated options, some studies reported significant group differences (e.g., Griffiths et al., 2019; Philip et al., 2010), while others did not (e.g., Baker et al., 2010; Rhodes et al., 2018). Studies using an open-ended labelling task, where participants spontaneously generate an emotional label that describes the target emotion, have similarly reported inconsistent findings (Boucher, Lewis, & Collis, 2000; Castelli, 2005; Hobson Ouston, & Lee, 1989). Verbal discrimination tasks, requiring participants to judge whether a stimulus displays the same or different emotion than the given label, have consistently reported group differences (e.g., Oerlemans et al., 2014; Waddington et al., 2018). Verbal matching tasks involving participants selecting a stimulus among a set that matches the given label have also reported significant group differences (e.g., Loth et al., 2018). However, verbal detection tasks which require participants to detect the stimulus displaying the target emotion among other distractors, have found no group differences (e.g., Shafritz, Bregman, Ikuta, & Szeszko, 2015).

In contrast to verbal tasks, nonverbal emotion recognition tasks tap only into the perceptual stage where emotions are discriminated based on perceptual properties alone (Adolphs, 2002; Palermo et al., 2013). Despite the seemingly reduced task demands, findings have also been inconsistent. Nonverbal discrimination tasks, requiring participants to indicate whether the stimuli within a pair display the same or different emotions, have reported group differences with human faces (e.g., Greimel et al., 2014; Sasson, Shasteen, & Pinkham, 2016b; Vannetzel, Chaby, Cautru, Cohen, & Plaza, 2011), but not with speech prosody (e.g., Lindström et al., 2018). Nonverbal matching tasks, where participants are shown an emotional stimulus and then choose a stimulus from a set that displays the same expression, have found group differences (e.g., Philip et al., 2010; Tanaka et al., 2012). By contrast, nonverbal detection tasks, which involve participants detecting the odd emotion expression among other distractors, have found no group differences (e.g., Isomura et al., 2014b; Kujala, Lepistö, Nieminen-Von Wendt, Näätänen, & Näätänen, 2005). Upon these mixed results, the moderating role of task demands in accounting for heterogeneity in the literature is yet to be elucidated.

1.2. Prior reviews of emotion perception in ASD

To date, three systematic reviews have investigated emotion recognition in ASD (Harms et al., 2010; Lozier et al., 2014; Uljarevic & Hamilton, 2013). The two meta-analyses, bringing together 48 papers on facial and body expressions (Uljarevic & Hamilton, 2013) and 43 papers on facial expressions (Lozier et al., 2014), revealed general emotion recognition deficits in ASD with varying severity across emotions. Age did not moderate emotion recognition in Uljarevic & Hamilton (2013) but moderated recognition of fear, sadness, and disgust in Lozier et al. (2014). The magnitude of deficits could not be accounted for by either full-scale IQ (Lozier et al., 2014; Uljarevic &Hamilton, 2013) or verbal IQ (Lozier et al., 2014). The effect of task on emotion recognition was examined in Uljarevic and Hamilton (2013) and was not found to be a significant moderator; that is, there appeared no overall differences in performance between emotion matching (i.e., requiring perceptual demands) and emotion labelling (i.e., requiring verbal and perceptual demands) tasks. To our knowledge, no reviews have yet evaluated emotion recognition in ASD across domains (human faces, nonhuman faces, speech prosody, and music) and modalities (visual and auditory). In addition, the effects of domain within modalities have not been systematically examined, especially when impairments in the human face and speech prosody domains do not seem to generalise to the nonhuman face nor to the music domain. Reviewing the growing body of literature, particularly with the increased attention to emotion recognition in the auditory domain during the past decade (Järvinen et al., 2016; Schelinski & von Kriegstein, 2019; Taylor et al., 2015; Waddington et al., 2018; Wang & Taso, 2015), may therefore further our understanding of the general emotion recognition ability in ASD as a whole.

1.3. Aims and purpose

This systematic review and meta-analysis aimed to investigate whether the ability to accurately recognise basic emotions in the different domains across visual (i.e., human and nonhuman face) and auditory modalities (i.e., speech prosody and music) in autistic individuals differs from that in NT individuals (see Supplementary Table S1 for a detailed summary of the similarities and differences between prior reviews and the present study). We also examined the possible influence of study quality factors on these observed group differences. One specific focus was on the influence of IQ matching, given that group differences in emotion recognition could be confounded by differences in IQ rather than diagnosis per se (see Harms et al., 2010 for a detailed discussion). Another specific focus was on the influence of stimulus presentation time, given that tasks implementing restriction on stimulus presentation may prevent the use of compensatory strategies by autistic individuals and hence have higher sensitivity in detecting group differences. This postulation is supported by studies with no restriction failing to reveal group differences (e.g., Akechi et al., 2009; Grossman et al., 2000; Fink, de Rosnay, Wierda, Koot, & Begeer, 2014; Lacroix, Guidetti, Rogé, & Reilly, 2014), as opposed to studies with restriction tending to find group differences (e.g., Clark et al., 2008; Ciaramidaro et al., 2018 Otsuka et al., 2017; Griffiths et al., 2019). To identify factors contributing to the mixed results, the present work investigated a number of potential moderators, including age, full-scale IQ, verbal IQ, nonverbal IQ, stimulus domain, and task demand. In addition to recognition accuracy, response times (where available) were also examined separately in the main meta-analyses to address the discrepancies in the literature that show either slower (Greimel et al., 2014; Ketelaars, In'T Velt, Mol, Swaab, & Van Rijn 2016; Lindström et al., 2018; Sawyer et al., 2012) or comparable (Akechi et al., 2010; Fink et al., 2014; Grossman et al., 2000) speed of emotion recognition in ASD relative to NT.

2. Methods

This study was carried out in accordance with the recommended procedures for conducting systematic reviews and meta-analyses (i.e., the Preferred Reporting Items for Systematic reviews and Meta-Analyses [PRISMA] guidelines; Moher et al., 2009). The protocol for this systematic review was registered on PROSPERO at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018091703.

2.1. Search strategy

To identify all eligible studies, a comprehensive search was conducted using Embase, Medline, PubMed, Web of Science and Google Scholar, from the earliest record to August 2019. Our search strategy was designed to combine ASD and emotion recognition relating to the domains of interest using three blocks of search terms: (a) synonyms and terms used to describe ASD; (b) emotion recognition; and (c) human face, nonhuman face, speech prosody, and music (see Supplementary Table S2 for full details). The initial search was completed on 27th April 2018; additional studies were included until 1st August 2019 via an update search.

2.2. Study selection and eligibility criteria

The inclusion criteria of the review required studies that:

1. included an ASD group with mean IQ ≥ 70 and a sample size ≥ 2. The ASD diagnosis encompassed terms including ASD, autism, Asperger's syndrome, and pervasive developmental disorder-not otherwise specified (PDD-NOS), depending on the diagnostic tools/scales used in the studies. In the autism literature, it is often difficult to match ASD and NT groups on pre-existing variables such as age, language level, or IQ (Bang et al., 2020; Jarrold & Brock, 2004). Given the significant influence of IQ on emotion recognition (Lawrence et al., 2015) and to ensure that the group differences observed were not due to differences in intellectual ability between groups, we used mean IQ ≥ 70 as the inclusion criterion for the ASD group, as IQ < 70 would lie two or more standard deviations below the general population (Boat & Wu, 2015).
2. included an NT control group for comparison.
3. implemented an emotion recognition paradigm.
4. used stimuli that conveyed emotions through human faces, nonhuman faces, speech prosody, or music.
5. included an objective measure of unimodal emotion recognition accuracy or response time.
6. provided summary statistics (i.e., means and standard deviations) of each group or inferential statistics (i.e., t, F, or z statistics) of group difference for at least one of the six basic emotions (anger, fear, happiness, sadness, disgust, and surprise), or for an overall composite with all six emotions combined. In cases where data were not present in eligible studies (e.g., due to reporting styles), first and/or corresponding authors were contacted in an attempt to obtain sufficient information to calculate effect sizes. If this

information was not obtained upon two request attempts, studies were excluded from our review and analysis.

7. published in English and in peer-reviewed journals. This criterion was set because (a) professional translators are needed for review teams to include non-English-language studies (Neimann Rasmussen & Montgomery, 2018), which requires extra costs and resources, given that translation software such as Google Translate produces uneven accuracy across different languages (Aiken, 2019); (b) there is no evidence of a language bias regarding effect estimates or conclusions in meta-analyses relying exclusively on English-language studies (Dobrescu et al., 2021; Morrison et al., 2012); and (c) there is currently no "gold standard" for systematic search of grey literature (Adams et al., 2016; Paez, 2017). However, despite our comprehensive search in multiple databases, we cannot exclude the possibility of missing non-English studies and grey literature, which is a limitation of the present study.

All citations retrieved from the searches were imported into EndNote X8 (Clarivate Analytics, 2016). The resulting set was screened for eligibility against inclusion criteria using a three-stage process of reviewing the titles, followed by abstracts, and full-texts. Each stage involved two review authors (FYNL and CD/AV/JO) independently screening the retrieved items. To ensure rigour, a random 25% of the total sample of all titles, abstracts and full texts were doubly and independently screened by two review authors. Any disagreement over the eligibility of particular studies was resolved through discussion by all review authors.

2.3. Data extraction

The data of full texts were extracted independently by four review authors (FYNL, CD, JO, and CZ). A random 10% of studies were doubly extracted by FYNL and CD to ensure accuracy of the data extracted. The predefined data extraction form included items of study setting, study characteristics, intellectual level of participants, diagnostic instruments, group matching procedures, type of task, stimulus type and domain, emotions assessed, and the study results. With regards to studies that included stimuli with more than one intensity level, the most prototypical and well-validated stimuli presented at the highest intensity were extracted. One study included stimuli with two different presentation durations (2000ms vs. 50ms; Otsuka et al., 2017), and only the data with the 2000ms presentation time were extracted. This was to ensure greater homogeneity of the data extracted across studies, since most studies used a presentation time greater than 50ms. The accuracy and response time results were extracted based on summary statistics (i.e., means and standard deviations) of each group or inferential statistics (i.e., *t, F, or z* statistics) of group difference in terms of performance on each individual emotion assessed (i.e., anger, fear, happiness, sadness, disgust, and/or surprise) and/or performance across all six emotions (i.e., the composite). Studies with multiple tasks, domains, and emotions were extracted as individual datasets (e.g., 1. verbal, face, angry; 2. verbal, face, fearful; 3. verbal, speech, angry; 4. verbal, speech, fearful) for separate analyses, as outlined in Section 2.5.

2.4. Quality assessment

The Critical Appraisal Skills Programme (CASP, 2017) case control study appraisal checklist was used to assess the methodological quality and the overall risk of bias of all included studies. The task was performed by four review authors (FYNL, CD, JO, and CZ), with a random 10% of studies doubly assessed by FYNL and CD. Discrepancies were discussed and resolved among the team.

2.5. Analysis plan

The main meta-analyses were conducted on the fourteen sets of effect sizes for accuracy and response time (seven each) separately, in order to examine group differences in emotion recognition for anger, fear, happiness, sadness, disgust, surprise, and the composite. All analyses were performed in R (RStudio Team, 2018). The standardised mean difference (SMD) was computed as the overall pooled effect for each dataset using the *compute.es* package (Deeks, Altman, & Bradburn, 2001; Del Re, 2013). SMDs (given as Hedges' *g*) and their corresponding sampling variances for each dataset were calculated from summary statistics (Hedges, 1981). In cases where non-positive sampling variances (i.e., 0) were reported, datasets were omitted from the meta-analysis (Deeks, Higgins, & Altman, 2019; Vesterinen et al., 2014).

Random-effects models were run to analyse the SMDs from individual datasets and to compute the estimated pooled SMD together with its 95% confidence interval (CI) using the *metafor* package in R (Viechtbauer, 2010). The greater the magnitude of the pooled SMD, the larger the difference in effect between the ASD and NT groups. We interpreted an SMD < 0.40 as a small effect size, 0.40 – 0.70 as moderate, and > 0.70 as large (Schünemann et al., 2019). Statistical heterogeneity was quantified using the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003) and the 95% prediction interval (PI; or "plausible value interval") of the estimated pooled SMD (IntHout, Ioannidis, Rovers, & Goeman, 2016; Spineli & Pandis, 2020). Values of I^2 were classified as $< 50\%$ (low; possibly unimportant) and $\geq 50\%$ (high; considerable concern). Since I^2 is not an absolute indicator of heterogeneity (Borenstein, Higgins, Hedges, & Rothstein, 2017), we also included prediction intervals, which estimate heterogeneity by providing the expected ranges of true effects in future similar studies (IntHout et al., 2016; Spineli & Pandis, 2020).

As there were often multiple datasets per study, we referred to the count of studies as *k* and the count of datasets as *N*. For multiple datasets in each study, we employed the following prioritisation criteria in our main meta-analyses so that effect sizes would not be computed multiple times based on data from the same sample: (a) where the same group of participants were tested on multiple domains, human faces were prioritised, followed by speech prosody, nonhuman faces and music; (b) in cases where multiple tasks were reported, priority was given to verbal over nonverbal tasks, and labelling over discrimination, matching and detection tasks; and (c) where multiple NT groups were present, priority was given to the NT group matched with the ASD group on chronological age (and full-scale IQ, if available), followed by verbal mental age and performance mental age (and verbal IQ and nonverbal IQ, if available). These prioritisation criteria were chosen in order to reduce heterogeneity of the data analysed in our main meta-analyses, since the majority of the included studies matched groups on chronological age (and full-scale IQ, if available) and examined facial emotion recognition using verbal labelling tasks.

To establish the significance of plausible moderating factors in emotion recognition, we conducted a series of planned subgroup analyses and meta-regressions as guided by previous literature. Planned subgroup analyses for each emotion were conducted based on domain (human faces vs. nonhuman faces vs. speech prosody vs. music) and task demand (verbal vs. nonverbal). Following the computation of the pooled SMDs for each subgroup, effect sizes were compared to determine differences between subgroups using the Cochran's *Q* test (Borenstein & Higgins, 2013; Çoğaltay & Karadağ, 2015). To examine the moderating effects of age, full-scale IQ, verbal IQ, and nonverbal IQ on group differences for each emotion, univariate meta-regressions were conducted to assess whether there were significant associations between each of these moderators and performance accuracy (Thompson & Higgins, 2002) when sufficient data were available (≥10 studies; Deeks et al., 2019).

We performed sensitivity analyses to test the robustness of the main meta-analysis results on accuracy, by including only studies meeting specific quality standards, i.e., those that had undertaken full-scale, verbal, and nonverbal IQ matching and/or stimulus presentation time restriction. Specifically, sensitivity analyses were first performed to assess the impact of IQ matching following the removal of datasets that did not undertake full-scale, nonverbal, and verbal IQ matching (on verbal tasks only, given that verbal IQ has been proposed to be an important contributor to performance on tasks requiring explicit verbal knowledge in ASD as discussed earlier). Likewise, sensitivity analyses were conducted to examine the impact of stimulus presentation time restriction following the exclusion of datasets without restricting stimulus presentation times. Results obtained from the sensitivity analyses were subsequently compared against the results of the main meta-analyses in order to assess the sustainability of the results after focusing on datasets with more controlled study designs. Moreover, through informal comparisons between the results of the main meta-analyses and sensitivity analyses, we evaluated whether pooled SMDs and heterogeneity were strengthened or weakened following exclusion of datasets aforementioned. In addition, to test the impact of IQ matching on the robustness of the results from the moderator analyses on full-scale, verbal, and nonverbal IQ, further sensitivity analyses were performed following the removal of datasets that did not undertake full-scale/verbal/nonverbal IQ matching, respectively.

Publication bias (Easterbrook, Gopalan, Berlin, & Matthews, 1991) was assessed with funnel plots and the Egger's test of funnel plot asymmetry (Egger, Smith, Schneider, & Minder, 1997; Page, Higgins, & Sterne, 2019). Trim-and-fill analyses using the R_0 estimator were conducted to establish how each respective mean effect size would change if any identified bias were to be removed (Duval & Tweedie, 2000), in addition to test of the null hypothesis that the number of missing studies is zero (Duval, 2005).

3. Results

3.1. Study characteristics

Figure 1 shows the PRISMA flowchart of the study selection process. A total of 72 papers with 1868 ASD and 2232 NT participants were included (see Supplementary Table S3 for a summary of study/participant characteristics and Supplementary Figure S1 for geographic distribution of the papers). Table 1 provides a detailed summary of the included studies. One paper (Fridenson-Hayo et al., 2016) reported results from three separate international sites that all met our inclusion criteria and were therefore considered to be three independent samples in the meta-analysis. This resulted in a total of 74 studies and 332 unique datasets: 73 studies contributed 259 unique accuracy datasets and 27 studies contributed 73 unique response time datasets for the main meta-analyses.

[Insert Figure 1 about here]

[Insert Table 1 about here]

3.2. Quality assessment

The results of the critical appraisal process are summarised in Figure 2. The overall quality of the included studies was high, with total scores ranging from 28 to 35 out of the maximum score of 36 on the CASP case control study quality assessment (CASP, 2017). The strengths of the literature included the formulation of a clearly focused research question and the implementation of the appropriate methods to address the research aims. Selection bias was determined by the recruitment criteria of autistic participants and NT controls. The use of standard diagnostic instruments during the selection of autistic participants was of particular importance to reflect the representativeness of the clinical group. Among the 11 papers with moderate to high risk of bias, eight did not specify how the clinical group was diagnosed and three had very small samples. The lack of consistent matching procedures across studies might have compromised the generalisability of the findings. Thirty-nine papers did not match the ASD and NT groups on all three of the essential background measures, namely age, gender, and IQ, while two papers did not match groups on any of these measures. Measurement bias was observed in eight papers which could be characterised by unvalidated emotional stimuli, unstandardized experimental methods implemented across participants or conditions, and inappropriate measurement methods. Inappropriate reporting styles lowered the quality of a substantial proportion of the included studies. These related to the reporting of effect sizes which were missing in 31 papers as well as the reporting of imprecise *p*-values of inferential statistical tests in 23 papers. The reliability of results was also of concern for nine papers. Among those, five papers concluded their findings that were not justified by the inferential statistical tests conducted. Two papers did not conduct follow-up post-hoc tests to determine group differences within an interaction. One paper included contradictory information, e.g., groups described as matched by verbal IQ yet with a statistically significant group difference in verbal IQ. The remaining paper drew results based on a limited number of trials. The generalisability of the study findings to clinical populations was driven by the coverage of the range of participant demographics in terms of age and gender. While the sample in two papers covered both genders from different age groups, 61 papers covered both genders from only one age group (e.g., both female and male children) and six papers covered only one gender from the same age group (e.g., male adults only).

[Insert Figure 2 about here]

3.3. Main meta-analyses

3.3.1. Group differences in emotion recognition accuracy

Results from the main meta-analyses on accuracy were summarised in Supplementary Table S4. The pooled SMDs for group differences in recognition accuracy were significant across all six basic emotions: anger (N = 52, SMD = -0.42, 95% CI [-0.59, -0.24], 95% PI [-1.54, 0.70], $p < 0.001$, $I^2 = 81.40\%$), fear (N = 46, SMD = -0.47, 95% CI [-0.66, -0.28], 95% PI [-1.60, 0.66], $p < 0.001$, $I^2 = 82.07\%$), happiness (N = 52, SMD = -0.45, 95% CI [-0.61, -0.28], 95% PI [-1.51, 0.61], $p < 0.001$, $I^2 = 79.03\%$), sadness (N = 48, SMD = -0.47, 95% CI [-0.65, -0.28], 95% PI [-1.57, 0.64], $p < 0.001$, $I^2 = 80.87\%$), disgust (N = 20, SMD = -0.41, 95% CI [-0.58, -0.24], 95% PI [-0.89, 0.07], $p < 0.001$, $I^2 = 38.89\%$), and surprise (N = 19, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$). Significant group differences in recognition accuracy were also observed for the six-emotion composite score (N = 22, SMD = -0.77, 95% CI [-1.03, -0.50], 95% PI [-1.87, 0.33], $p < 0.001$, I^2 =

77.56%), as shown in Figure 3. Depending on the emotion type, the pooled SMDs represented small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), and large effects (i.e., composite), as can be seen in Figure 4. While heterogeneity was not observed for surprise, it was low for disgust but considerably high for anger, fear, happiness, sadness, and the composite. These results indicate that overall, the ASD group showed lower accuracy than the NT group in emotion recognition across all emotion types, and the size of these group differences varied by emotion type. Importantly, there was substantial heterogeneity across studies for most emotion types. Exploration for potential contributors to the observed high heterogeneity is reported in Section 3.4 via moderator analyses.

[Insert Figure 3 about here]

[Insert Figure 4 about here]

The influence of IQ matching on meta-analysis results

To examine the impact of IQ matching on the robustness of the results in the main meta-analyses on accuracy, sensitivity analyses were performed on datasets that had implemented full-scale IQ matching (on all tasks), nonverbal IQ matching (on all tasks), and verbal IQ matching (on verbal tasks only) for each emotion category (see Supplementary Table S5 for full results).

In the sensitivity analyses including datasets that employed full-scale IQ matching, the pooled SMDs remained significant for all emotions: anger (N = 30, SMD = -0.28, 95% CI [-0.44, -0.12], 95% PI [-0.93, 0.37], $p < 0.001$, I^2 = 55.43%), fear (N = 26, SMD = -0.25, 95% CI [-0.42, 0.08], 95% PI [-0.91, 0.40], $p = 0.003$, I^2 = 55.73%), happiness (N = 27, SMD = -0.38, 95% CI [-0.57, -0.19], 95% PI [-1.18, 0.41], $p < 0.001$, I^2 = 64.04%), sadness (N = 23, SMD = -0.27, 95% CI [-0.46, -0.07], 95% PI [-1.00, 0.47], $p = 0.007$, I^2 = 61.87%), disgust (N = 13, SMD = -0.31, 95% CI [-0.52, 0.11], 95% PI [-0.80, 0.17], $p = 0.002$, I^2 = 38.42%), surprise (N = 13, SMD = -0.28, 95% CI [-0.44, -0.13], 95% PI [-0.44, -0.13], $p < 0.001$, I^2 =

0.00%), and the composite (N = 9, SMD = -0.79, 95% CI [-1.31, -0.26], 95% PI [-2.34, 0.77], p = 0.003, I^2 = 88.24%). In comparison with the main meta-analysis results, the size of the pooled SMDs decreased from moderate to small for anger, fear, happiness, sadness, and disgust after removing datasets without implementing full-scale IQ matching. The small effect for surprise and the large effect for the composite remained unchanged. Heterogeneity, although reduced, remained considerably high for anger, fear, happiness, and sadness, while the low heterogeneity for disgust and nil heterogeneity for surprise remained unchanged. Heterogeneity for the composite remained considerably high. These results indicate that, in comparison to the main meta-analysis results based on the full datasets, the sizes of the group differences were weakened when only including studies that matched groups on full-scale IQ. In addition, the reduced, yet high, heterogeneity observed here suggests that full-scale IQ matching may explain some of the variability across studies, but only to a certain extent.

In the sensitivity analyses including datasets that employed nonverbal IQ matching, the pooled SMDs remained significant for all emotions: anger (N = 13, SMD = -0.35, 95% CI [-0.53, -0.17], 95% PI [-0.67, -0.03], p < 0.001, I^2 = 17.73%), fear (N = 17, SMD = -0.41, 95% CI [-0.62, -0.21], 95% PI [-1.02, 0.19], p < 0.001, I^2 = 48.22%), happiness (N = 15, SMD = -0.47, 95% CI [-0.64, -0.30], 95% PI [-0.81, -0.13], p < 0.001, I^2 = 20.21%), sadness (N = 11, SMD = -0.44, 95% CI [-0.70, -0.18], 95% PI [-1.11, 0.24], p = 0.001, I^2 = 55.21%), disgust (N = 9, SMD = -0.44, 95% CI [-0.73, -0.14], 95% PI [-1.17, 0.29], p = 0.004, I^2 = 59.22%), surprise (N = 7, SMD = -0.28, 95% CI [-0.47, -0.08], 95% PI [-0.47, -0.08], p = 0.005, I^2 = 0.00%), and the composite (N = 12, SMD = -0.86, 95% CI [-1.23, -0.49], 95% PI [-2.05, 0.33], p < 0.001, I^2 = 80.70%). In comparison with the main meta-analysis results, the size of the pooled SMDs decreased from moderate to small for anger upon removing datasets without implementing nonverbal IQ matching. The small effect for surprise, moderate effects for fear, happiness, sadness, and disgust, and the large effect for the composite remained unchanged. Heterogeneity substantially reduced from high to low for anger, fear, and happiness but increased from low to high for disgust. The nil heterogeneity for surprise and the high heterogeneity for sadness and the composite remained unchanged. These results indicate that nonverbal IQ matching may explain some of the variability across studies, but only for certain emotions.

Given that verbal ability is required to label emotions in verbal tasks, we conducted sensitivity analyses examining the influence of verbal IQ matching on datasets of verbal tasks. The analyses showed that the pooled SMDs from VIQ-matched datasets involving verbal tasks remained significant for all emotions: anger ($N = 17$, SMD = -0.35, 95% CI [-0.51, -0.18], 95% PI [-0.75, 0.05], $p < 0.001$, $I^2 = 28.58\%$), fear ($N = 18$, SMD = -0.39, 95% CI [-0.58, -0.20], 95% PI [-0.93, 0.15], $p < 0.001$, $I^2 = 41.49\%$), happiness ($N = 18$, SMD = -0.42, 95% CI [-0.57, -0.28], 95% PI [-0.65, -0.20], $p < 0.001$, $I^2 = 7.85\%$), sadness ($N = 13$, SMD = -0.35, 95% CI [-0.55, -0.15], 95% PI [-0.82, 0.12], $p = 0.001$, $I^2 = 35.71\%$), disgust ($N = 8$, SMD = -0.39, 95% CI [-0.69, -0.08], 95% PI [-1.11, 0.34], $p = 0.014$, $I^2 = 59.30\%$), surprise ($N = 9$, SMD = -0.27, 95% CI [-0.45, -0.091], 95% PI [-0.45, 0.09], $p = 0.004$, $I^2 = 0.00\%$), and the composite ($N = 11$, SMD = -0.95, 95% CI [-1.39, -0.51], 95% PI [-2.33, 0.43], $p < 0.001$, $I^2 = 83.74\%$). In comparison with the main meta-analysis results, the sizes of the pooled SMDs for anger, fear, sadness, and disgust decreased from moderate to small upon removing datasets without implementing verbal IQ matching. The effect for the composite was particularly strengthened and remained large, while the small effect for surprise and moderate effect for happiness remained unchanged. Heterogeneity substantially reduced from high to low for anger, fear, happiness, and sadness but increased from low to high for disgust. The nil heterogeneity for surprise and high heterogeneity for the composite remained unchanged. These results indicate that verbal IQ matching for verbal tasks may explain the variability across studies for most emotions.

Overall, these three sets of sensitivity analyses confirmed the robustness of the results of the main meta-analyses on accuracy: the pooled SMDs for group differences in recognition accuracy remained significant across all six basic emotions and the composite after the removal of datasets without IQ matching. The magnitude of group differences in emotion recognition was, nevertheless, weakened for most emotions when groups were matched on full-scale IQ (anger, fear, happiness, sadness, and disgust) and verbal IQ (anger, fear, sadness, and disgust), but remained relatively unchanged with nonverbal IQ matching (except for anger). Notably, group differences were particularly strengthened for the composite upon verbal IQ matching. Moreover, whether groups were matched on IQ appears to be important sources of variability across studies for anger, fear, happiness, and sadness, with verbal IQ matching showing a greater contribution above and beyond that by nonverbal and full-scale IQ matching.

The influence of stimulus presentation time restriction on meta-analysis results

Given the potential influence of stimulus presentation time restriction on emotion recognition performance, we conducted sensitivity analyses to investigate if the main meta-analysis results remained after excluding datasets with unrestricted presentation times (see Supplementary Table S6 for full results). Using only datasets with restricted presentation times, sensitivity analyses revealed that the pooled SMDs for group differences remained significant for all emotions: anger (N = 21, SMD = -0.41, 95% CI [-0.66, -0.16], 95% PI [-1.38, 0.57], p = 0.001, I^2 = 71.45%), fear (N = 23, SMD = -0.45, 95% CI [-0.72, -0.18], 95% PI [-1.60, 0.69], p = 0.001, I^2 = 76.74%), happiness (N = 25, SMD = -0.36, 95% CI [-0.51, -0.20], 95% PI [-0.87, 0.16], p < 0.0001, I^2 = 39.80%), sadness (N = 20, SMD = -0.41, 95% CI [-0.68, -0.14], 95% PI [-1.45, 0.64], p = 0.003, I^2 = 72.79%), disgust (N = 11, SMD = -0.45, 95% CI [-0.72, -0.18], 95% PI [-1.12, 0.23], p = 0.001, I^2 = 50.98%), surprise (N = 12, SMD = -0.26, 95% CI [-0.43, -0.08], 95% PI [-0.43, -0.08], p = 0.004, I^2 = 0.00%), and the composite (N = 11, SMD = -0.83, 95% CI [-1.32, -0.33], 95% PI [-2.40, 0.75], p = 0.001, I^2 = 86.63%). Compared to the results from the main meta-analyses, the sizes of the pooled SMDs reduced from moderate to small for happiness, while the moderate effects for anger, fear, sadness, and disgust and the large effect for the composite remained unchanged. Heterogeneity was reduced from high to low for happiness but increased from low to high for disgust. The low heterogeneity for surprise and the high heterogeneity for anger, fear, sadness, and the composite remained unchanged. These results suggest that stimulus presentation time restriction may not be an important contributor to the mixed findings across studies.

3.3.2. Group differences in emotion recognition response time

The pooled SMDs for group differences in recognition response time were significant and represented moderate effects for anger ($N = 18$, SMD = 0.45, 95% CI [0.02, 0.88], 95% PI [-1.34, 2.24], $p = 0.041$, $I^2 = 93.08\%$), fear ($N = 17$, SMD = 0.57, 95% CI [0.07, 1.07], 95% PI [-1.48, 2.62], $p = 0.026$, $I^2 = 94.46\%$), sadness ($N = 12$, SMD = 0.70, 95% CI [0.08, 1.32], 95% PI [-1.37, 2.88], $p = 0.027$, $I^2 = 95.57\%$), and the composite ($N = 3$, SMD = 0.45, 95% CI [0.09, 0.81], 95% PI [0.07, 0.85], $p = 0.014$, $I^2 = 0.00\%$). Heterogeneity was not observed for the composite but was substantially high for anger, fear, and sadness. The group differences in recognition response time, however, did not reach significance for happiness, disgust, and surprise (see Supplementary Table S7 for full results). These results indicate that the ASD group was generally slower than the NT group at recognising anger, fear, sadness, and the composite emotions, with substantial amount of heterogeneity across studies. The two groups, nevertheless, showed comparable response time when recognising emotions of happiness, disgust, and surprise.

3.4. Moderator analyses

The following moderator analyses were only performed on recognition accuracy, but not on response time (due to limited data available).

3.4.1. Age

Meta-regressions revealed that age (i.e., mean age of the ASD groups) was a significant predictor of the magnitude of the pooled SMD for the composite ($N = 22$, Q_M (1) = 16.22, $p <$ 0.001), accounting for 57.43% of the variability in the difference in SMDs among studies. As seen in Figure 5, age was negatively associated with the SMDs for the composite ($\beta = -0.05$, SE = 0.01): the older the autistic participants, the greater difference in recognition accuracy of the composite emotions between ASD and NT groups. Age was, however, not a significant predictor for the recognition of other emotions (see Supplementary Table S8 for full results). These results suggest that the group difference increases with age for the composite measure, but not for the individual emotions.

[Insert Figure 5 about here]

3.4.2. IQ

Separate meta-regressions across all the available datasets revealed that full-scale IQ, verbal IQ, and nonverbal IQ (i.e., mean standard scores of the ASD groups) were not significant moderators of the magnitude of the pooled SMDs for any of the seven emotion types (see Supplementary Table S9 for full results). These results indicate that group differences in emotion recognition accuracy were not moderated by full-scale IQ, verbal IQ, or nonverbal IQ.

The influence of IQ matching on meta-regression results

To examine the impact of IQ matching on the robustness of the meta-regression results regarding IQ as shown above, sensitivity analyses were carried out on datasets that had implemented full-scale, verbal, or nonverbal IQ matching for each emotion category (see Supplementary Table S10 for full results).

In the sensitivity analyses including datasets that employed full-scale IQ matching and nonverbal IQ matching, full-scale IQ and nonverbal IQ remained as nonsignificant predictors of the magnitude of the pooled SMDs for all emotions.

In the sensitivity analyses including datasets that employed verbal IQ matching, verbal IQ was a significant predictor of the magnitude of the pooled SMDs for happiness ($N = 15$, Q_M $(1) = 5.68$, $p = 0.017$), but not for other emotions. The negative association between verbal IQ and the SMDs for happiness (β = -0.03, SE = 0.01) indicated that the higher the verbal IQ of the autistic participants, the lower the SMDs (i.e., poorer recognition accuracy of happiness in the ASD group relative to the NT group) (see Supplementary Figure S3 for a scatter plot).

3.4.3. Stimulus domain

Subgroup analyses were performed on human face, nonhuman face, speech prosody, and music domains separately for each emotion type (see Supplementary Table S11 for full results).

For studies on emotion recognition in human faces, subgroup analyses revealed that the ASD groups were significantly worse than NT controls across all emotions: anger ($N = 43$, SMD = -0.47, 95% CI [-0.68, -0.27], 95% PI [1.68, 0.73], $p < 0.001$, $I^2 = 84.26\%$), fear ($N = 41$, SMD = -0.47, 95% CI [-0.68, -0.26], 95% PI [-1.69, 0.74], $p < 0.001$, $I^2 = 84.26\%$), happiness ($N = 41$, SMD = -0.47, 95% CI [-0.66, -0.27], 95% PI [-1.57, 0.64], $p < 0.001$, $I^2 = 81.36\%$), sadness ($N = 37$, SMD = -0.48, 95% CI [-0.71, -0.26], 95% PI [-1.72, 0.75], $p < 0.001$, $I^2 = 85.19\%$), disgust ($N = 19$, SMD = -0.40, 95% CI [-0.57, -0.23], 95% PI [-0.89, 0.10], $p < 0.001$, $I^2 = 40.73\%$), surprise ($N = 19$, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$), and the composite ($N = 20$, SMD = -0.72, 95% CI [-0.95, -0.49], 95% PI [-1.57, 0.13], $p < 0.001$, $I^2 = 67.76\%$). The pooled SMDs represented small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), to large effects (i.e., composite). The results for disgust and surprise were associated with low heterogeneity, whereas the results for the other emotions were associated with high heterogeneity. These results suggest that, for studies investigating human faces alone, group differences were statistically significant across all emotions: (a) the ASD group showed lower recognition accuracy compared to the NT group across all seven emotion types; (b) the size of these group differences varied by emotion type; and (c) substantial heterogeneity across studies remained for most emotions.

In contrast to the above results on human faces, no significant group differences were found in the nonhuman face subgroup analyses for emotions with sufficient datasets, including anger ($N = 3$, SMD = -0.11, 95% CI [-0.49, 0.28], 95% PI [-0.53, 0.31], $p = 0.591$, $I^2 = 6.09\%$) and happiness ($N = 3$, SMD = 0.34, 95% CI [-0.74, 1.43], 95% PI [-1.71, 2.40], $p = 0.535$, $I^2 = 86.85\%$). These results indicate that group differences were not evident in studies investigating nonhuman faces alone, where the ASD and NT groups showed comparable accuracy in the recognition of anger and happiness, specifically.

Within the auditory modality, mixed results were obtained from the speech prosody subgroup analyses. The ASD groups were significantly worse than NT controls at recognition of anger ($N = 12$, SMD = -0.34, 95% CI [-0.64, -0.03], 95% PI [-1.26, 0.59], $p = 0.030$, $I^2 = 74.48\%$), happiness ($N = 12$, SMD = -0.52, 95% CI [-0.76, -0.28], 95% PI [-1.16, 0.13], $p < 0.001$, $I^2 = 57.90\%$), and disgust ($N = 2$, SMD = -0.93, 95% CI [-1.40, -0.46], 95% PI [-1.40, -0.46], $p < 0.001$, $I^2 = 0.00\%$) in speech prosody. These significant pooled SMDs represented small (i.e., anger), moderate (i.e., happiness), and large effects (i.e., disgust). While no heterogeneity was observed for disgust, results for anger and happiness were associated with high heterogeneity. The pooled SMDs for fear, sadness, and the composite, however, did not reach significance, while no datasets on surprise were available precluding analysis. These results indicate that, for studies investigating speech prosody alone, lower accuracy by the ASD group compared to the NT group was consistently found in the recognition of anger, happiness, and disgust, with the size of these group differences and amount of heterogeneity varying by emotion type. By contrast, the ASD and NT groups showed comparable accuracy in the recognition of fear, sadness, and the composite.

Similar to the speech domain, mixed results were obtained from the music subgroup analyses. The ASD groups were significantly worse than NT controls at recognising fear (N = 2, SMD = -0.50, 95% CI [-0.92, -0.08], 95% PI [-0.92, -0.08], p = 0.021, I^2 = 0.00%) and sadness (N = 2, SMD = -0.63, 95% CI [-1.06, -0.21], 95% PI [-1.06, -0.21], p = 0.004, I^2 = 0.00%) in music. The pooled SMD for happiness was, however, not significant, while there were no available datasets for anger, disgust, surprise, nor the composite. These results indicate that, for studies investigating music alone, lower accuracy by the ASD group compared to the NT group was consistently found in the recognition of fear and sadness, both with a moderate effect size and no heterogeneity across studies. By contrast, the ASD and NT groups showed comparable accuracy in the recognition of happiness in music.

The test for subgroup differences revealed that the pooled SMDs did not differ significantly across human faces, speech prosody, nonhuman faces, and music for any of the emotions (see Supplementary Table S11 for full results). It should, however, be noted that subgroup comparison was not available for surprise due to a lack of datasets on domains other than human faces. Subgroup comparison incorporating all four domains was only feasible for happiness, with significant group differences observed for human faces and speech prosody, but not for nonhuman faces or music (see Supplementary Figure S2 for forest plots).

3.4.4. Task demand

Subgroup analyses were done on verbal and nonverbal tasks separately for each emotion type (see Supplementary Table S12 for full results).

For verbal tasks, the pooled SMDs for group differences in recognition accuracy were significant across all emotions: anger (N = 47, SMD = -0.44, 95% CI [-0.62, -0.27], 95% PI [-1.49, 0.61], $p < 0.001$, I^2 = 79.75%), fear (N = 45, SMD = -0.50, 95% CI [-0.68, -0.32], 95% PI [-1.54, 0.55], $p < 0.001$, I^2 = 79.56%), happiness (N = 47, SMD = -0.49, 95% CI [-0.66, -0.32], 95% PI [-1.50, 0.52], $p < 0.001$, I^2 = 77.85%), sadness (N = 45, SMD = -0.48, 95% CI

[-0.66, -0.31], 95% PI [-1.52, 0.55], $p < 0.001$, $I^2 = 78.88\%$), disgust ($N = 20$, SMD = -0.41, 95% CI [-0.58, -0.24], 95% PI [-0.89, 0.07], $p < 0.001$, $I^2 = 38.89\%$), surprise ($N = 19$, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$), and the composite ($N = 22$, SMD = -0.77, 95% CI [-1.03, -0.50], 95% PI [-1.87, 0.33], $p < 0.001$, $I^2 = 77.56\%$). These results were associated with small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), and large effects (i.e., composite). While heterogeneity was not observed for surprise, it was low for disgust and high for anger, fear, happiness, sadness, and the composite. Thus, focusing on studies employing a verbal task, subgroup analyses suggested significant group differences across all emotion types: (a) the ASD group showed lower accuracy than the NT group; (b) the sizes of these group differences varied by emotion type; and (c) substantial heterogeneity remained for most emotions.

For nonverbal tasks, the pooled SMDs for group differences in recognition accuracy were only significant for disgust, with a moderate effect and no heterogeneity ($N = 2$, SMD = -0.53, 95% CI [-0.83, -0.23], 95% PI [-0.83, -0.23], $p = 0.001$, $I^2 = 0.00\%$). No significant group differences were found for anger, fear, happiness, or sadness, while no datasets were available for surprise or the composite precluding analysis. These results indicate that, for nonverbal tasks, the ASD group performed worse than the NT group only for recognition of disgust, but not for recognition of anger, fear, happiness, or sadness.

The tests for subgroup differences revealed that the pooled SMDs differed between verbal and nonverbal tasks only for fear ($Q(1) = 25.38$, $p < 0.001$), but not for anger, happiness, or sadness. Due to the limited datasets available for nonverbal tasks relating to surprise and the composite and prioritisation of multiple datasets from the same sample for disgust, subgroup comparisons were precluded for these emotions.

3.5. Publication bias

The Egger's test identified significant potential bias in the studies contributing to the pooled SMDs for fear (z = -4.10, $p < 0.0001$) and the composite (z = -2.60, $p = 0.009$), but not for anger (z = -1.26, $p = 0.207$), happiness (z = -0.47, $p = 0.635$), sadness (z = -1.67, $p = 0.095$), disgust (z = -0.75, $p = 0.453$), or surprise (z = -0.47, $p = 0.637$). For fear, an estimated number of one study was missing on the right side of the funnel plot. A trim-and-fill procedure corrected the observed pooled SMD in the meta-analysis to -0.44 (95% CI [-0.65, -0.22], 95% PI [-1.77, 0.90], $p < 0.001$), which remained significant albeit with non-significant results in the test of null hypothesis that the number of missing studies is zero ($p = 0.250$). For the composite, the test of null hypothesis that the number of missing studies is zero was significant ($p = 0.016$). An estimated number of five studies were missing on the right side of the funnel plot (Figure 6). The trim-and-fill imputed mean effect for the composite remained significant with the observed pooled SMD corrected to -0.53 (95% CI [-0.84, -0.22], 95% PI [-2.02, 0.96], $p = 0.001$).

[Insert Figure 6 about here]

4. Discussion

The current systematic review and meta-analysis evaluated the recognition of six basic emotions and their composite in ASD relative to NT, across domains of human and nonhuman faces, speech prosody, and music, while identifying a number of potential moderating factors (age, IQ, domain, and task demand) that might have contributed to the mixed findings in the literature.

Combining non-overlapping datasets across the four domains, our main meta-analyses suggested emotion recognition impairments across all emotions in the ASD group, who also showed longer response times than NT controls for anger, fear, sadness, and the composite. Sensitivity analyses confirmed the robustness of the observed impairments in emotion recognition accuracy, as significant group differences remained for all emotions after removing datasets without implementing IQ matching or restricting stimulus presentation time. Nevertheless, the magnitude of these group differences was weakened for a subset of emotions across datasets with full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust), but less so for nonverbal IQ matching (i.e., anger). This indicates that while group differences in emotion recognition accuracy are not due to an absence of IQ matching, the magnitude of these differences could have been inflated due to the lack of IQ matching.

Moderator analyses indicated that age predicted the magnitude of group differences for the composite (but not for the individual emotions): the older the participants, the more pronounced the group differences. Domain was another significant moderator, as autistic individuals showed impaired recognition accuracy compared to controls across all emotions with human faces, but only for particular emotions with speech prosody (i.e., anger, happiness, and disgust) and music (i.e., fear and sadness), while no impairment was observed with nonhuman faces. Task demand also modulated group differences, with verbal tasks revealing group differences across all emotions and nonverbal tasks suggesting impairment of the ASD group for the recognition of disgust only. Finally, moderator analyses suggested no significant moderating effect of full-scale IQ, verbal IQ, and nonverbal IQ on group differences for any of the emotions. The nil effect of these IQ measures remained for datasets that had undertaken IQ matching, with one exception for verbal IQ, which significantly predicted the magnitude of group differences for happiness. Specifically, the higher the verbal IQ of the autistic group, the more prominent the group differences when recognising happiness. These findings will be further discussed in the subsections below.

4.1. Age-related factors

The current finding of more pronounced deficits in adults for the composite has been described in Harms et al. (2010). Indeed, cross-sectional literature shows a lack of improvement in emotion recognition skills in ASD beyond late childhood, whereas maturation of skills continues through adulthood in NT (Rump et al., 2009; Uono, Sato, & Toichi, 2011). In the present work, the age effect was only seen in the composite measure, which reflects greater task complexity as sophisticated categorical skills are required to distinguish all six basic emotions. Our focus on the composite score also explains the discrepancy in the age effect between the current and prior reviews. In prior reviews, the overall measure comprised any numbers of combinations of emotions examined in the individual studies (Lozier et al., 2014; Uljarevic & Hamilton, 2013), which may reflect a reduced task complexity as fewer emotions were involved. As a result, no effect of age was observed on the overall measures in prior reviews (Lozier et al., 2014; Uljarevic & Hamilton, 2013). Nevertheless, effects of age were found for individual emotions such as fear, sadness, and disgust in Lozier et al. (2014), where the data came from studies that examined all six emotions. Thus, it is plausible that studies involving more emotions are more sensitive in detecting group differences among adult participants.

In addition, research on the development of emotion recognition skills suggests that children generally achieve adult-level performance with prototypical expressions around 10 – 11 years of age across domains (Bruce et al., 2000; Chronaki, Hadwin, Garner, Maurage, & Sonuga-Barke, 2015; Mondloch, Geldart, Maurer, & Le Grand, 2003; Tonks, Williams, Frampton, Yates, & Slater, 2007; Van Lancker et al., 1989; Vidas, Dingle, & Nelson, 2018). However, the recognition of more subtle expressions requires considerable improvement in proficiency beyond this age (Chronaki et al., 2015; Rump et al., 2009). Given autistic children's preference for non-social over social stimuli (Gale et al., 2019) and musical over speech stimuli

(Blackstock, 1978), as well as their lack of social motivation (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012), it might be the case that they are not exposed to enough opportunities to develop and harness sophisticated emotional understanding like NT children do throughout adolescence and adulthood. Future longitudinal studies are needed to track the emotional development in autistic individuals, so that effective interventions can be employed to increase emotional skills in ASD (Vogan et al., 2018).

4.2. IQ-related factors

To investigate the influence of the study quality issue of IQ matching between groups on the general emotion recognition impairments in ASD, sensitivity analyses were conducted on datasets with full-scale/verbal/nonverbal IQ matching. The results confirmed the robustness of the main meta-analysis results on the whole datasets (with or without IQ matching), as the group differences for all emotions remained significant on the restricted datasets (with IQ matching). However, weakened magnitudes of group differences were observed for the majority of emotions among datasets with full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust), but less so for datasets with nonverbal IQ matching (i.e., anger). This indicates that matching groups on IQ could lower the magnitude of the observed group differences, as well as reduce heterogeneity across studies. Thus, it is important for future studies to incorporate IQ matching procedures in the design, in order to provide more precise estimates of group effects on emotion recognition performance that are not influenced by differences in IQ between the ASD and NT groups.

In the moderator analysis regarding IQ, our meta-regressions on all available datasets corroborated previous meta-analytic works in failing to detect significant effects of full-scale IQ and verbal IQ on group differences for any emotions (Lozier et al., 2014; Uljarevic & Hamilton, 2013). Additionally, our current results also suggested no significant effects of nonverbal IQ on group differences across studies. Focusing on datasets that matched groups on full-scale/verbal/nonverbal IQ, sensitivity analyses still suggested non-significant effects of IQ on group differences across all emotions with only one exception – the higher verbal IQ of the autistic participants, the larger group differences in the recognition of happiness. These findings indicate that the magnitude of group differences in emotion recognition is largely unrelated to the intellectual level of autistic individuals. It should, however, be noted that this lack of IQ effect does not imply that there is no relationship between IQ and emotion recognition in *absolute* terms; in other words, it does not mean that autistic individuals with lower IQ perform similarly to autistic individuals with higher IQ. Rather, these results indicate that the difference in performance between autistic individuals and their NT counterparts is largely unaffected by IQ differences in *relative* terms.

In fact, the effects of IQ on emotion recognition have predominantly been studied in *absolute* terms but have rarely been studied in *relative* terms in the autism literature. Among the limited studies investigating the *relative* effects of IQ on emotion recognition, Rommelse et al. (2015) found that group differences in performance on social cognition tasks (encompassing face recognition, facial and prosodic emotion recognition) were larger for autistic individuals with higher full-scale IQ in comparison to autistic individuals with lower full-scale IQ. The present finding of a lack of IQ effects across studies thus stands in contrast to the findings of Rommelse et al. (2015) – except for the observation of more prominent impairments for the recognition of happiness in autistic individuals with higher verbal IQ when the groups are matched on this measure. It is particularly curious as to why such effects were only observed for happiness but not for other emotions across studies. Notably, in Rommelse et al. (2015), the ASD and NT groups were carefully matched on full-scale IQ, verbal IQ, nonverbal IQ, as well as verbal-nonverbal IQ discrepancy. The lack of IQ effects in previous reviews (Lozier et al., 2014; Uljarevic & Hamilton, 2013) and the present study may be explained by the potential confounds of verbal-nonverbal IQ discrepancy between ASD and NT groups, since IQ matching did not (mostly) influence the moderating role of IQ in the present findings. Given the substantial variability in cognitive profiles in ASD (Nowell, Schanding, Kanne, & Goin-Kochel, 2015; Tager-Flusberg & Joseph, 2003), it is plausible that different cognitive profiles influence performance in different ways such that individuals with discrepantly higher verbal IQ may make use of verbal abilities to succeed on labelling tasks, while performance may be hindered in individuals with discrepantly lower verbal IQ. For example, within the same pair of autistic and NT participants, despite having a similar full-scale IQ score, if the autistic participant had a verbal > nonverbal IQ profile and the NT participant had a nonverbal > verbal IQ profile, the autistic participant may be able to use their verbal ability to compensate for their emotion recognition difficulties and perform comparably to their NT counterpart. By contrast, if the autistic participant had a nonverbal > verbal IQ profile while the NT participant had a verbal > nonverbal IQ profile, the autistic participant may not be as readily able to make use of their verbal ability as a compensatory strategy and thus perform worse than their NT counterpart. The separate effects of full-scale IQ, verbal IQ, and nonverbal IQ on group differences may not fully account for the potential effects of verbal-nonverbal IQ discrepancy. Given that research investigating the effects of IQ on group differences in *relative* terms is limited, more research is warranted. In addition, future research should consider the intertwining relationship between verbal-nonverbal IQ discrepancy and IQ characteristics on emotion recognition in autistic individuals *relative* to NT individuals. This area of research has important implications for clinical practice. If autistic individuals with higher IQ indeed had more severe emotion impairments in *relative* terms, but their emotion recognition capacities were overestimated by their social environment because of their high cognitive abilities, this can contribute to the development of behavioural problems (Howlin, 1998).

4.3. Domain- and emotion-related factors

The present work revealed general emotion recognition deficits for human faces in ASD, similar to the findings of previous meta-analyses (Lozier et al., 2014; Uljarevic & Hamilton, 2013). However, we observed no consistent group differences for nonhuman faces. The different results for human faces versus nonhuman faces may be related to the specific perceptual processing strategies used to process these stimuli in autistic versus NT individuals. Configuration information has been found to play a particularly prominent role in typical emotion recognition for both human faces (Bombari et al., 2013; Calder, Keane, Young, & Dean, 2000; Derntl, Seidel, Kainz, & Carbon, 2009; McKelvie, 1995; Prkachin, 2003; Rosset et al., 2008) and nonhuman faces (Rosset et al., 2008). In particular, the visual scan paths for emotional faces are strategic and controlled for NT individuals, who generally trace a triangle subtending the eyes, nose, and mouth (Pelphrey et al., 2002). Conversely, autistic individuals use atypical featural processing for human faces (Deruelle, Rondan, Gepner, & Tardif, 2004; Hernandez et al., 2009), and show disorganised visual scan paths (Pelphrey et al., 2002) and avoidance of the eyes (Frazier et al., 2017; Tanaka & Sung, 2016). However, for nonhuman faces, autistic individuals have shown both the use of typical configural processing with human cartoon faces (Rosset et al., 2008) and atypical featural processing with schematic faces (Isomura et al., 2014b). Interestingly, regardless of the processing strategies, autistic individuals performed comparably to their NT counterparts on recognition (Rosset et al., 2008) and detection (Isomura et al., 2014b) of emotions in nonhuman faces.

Taken together, these findings suggest that impairments in emotion processing of human faces in ASD are likely due to atypical featural processing strategies employed by autistic individuals. However, the relationship between processing strategy and emotion processing of nonhuman faces is less clear. Critically, two questions are yet to be scrutinised in future research. First, given the disparity in the use of different processing strategies for nonhuman faces reported (Rosset et al., 2008; Isomura et al., 2014b), how emotional expressions are decoded from nonhuman facial stimuli and whether they differ as a function of different types of nonhuman faces in autistic individuals need to be further examined, particularly through eye tracking techniques for more precise comparisons with that for human faces. Secondly, since intact emotion processing of nonhuman faces was observed regardless of the processing strategy employed by autistic individuals (Rosset et al., 2008; Isomura et al., 2014b), it raises the question as to whether the advantage of configural processing is less critical for nonhuman faces (which do not share the same special status in perception as human faces; e.g., Akdeniz, 2020; Rosset et al., 2010), and thus the use of either processing strategy may result in comparable performance. All in all, the literature on emotion processing of nonhuman faces in ASD is yet to be expanded, while further investigation should be directed to uncovering the processing strategies across the two visual domains and how they relate to emotion processing specifically.

Compared to the visual modality, less is known about how auditory emotions are processed in ASD. The current findings outline the preserved skills to decode some but not all basic emotions from prosody and music in ASD. Acoustic cues, such as pitch, are used similarly to infer emotions from prosody and music (Juslin & Laukka, 2003). Vocal emotion recognition has been associated with non-vocal pitch processing (Globerson, Amir, Kishon-Rabin, & Golan, 2015), but not with vocal pitch processing in ASD (Schelinski & von Kriegstein, 2019). Therefore, it is yet to be elucidated how pitch is integrated in informing prosodic emotion recognition in ASD. In addition, autistic individuals tended to rely on contextual cues (Le Sourn-Bissaoui et al., 2013) or verbal content (Lindner & Rosén, 2006; Stewart, McAdam, Ota, Peppé, & Cleland, 2013) to infer the speaker's emotions rather than using the emotional prosody. Given their intact ability to recognise emotions solely cued by prosody (Brennand et al., 2011; Le Sourn-Bissaoui et al., 2013), it is plausible that autistic individuals are able to extract basic sensory features of emotional prosody but fail to orient consistently to such vocal pitch cues. The way in which musical emotions are processed and whether such processing strategies are shared across these auditory domains are even less understood, and therefore more research is warranted in this area (Molnar-Szakacs & Heaton, 2012).

Examining data from four different domains, we reached different conclusions regarding whether emotion recognition deficits (if any) in ASD are specific to certain emotions. In the domain of human faces, deficits were observed across all emotions. For nonhuman faces, however, no deficits were observed for any emotions. Whereas recognition of anger, happiness, and disgust was impaired for speech prosody, fear and sadness were the most difficult for autistic individuals to recognise when it came to music. Thus, in the visual modality, autistic individuals seem to have a general ability (in cases of nonhuman faces) or disability (in cases of human faces) to recognise different emotions. In the auditory modality (in cases of speech and music), however, this (dis)ability is specific to certain emotions. These findings have implications for the debates over domain- and emotion-specificity vs. generality of emotion processing in ASD and NT (Baron-Cohen et al., 2000; Kanwisher, 2000; Nuske et al., 2013; Weigelt et al., 2013). Given the limited data available in domains other than human faces in the current literature, further comparative studies examining emotion recognition across different domains in the same sample of participants are needed to inform these debates.

4.4. Experimental factors

Contrary to Uljarevic and Hamilton's (2013) findings of a lack of differences between labelling and matching paradigms, our results revealed consistent deficits for verbal tasks but not for nonverbal tasks when detection and discrimination paradigms were also included. In addition, the moderating effect of task demand was particularly evident for the recognition of fear, with group differences being significantly more pronounced for verbal tasks compared to nonverbal tasks. Notably, although verbal and nonverbal tasks are thought to involve the same core emotion recognition systems (Herba & Phillips, 2004; Phan, Wager, Taylor, & Liberzon, 2002), these two types of tasks also have substantial differences in their demands. Specifically, nonverbal tasks could be completed by discriminating perceptual characteristics between emotional expressions without necessarily understanding the emotional meaning of these expressions (Adolphs, 2002; Palermo et al., 2013). Thus, nonverbal tasks may not tap into group differences in the emotion understanding of different expressions. By contrast, verbal tasks not only involve decoding emotion expressions based on perceptual properties as do nonverbal tasks, but also require access to emotion vocabulary for assigning emotional labels to the expressions (Palermo et al., 2013). The present findings therefore highlight that the particular difficulties autistic individuals have may be due to the linguistic demands for labelling of emotion expressions. It is however worth noting that the inconsistent findings for nonverbal tasks may also be due to a lack of sufficient data. More research is needed to establish whether and how different tasks moderate group differences.

Studies of emotion recognition in ASD have employed different stimulus presentation times, ranging from brief (< 50 ms; Clark et al., 2008; Otsuka et al., 2017) to restricted (ranging from 300 – 3000 ms; Brewer et al., 2016; Ciaramidaro et al., 2018; Corden et al., 2008; Greimel et al., 2014; Griffiths et al., 2019) to unlimited periods (Akechi et al., 2009; Doi et al., 2013; Evers, Kerkhof, Steyaert, Noens, & Wagemans, 2014; Fink et al., 2014; Grossman et al., 2000; Lacroix et al., 2014). For our sensitivity analyses on stimulus presentation time, we hypothesised that extended stimulus presentation time might make it possible for autistic individuals to use compensatory strategies to process emotions, thus reducing the sensitivity of studies to detect any subtle group differences. However, our results showed that significant group differences were present even prior to the removal of datasets without restriction. Thus, the restriction of stimulus presentation time may not be an important contributor to the heterogeneous findings in the literature.

One possible explanation for the lack of influence of stimulus presentation restriction on group differences may relate to the fact that, even with restricted stimulus presentation times in some studies (e.g., 300 – 3000 ms), the durations of stimulus presentation already exceed those required for typical emotion recognition. Indeed, it has been shown that the recognition of discrete emotions first requires the analysis of the perceptual features of emotional expressions (occurring within 200 ms from stimulus onset; Ashley, Vuilleumier, & Swick, 2004; Eimer & Holmes, 2003; Paulmann & Pell, 2009), followed by early semantic analysis in distinguishing between emotional vs. non-emotional input (occurring between 220-300 ms; Bostanov & Kotchoubey, 2004; Paulmann & Pell, 2009). Thus, the time course for recognising emotions has been reported to fall somewhere within the 500-600-ms window for faces (Palermo & Coltheart, 2004; Tracy & Robins, 2008), though it may vary considerably depending on the emotion category (e.g., ranging from 634 ms for happy faces to 2389 ms for fearful faces; Palermo & Coltheart, 2004). In contrast, prosodic and musical emotions can be recognised at above chance level following very brief exposures to the stimuli (e.g., within 100 ms; Nordström & Laukka, 2019). The time course for recognition accuracy rates of prosodic emotions to stabilise has been found to fall within the 500-600 ms window, with varying lengths across emotions (e.g., ranging from 517 ms for fearful prosody to 1486 for disgust prosody; Pell & Kotz, 2011). For music, recognition accuracy rates generally take longer to stabilise (e.g., ranging from 1281 ms for happiness to 1656 for sadness; Nordström & Laukka, 2019). Considering the time courses for accurate emotion recognition across different domains, the presentation times in a great proportion of the datasets with restricted presentation times analysed in the present study (e.g., 300 – 3000 ms) were much beyond the required timeframes for emotion recognition regardless of domain. Thus, it cannot be ruled out that autistic individuals might have used compensatory strategies in processing emotions in these studies, leading to a high amount of heterogeneity retained in the results.

Another explanation for the lack of effect of stimulus presentation time on group differences in our sensitivity analyses may relate to the presentation nature of different stimulus types. For instance, static visual stimuli provide constant emotional cues throughout presentation; thus, longer presentation times would allow for more focused attention to specific features. Dynamic visual stimuli, however, present instantaneous emotional information with varying speeds and intensities from multiple sources, resembling expression of emotions in real life. It has been suggested that in addition to static properties, temporal information such as the temporal order in which a facial expression unfolds (e.g., perhaps the mouth moves first, then eyebrows; Delis et al., 2016; Jack, Garrod, & Schyns, 2014; Jack & Schyns, 2015) and the speed at which the features move (Kamachi et al., 2013; Sato & Yoshikawa, 2004; Sowden, Schuster, Keating, Fraser, & Cook, 2021) also add to emotional judgments of dynamic facial stimuli. Similarly, auditory stimuli, which are inherently dynamic, require listeners to track and integrate acoustic information that develops over time (Schirmer & Adolphs, 2017). Despite that emotions can be identified during brief exposure to the input, recognition accuracy improves with increased durations for both speech prosody (Nordström & Laukka, 2019; Rigoulot, Wassiliwizky, & Pell, 2013) and music (Nordström & Laukka, 2019; Peretz, Gagnon, & Bouchard, 1998). In particular, acoustic cues that develop over time play a larger role for music than for speech, as recognition rates for emotional music take longer to stabilise than for emotional speech under a gating paradigm (Nordström & Laukka, 2019). Thus, different stimulus subtypes, such as words vs. sentences vs. excerpts in speech, or musical chords vs. segments vs. excerpts in music, may bring possible heterogeneity to the findings, obscuring the effects of stimulus presentation time restriction. Taken together, the variability in presentation times and stimulus types across studies may combine to produce mixed results in the literature.

Unfortunately, among the included studies in the current review, only a handful used presentation times within the time windows reported for typical emotion recognition, and there were insufficient data for each stimulus subtype. Thus, it was not possible to partition the datasets with restricted presentation times further for sensitivity analyses. Future studies are required to investigate the interrelationship between presentation time and stimulus type on emotion recognition across domains and modalities.

Finally, a related note on emotion recognition skills in ASD is about response time. Our main meta-analyses on the full datasets suggested that autistic individuals showed significantly longer response times than NT controls when recognising anger, fear, sadness, and the composite emotions. The slower recognition speed could be attributed to a more cautious and time-consuming cognitive approach to the task by autistic individuals, as opposed to the more intuitive strategies that NT individuals use in effortless recognition of emotions (Grossman et la., 2000; Livingston & Happé, 2017; Rutherford & McIntosh, 2007). For instance, autistic individuals may extract facial information from local features, rather than using a higher-level configural processing strategy (Behrmann, Thomas, & Humphreys, 2006; Pelphrey et al., 2002). Another plausible explanation for the slower recognition speed in autistic individuals may reflect the amount of cognitive resources needed for the processing of emotional stimuli in ASD. In relation to this, atypical orientation towards face-like objects (Akechi et al., 2014; Guillon et al., 2016; Pavlova et al., 2017; Ryan, Stafford, & King, 2016) and speech sounds (Čeponienė et al., 2003; Lepistö et al., 2005), as well as reduced sensitivity towards prosodic and musical expressivity (Bhatara et al., 2010; Brown, 2017; Gebauer, Skewes, Hørlyck, & Vuust, 2014) may contribute to such delays. Alternatively, the slower recognition speed across domains might be the result of overall slower processing speed in ASD. In line with this, prior studies have reported a fundamental impairment in the speed at which autistic individuals can process information (Haigh, Walsh, Mazefsky, Minshew, & Eack, 2018; Mayes et al., 2007;

Velikonja, Fett, & Velthorst, 2019). Furthermore, the relationship between slower processing speed and measures of social cognition have also been illustrated (Haigh et al., 2018; Hedvall et al., 2013; Oliveras-Rentas, Kenworthy, Roberson, Martin, & Wallace, 2012; Russo-Ponsaran et al., 2015). These possibilities, however, require further investigations into whether and how processing strategy, cue sensitivity, and processing speed contribute to the accuracy and speed of emotion recognition in ASD.

4.5. Limitations and future research directions

Insufficient reporting of study data (k = 138) has resulted in limitation of our analyses. In an era of open science, it is recommended that authors make their data accessible, in order to optimise data usage in the field. A considerable number of papers did not (consistently) report exact *p*-values, undermining accurate interpretation of results in relation to study hypotheses. To infer the importance of results, it is also recommended that researchers shift towards a meta-analytic thinking orientation and report effect sizes and confidence intervals along with statistical test results (Henson, 2006). The reporting of effect sizes (including nonsignificant results) allows explicit comparisons to be made between studies and enables all relevant data to be included in future syntheses, thus reducing potential influence of publication bias.

The generalisability of our study sample characteristics may be compromised due to a lack of coverage of ages within individual studies. Only 8% of the papers studied both autistic children and adults. When investigating the developmental trajectory of emotion recognition in ASD, it is important to control for heterogeneity brought by different tasks from different studies. More cross-sectional and longitudinal studies would be favoured to examine the effects of age on emotion recognition in ASD. Furthermore, with the growing number of interventions available for enhancing social communication skills in ASD (see Berggren et al., 2018 and Kouo & Egel, 2016 for reviews), it is likely that these studies have included participants who had undergone training programs prior to the experiments. We speculate that any long-term training effects brought into experimental settings may have obscured the true effects of emotion recognition ability in ASD observed across studies in a subtle and inconsistent way. It may therefore be worthwhile for new studies to take note of the interventions that autistic participants may have undertaken, such as in Wright et al. (2008).

The results of our main meta-analyses were based on evidence disproportionately distributed across domains (i.e., with the majority on human faces) and tasks (i.e., with the majority employing a verbal labelling task). The lack of data for the different categories further led to our decision on prioritising datasets based on domains and tasks that were most commonly used across studies, in an attempt to reduce heterogeneity. Subgroup comparisons, nonetheless, did not find significant differences between domain subgroups across emotions nor between task subgroups for happiness, sadness, and anger. This suggests that the impact of such prioritisation on the overall results is likely to be limited. Future meta-analytic work with more data in these under-researched areas will provide clearer indications of results and further insights into the effects of these moderators with increased statistical power.

A number of studies have explored emotion recognition ability at varying intensities in ASD. As intensity level was based on either ratings obtained from validation or different morphed continua used (e.g., between neutral and an emotional expression, Law Smith et al., 2010; Otsuka, Uono, Yoshimura, Zhao, & Toichi, 2017, or between emotions within a pair such as fear vs. happiness, Wang & Adolphs, 2017), we could not reliably compare results at low (25%) or intermediate (50%) intensity levels across studies. Thus, only findings at the highest intensity level were included in the current meta-analyses. One may therefore expect stronger effects to be revealed with data collating from stimuli with lower intensities. However, despite taking a potentially less sensitive measure in detecting subtle group differences

(Mazefsky & Oswald, 2007; Rump et al., 2009), general emotion recognition deficits across emotions were still observed in the ASD groups.

Finally, there has been an ongoing debate on whether co-occurring alexithymia, but not ASD per se, may be responsible for emotion recognition difficulties documented in the autistic population (Bird & Cook, 2013; Cook, Brewer, Shah, & Bird, 2013; Kinnaird, Stewart, & Tchanturia, 2019; see also Sivathasan, Fernandes, Burack, & Quintin, 2020 for a recent review). Alexithymia, characterised by a lack of fluency in identifying and describing one's own emotions and feelings (Bird & Cook, 2013), has been reported to be highly comorbid with ASD, affecting approximately 50% of the autistic population (Berthoz & Hill, 2005; Milosavljevic et al., 2016). Several studies have provided supporting evidence for the alexithymia hypothesis, by noting the significant relevance of alexithymia to emotion recognition difficulties with facial and prosodic expressions in ASD (Cook et al., 2013; Heaton et al., 2012; Ketelaars, In'T Velt, Mol, Swaab, & Van Rijn, 2016). Conversely, other studies have reported a lack of contribution of alexithymia to emotion recognition difficulties in ASD (Kliemann, Rosenblau, Bölte, Heekeren, & Dziobek, 2013; Stephenson, Luke, & South, 2019). Although there have been increasing efforts to consider alexithymia as a potential candidate in accounting for emotion recognition performance in autistic individuals over the past decade (Cook et al., 2013; Heaton et al., 2012; Keating, Fraser, Sowden, & Cook, 2021; Ketelaars et al., 2016; Kliemann et al., 2013; Milosavljevic et al., 2016), available data for the different emotion types examined in the current review remain scarce. Thus, adhering to the recommendations that meta-regressions should generally not be performed on data from less than 10 studies (Deeks et al., 2019), we did not examine the moderating effect of alexithymia on group differences in emotion recognition in the present study. To enable a meta-analysis of the findings, future studies would need to include measures of alexithymia when investigating emotion recognition in ASD.

5. Conclusions

In conclusion, this quantitative synthesis of the current literature found that autistic individuals demonstrate general emotion recognition deficits across six basic emotions and their composite, while longer response times were also found for anger, fear, sadness, and the composite in ASD versus NT. The general impairments in recognition accuracy were shown to be robust and were not driven by differences in IQ matching and stimulus presentation time restriction – though the severity of these impairments was less pronounced for a subset of emotions when full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust) had been undertaken. These results suggest that sample characteristics and experimental designs interact to give rise to the heterogeneity seen in the literature. By investigating all these factors simultaneously in a large dataset, together with rigorous inclusion criteria and robust analysis procedures, we show the moderating effects of age, domain, task demand on emotion recognition in ASD relative to NT. The group effect was more pronounced in adults with increased task demands. Although insufficient data prevent reliable conclusions to be drawn on the effect of domain, deficits were consistently found for human faces but not for nonhuman faces, with deficits for speech prosody and music specific to certain emotions. Task demand moderated emotion recognition, with autistic individuals performing worse on tasks requiring verbal knowledge about emotions than nonverbal tasks. Full-scale/verbal/nonverbal IQ measures, by contrast, were not important moderators of group effects. Further work is needed to extend the literature particularly on emotion recognition of prosodic, musical, and nonhuman facial emotions in order to draw an unbiased comparison across domains. Future research should also focus on processing strategies that autistic individuals employ for different types of stimuli in both bottom-up and top-down processes, as well as investigate the interactions between stimulus type and stimulus presentation time on emotion recognition. Given the positive consequences of learned strategies in fulfilling life experiences, such investigations will provide insights into the optimal contexts for autistic individuals to accomplish successful social interactions in daily life.

Acknowledgments

This study was supported by a European Research Council (ERC) Starting Grant, ERC-StG-2015, CAASD, 678733, to F.L. We thank all of the authors of primary literature who responded to our emails and requests for further details, and Kathryn Ansell for help with preparing the summary table. We would also like to thank two anonymous reviewers and the Associate Editor, Dr. Kinga Morsanyi, for their insightful comments and suggestions on the earlier versions of the manuscript.

References (References marked with an asterisk indicate studies included in the meta-analysis)

Adams, J., Hillier-Brown, F. C., Moore, H. J., Lake, A. A., Araujo-Soares, V., White, M., & Summerbell, C. (2016). Searching and synthesising 'grey literature' and 'grey information' in public health: Critical reflections on three case studies. *Systematic Reviews*, *5*(1), 164. https://doi.org/10.1186/s13643-016-0337-y

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62. https://doi.org/10.1177/1534582302001001003

Aiken, M. (2019). An updated evaluation of Google Translate accuracy. *Studies in Linguistics and Literature*, *3*, p253. https://doi.org/10.22158/sll.v3n3p253

*Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2009). Does gaze direction modulate facial expression processing in children with autism spectrum disorder? *Child Development*, *80*(4), 1134–1146. https://doi.org/10.1111/j.1467-8624.2009.01321.x

*Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48*(10), 2841–2851. https://doi.org/10.1016/j.neuropsychologia.2010.05.026

Akdeniz, G. (2020). Brain activity underlying face and face pareidolia processing: an ERP study. *Neurological Sciences*, *41*(6), 1557-1565. https://doi.org/10.1007/s10072-019-04232-4

Allen, R., Davis, R., & Hill, E. (2013). The effects of autism and alexithymia on physiological and verbal responsiveness to music. *Journal of Autism and Developmental Disorders*, *43*(2), 432–444. https://doi.org/10.1007/s10803-012-1587-8

Alshaban, F., Aldosari, M., Al-Shammari, H., El-Hag, S., Ghazal, I., Tolefat, M., … Fombonne, E. (2019). Prevalence and correlates of autism spectrum disorder in Qatar: a national study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *60*(12), 1254–1268. https://doi.org/10.1111/jcpp.13066

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition (DSM-5)*. Arlington, VA: American Psychiatric Publishing.

Argstatter, H. (2016). Perception of basic emotions in music: Culture-specific or multicultural? *Psychology of Music*, *44*(4), 674–690. https://doi.org/10.1177/0305735615589214

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport*, *15*(1), 211–216. https://doi.org/10.1097/01.wnr.0000091411.19795.f5

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. https://doi.org/10.1080/17470910601040772

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia*, *47*(13), 3023-3029. https://doi.org/10.1016/j.neuropsychologia.2009.05.019

*Baker, K. F., Montgomery, A. A., & Abramson, R. (2010). Brief report: Perception and lateralization of spoken emotion by youths with high-functioning forms of autism. *Journal of Autism and Developmental Disorders*, *40*(1), 123–129. https://doi.org/10.1007/s10803-009-0841-1

Bang, J. Y., Sharda, M., & Nadig, A. S. (2020). Towards greater transparency in neurodevelopmental disorders research: Use of a proposed workflow and propensity scores to facilitate selection of matched groups. *Journal of Neurodevelopmental Disorders*, *12*(1), 20. https://doi.org/10.1186/s11689-020-09321-6

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience and Biobehavioral Reviews*, *24*(3), 355–364. https://doi.org/10.1016/S0149-7634(00)00011-7

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *British Journal of Psychiatry*, *194*(6), 500–509. https://doi.org/10.1192/bjp.bp.108.059345

*Bast, N., Banaschewski, T., Dziobek, I., Brandeis, D., Poustka, L., & Freitag, C. M. (2019). Pupil Dilation Progression Modulates Aberrant Social Cognition in Autism Spectrum Disorder. *Autism Research*, *12*(11), 1680–1692. https://doi.org/10.1002/aur.2178

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: Visual processing in autism. *Trends in Cognitive Sciences*, *10*(6), 258–264. https://doi.org/10.1016/j.tics.2006.05.001

*Berggren, S., Engström, A. C., & Bölte, S. (2016). Facial affect recognition in autism, ADHD and typical development. *Cognitive Neuropsychiatry*, *21*(3), 213–227. https://doi.org/10.1080/13546805.2016.1171205

Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P. B., Bölte, S., & Jonsson, U. (2018). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation*, *21*(3), 141–154. https://doi.org/10.1080/17518423.2017.1305004

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European psychiatry*, *20*(3), 291-298. https://doi.org/10.1016/j.eurpsy.2004.06.013

Bhatara, A., Quintin, E.-M., Levy, B., Bellugi, U., Fombonne, E., & Levitin, D. J. (2010). Perception of Emotion in Musical Performance in Adolescents with Autism Spectrum Disorders. *Autism Research*, *3*(5), 214–225. https://doi.org/10.1002/aur.147

Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational psychiatry*, *3*(7), https://doi.org/10.1038/tp.2013.61

Blackstock, E. G. (1978). Cerebral asymmetry and the development of early infantile autism. *Journal of Autism and Childhood Schizophrenia*, *8*(3), 339–353.

Blair, R. J. R., Frith, U., Smith, N., Abell, F., & Cipolotti, L. (2002). Fractionation of visual memory: Agency detection and its impairment in autism. *Neuropsychologia*, *40*(1), 108–118. https://doi.org/10.1016/S0028-3932(01)00069-0

Boat, T. F., & Wu, J. T. (2015). Clinical characteristics of intellectual disabilities. In *Mental disorders and disabilities among low-income children* (pp. 169–178). Washington D.C: The National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK332877/

*Boggs, K. M., & Gross, A. M. (2010). Cue Salience in Face Processing by High Functioning Individuals with Autism Spectrum Disorders. *Journal of Developmental and Physical Disabilities*, *22*(6), 595–613. https://doi.org/10.1007/s10882-010-9198-9

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, *66*(12), 2426-2442. https://doi.org/10.1080/17470218.2013.789065

Boraston, Z., Blakemore, S. J., Chilvers, R., & Skuse, D. (2007). Impaired sadness recognition is linked to social interaction deficit in autism. *Neuropsychologia*, *45*(7), 1501–1510. https://doi.org/10.1016/j.neuropsychologia.2006.11.010

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Borenstein, M., & Higgins, J. P. T. (2013). Meta-analysis and subgroups. *Prevention Science*, *14*(2), 134–143. https://doi.org/10.1007/s11121-013-0377-7

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology*, *41*(2), 259–268. https://doi.org/10.1111/j.1469-8986.2003.00142.x

*Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *41*(7), 847–857. https://doi.org/10.1111/1469-7610.00672

Braverman, M., Fein, D., Lucci, D., & Waterhouse, L. (1989). Affect comprehension in children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *19*(2), 301-316.

*Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, *5*(4), 1567–1573. https://doi.org/10.1016/j.rasd.2011.03.002

*Brewer, R., Biotti, F., Catmur, C., Press, C., Happé, F., Cook, R., & Bird, G. (2016). Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders. *Autism Research*, *9*(2), 262–271. https://doi.org/10.1002/aur.1508

Bruce, V., Campbell, R. N., Doherty-Sneddon, G., Import, A., Langton, S., McAuley, S., & Wright, R. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology*, *18*(3), 319–333. https://doi.org/10.1348/026151000165715

Brosnan, M., Johnson, H., Grawmeyer, B., Chapman, E., & Benton, L. (2015). Emotion recognition in animated compared to human stimuli in adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(6), 1785–1796. https://doi.org/10.1007/s10803-014-2338-9

Brown, L. S. (2017). The influence of music on facial emotion recognition in children with autism spectrum disorder and neurotypical children. *Journal of Music Therapy*, *54*(1), 55–79. https://doi.org/10.1093/jmt/thw017

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human perception and performance*, *26*(2), 527. https://doi.org/10.1037/0096-1523.26.2.527

*Campbell, R., Lawrence, K., Mandy, W., Mitra, C., Jeyakuma, L., & Skuse, D. (2006). Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. *Development and Psychopathology*, *18*(1), 99–118. https://doi.org/10.1017/S0954579406060068

Caria, A., Venuti, P., & De Falco, S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, *21*(12), 2838–2849. https://doi.org/10.1093/cercor/bhr084

CASP, U. (2017). Critical Appraisal Skills Programme (CASP). *Qualitative research checklist, 31,* 13.

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

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., Yaguchi, K., & Lepisto, T. (2003). Speech–sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proceedings of the National Academy of Sciences*, *100*(9), 5567–5572. https://doi.org/10.1073/pnas.0835631100

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. https://doi.org/10.1016/j.tics.2012.02.007

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, *33*(2), 218–236. https://doi.org/10.1111/bjdp.12075

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, *8*(1), 1–17. https://doi.org/10.1038/s41598-018-26889-1

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), 803–809. https://doi.org/10.1037/a0014124

*Ciaramidaro, A., Bölte, S., Schlitt, S., Hainz, D., Poustka, F., Weber, B., … Walter, H. (2018). Transdiagnostic deviant facial recognition for implicit negative emotion in autism and schizophrenia. *European Neuropsychopharmacology*, *28*(2), 264–275. https://doi.org/10.1016/j.euroneuro.2017.12.005

Clarivate Analytics (2016). EndNote X8. Available at: http://endnote.com/sites/en/files/m/pdf/en-x8-qrg-mac.pdf

Çoğaltay, N., & Karadağ, E. (2015). Introduction to meta-analysis. In E. Karadag (Ed.), *Leadership and Organizational Outcomes: Meta-Analysis of Empirical Studies* (pp. 19–28). https://doi.org/10.1007/978-3-319-14908-0

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological science*, *24*(5), 723-732. https://doi.org/10.1177/0956797612463582

*Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia*, *46*(1), 137–147. https://doi.org/10.1016/j.neuropsychologia.2007.08.005

Corden, B., Chilvers, R., & Skuse, D. (2008). Emotional modulation of perception in Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, *38*(6), 1072–1080. https://doi.org/10.1007/s10803-007-0485-y

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition and Emotion*, *27*(4), 658–684. https://doi.org/10.1080/02699931.2012.732559

*Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: More convergence than divergence. *Psychological Medicine*, *40*(4), 569–579. https://doi.org/10.1017/S003329170999078X

*Davidson, D., Hilvert, E., Misiunaite, I., Kerby, K., & Giordano, M. (2019). Recognition of facial emotions on human and canine faces in children with and without autism spectrum disorders. *Motivation and Emotion*, *43*(1), 191–202. https://doi.org/10.1007/s11031-018-9736-9

Davitz, J. R., & Beldoch, M. (1964). The communication of emotional meaning. McGraw-Hill.

de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 149–158. https://doi.org/10.1002/wcs.1335

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op her Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 200–213. https://doi.org/10.1076/jcen.24.2.200.989

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2001). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. *Systematic reviews in health care: meta-analysis in context*, 285-312. https://doi.org/10.1002/9780470693926.ch15

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook

Del Re, A.C. (2013). compute.es: Compute Effect Sizes. R package version 0.2-2. URL http://cran.r-project.org/web/packages/compute.es

Delis, I., Chen, C., Jack, R. E., Garrod, O. G. B., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision*, *16*(8), 14–14. https://doi.org/10.1167/16.8.14

Derntl, B., Seidel, E. M., Kainz, E., & Carbon, C. C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, *38*(12), 1849-1862. https://doi.org/10.1068/p6448

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*(2), 199–210. https://doi.org/10.1023/B:JADD.0000022610.09668.4c

Dobrescu, A., Nussbaumer, S. B., Klerings, I., Wagner, G., Persad, E., Sommer, I., Herkner, H., & Gartlehner, G. (2021). Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: A systematic review: Excluding English-language publications a valid shortcut. *Journal of Clinical Epidemiology*. https://doi.org/10.1016/j.jclinepi.2021.04.012

*Doi, H., Fujisawa, T. X., Kanai, C., Ohta, H., Yokoi, H., Iwanami, A., … Shinohara, K. (2013). Recognition of facial expressions and prosodic cues with graded emotional intensities in adults with asperger syndrome. *Journal of Autism and Developmental Disorders*, *43*(9), 2099–2113. https://doi.org/10.1007/s10803-013-1760-8

Donath, J. (2001). Mediated faces. In M. Beynon, C. L. Nehaniv, & K. Dautenhahn (Eds.), *Cognitive Technology: Instruments of Mind Proceedings of the 4th International Conference*. https://doi.org/10.1007/3-540-44617-6_34

Duggleby, W., Holtslander, L., Kylma, J., Duncan, V., Hammond, C., & Williams, A. (2010). Metasynthesis of the hope experience of family caregivers of persons with chronic illness. *Qualitative Health Research*, *20*(2), 148–158. https://doi.org/10.1177/1049732309358329

Duval, S. (2005). The trim and fill method. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments* (pp. 127–144). Chichester, England: John Wiley & Sons. https://doi.org/10.1002/0470870168

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method. *Biometrics*, *56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

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. https://doi.org/10.1207/s15374424jccp3501

*Eack, S. M., Mazefsky, C. A., & Minshew, N. J. (2015). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism*, *19*(3), 308–315. https://doi.org/10.1177/1362361314520755

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867-872. https://doi.org/10.1016/0140-6736(91)90201-Y

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.316.7129.469

Eimer, M., & Holmes, A. (2003). The role of spatial attention in the processing of facial expression: An ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective and Behavioral Neuroscience*, *3*(2), 97–110.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75. https://doi.org/10.1007/BF01115465

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guide-lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings*. New York: Pergamon.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240.

*Evers, K., Kerkhof, I., Steyaert, J., Noens, I., & Wagemans, J. (2014). No differences in emotion recognition strategies in children with autism spectrum disorder: Evidence from hybrid faces. *Autism Research and Treatment*, *2014*, 1–8. https://doi.org/10.1155/2014/345878

*Fink, E., de Rosnay, M., Wierda, M., Koot, H. M., & Begeer, S. (2014). Brief report: Accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(9), 2363–2368. https://doi.org/10.1007/s10803-014-2084-z

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A meta-analysis of gaze differences to social and nonsocial information between Individuals with and without autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(7), 546–555. https://doi.org/10.1016/j.jaac.2017.05.005

*Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., … Golan, O. (2016). Basic and complex emotion recognition in children with autism: Cross-cultural findings. *Molecular Autism*, *7*(1), 52. https://doi.org/10.1186/s13229-016-0113-9

Gale, C. M., Eikeseth, S., & Klintwall, L. (2019). Children with autism show atypical preference for non-social stimuli. *Scientific Reports*, *9*(1), 10355. https://doi.org/10.1038/s41598-019-46705-8

Gebauer, L., Skewes, J., Hørlyck, L., & Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*, *6*, 370–378. https://doi.org/10.1016/j.nicl.2014.08.025

Gebauer, L., Skewes, J., Westphael, G., Heaton, P., & Vuust, P. (2014). Intact brain processing of musical emotions in autism spectrum disorder, but more cognitive load and arousal in happy vs. sad music. *Frontiers in Neuroscience*, *8*, 1–10. https://doi.org/10.3389/fnins.2014.00192

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. https://doi.org/10.1023/A:1005609629218

Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody recognition in adults with high-functioning autism spectrum disorders: From psychoacoustics to cognition. *Autism Research*, *8*(2), 153–163. https://doi.org/10.1002/aur.1432

*Greimel, E., Schulte-Rüther, M., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., & Konrad, K. (2014). Impairment in face processing in autism spectrum disorder: A developmental perspective. *Journal of Neural Transmission*, *121*(9), 1171–1181. https://doi.org/10.1007/s00702-014-1206-2

Grelotti, D. J., Klin, A. J., Gauthier, I., Skudlarski, P., Cohen, D. J., Gore, J. C., … Schultz, R. T. (2005). fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia*, *43*(3), 373–385. https://doi.org/10.1016/j.neuropsychologia.2004.06.015

*Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2019). Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity. *Journal of Autism and Developmental Disorders*, 1–11. https://doi.org/10.1007/s10803-017-3091-7

Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech Language and Hearing Research*, *53*(3), 778–793. https://doi.org/10.1044/1092-4388(2009/08-0127)

*Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal bias in recognition of facial emotions in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *41*(3), 369–379. https://doi.org/10.1111/1469-7610.00621

Guillon, Q., Rogé, B., Afzali, M. H., Baduel, S., Kruck, J., & Hadjikhani, N. (2016). Intact perception but abnormal orientation towards face-like objects in young children with ASD. *Scientific Reports*, *6*, 1–9. https://doi.org/10.1038/srep22119

Haigh, S. M., Walsh, J. A., Mazefsky, C. A., Minshew, N. J., & Eack, S. M. (2018). Processing Speed is Impaired in Adults with Autism Spectrum Disorder, and Relates to Social Communication Abilities. *Journal of Autism and Developmental Disorders*, *48*(8), 2653–2662. https://doi.org/10.1007/s10803-018-3515-z

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. https://doi.org/10.1007/s11065-010-9138-6

*He, Y., Su, Q., Wang, L., He, W., Tan, C., Zhang, H., … Chen, Y. (2019). The characteristics of intelligence profile and eye gaze in facial emotion recognition in mild and moderate preschoolers with autism spectrum disorder. *Frontiers in Psychiatry*, *10*(JUN), 1–15. https://doi.org/10.3389/fpsyt.2019.00402

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1443–1447. https://doi.org/10.1098/rstb.2008.0327

Heaton Hermelin, B., Pring, L., P. (1999). Can children with autism spectrum disorder perceive affect in music? An experimental investigation. *Psychological Medicine*, *29*(6), 1405–1410. https://doi.org/10.1017/S0033291799001221

*Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., & Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum

disorder and typical development. *Psychological Medicine*, *42*(11), 2453–2459. https://doi.org/10.1017/S0033291712000621
Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107
Hedvall, Å., Fernell, E., Holm, A., Åsberg Johnels, J., Gillberg, C., & Billstedt, E. (2013). Autism, processing speed, and adaptive functioning in preschool children. *The Scientific World Journal*, *2013*. https://doi.org/10.1155/2013/158263
*Heikkinen, J., Jansson-Verkasalo, E., Toivanen, J., Suominen, K., Väyrynen, E., Moilanen, I., & Seppänen, T. (2010). Perception of basic emotions from speech prosody in adolescents with Asperger's syndrome. *Logopedics Phoniatrics Vocology*, *35*(3), 113–120. https://doi.org/10.3109/14015430903311184
Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, *34*(5), 601–629. https://doi.org/10.1177/0011000005283558
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 and Allied Disciplines*, *45*(7), 1185–1198. https://doi.org/10.1111/j.1469-7610.2004.00316.x
Hernandez, N., Metzger, A., Magné, R., Bonnet-Brilhault, F., Roux, S., Barthelemy, C., & Martineau, J. (2009). Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *Neuropsychologia*, *47*(4), 1004–1012. https://doi.org/10.1016/j.neuropsychologia.2008.10.023
Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557
Hobson, R. P., Ouston, J., & Lee, A. (1989). Naming emotion in faces and voices: Abilities and disabilities in autism and mental retardation. *British Journal of Developmental Psychology*, *7*(3), 237–250. https://doi.org/10.1111/j.2044-835x.1989.tb00803.x
Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., & Roberts, N. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *NeuroReport*, *11*(13), 2931–2935. https://doi.org/10.1097/00001756-200009110-00020
Howlin, P. (1998). Psychological and educational treatments for autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(3), 307-322. https://doi.org/10.1017/S0021963097002138
*Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research*, *10*(12), 1991–2001. https://doi.org/10.1002/aur.1847
IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247
*Isomura, T., Ito, H., Ogawa, S., & Masataka, N. (2014a). Absence of predispositional attentional sensitivity to angry faces in children with autism spectrum disorders. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep07525
*Isomura, T., Ogawa, S., Yamada, S., Shibasaki, M., & Masataka, N. (2014b). Preliminary evidence that different mechanisms underlie the anger superiority effect in children with and without Autism Spectrum Disorders. *Frontiers in Psychology*, *5*, 1–8. https://doi.org/10.3389/fpsyg.2014.00461

Iwanaga, M., Ikeda, M., & Iwaki, T. (1996). The effects of repetitive exposure to music on subjective and physiological responses. *Journal of Music Therapy*, *33*(3), 219–230. https://doi.org/10.1093/jmt/33.3.219

Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. *Current Biology*, *24*(2), 187–192. https://doi.org/10.1016/j.cub.2013.11.064

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Jarrold, C., & Brock, J. (2004). To Match or Not to Match? Methodological Issues in Autism-Related Research. *Journal of Autism and Developmental Disorders*, *34*(1), 81–86. https://doi.org/10.1023/B:JADD.0000018078.82542.ab

*Järvinen, A., Ng, R., Crivelli, D., Neumann, D., Arnold, A. J., Woo-Vonhoogenstyn, N., … Bellugi, U. (2016). Social functioning and autonomic nervous system sensitivity across vocal and musical emotion in Williams syndrome and autism spectrum disorder. *Developmental Psychobiology*, *58*(1), 17–26. https://doi.org/10.1002/dev.21335

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

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2013). Dynamic properties influence the perception of facial expressions. *Perception*, *42*(11), 1266–1278. https://doi.org/10.1068/p3131n

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759–763. https://doi.org/10.1038/77664

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2021). Differences Between Autistic and Non-Autistic Adults in the Recognition of Anger from Facial Motion Remain after Controlling for Alexithymia. *Journal of Autism and Developmental Disorders*, *0123456789*. https://doi.org/10.1007/s10803-021-05083-9

Keltner, D., Tracy, J., Sauter, D., Cordaro, D., & McNeil, G. (2016). Emotional expression. *Handbook of Emotion*.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

*Ketelaars, M. P., Mol, A., Swaab, H., & van Rijn, S. (2016). Emotion recognition and alexithymia in high functioning females with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *21*, 51–60. https://doi.org/10.1016/j.rasd.2015.09.006

*Kim, K., Rosenthal, M. Z., Gwaltney, M., Jarrold, W., Hatt, N., McIntyre, N., … Mundy, P. (2015). A virtual joy-stick study of emotional responses and social motivation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(12), 3891–3899. https://doi.org/10.1007/s10803-014-2036-7

Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., … Grinker, R. R. (2011). Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, *168*(9), 904–912. https://doi.org/10.1176/appi.ajp.2011.10101532

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: a systematic review and meta-analysis. *European Psychiatry*, *55*, 80-89. https://doi.org/10.1016/j.eurpsy.2018.09.004

Kliemann, D., Rosenblau, G., Bölte, S., Heekeren, H. R., & Dziobek, I. (2013). Face puzzle—two new video-based tasks for measuring explicit and implicit aspects of facial emotion recognition. *Frontiers in psychology*, *4*, 376. https://doi.org/10.3389/fpsyg.2013.00376

*Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *32*(28), 9469–9476. https://doi.org/10.1523/jneurosci.5294-11.2012

*Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *30*(37), 12281–12287. https://doi.org/10.1523/JNEUROSCI.0688-10.2010

*Król, M. E., & Król, M. (2019). A novel machine learning analysis of eye-tracking data reveals suboptimal visual information extraction from facial stimuli in individuals with autism. *Neuropsychologia*, *129*(April), 397–406. https://doi.org/10.1016/j.neuropsychologia.2019.04.022

Koldewyn, K., Whitney, D., & Rivera, S. M. (2010). The psychophysics of visual motion and global form processing in autism. *Brain*, *133*(2), 599-610. https://doi.org/10.1093/brain/awp272

Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech: A review. *International Journal of Speech Technology*, *15*(2), 99–117. https://doi.org/10.1007/s10772-011-9125-1

Kopec, J., Hillier, A., & Frye, A. (2014). The valency of music has different effects on the emotional responses of those with autism spectrum disorders and a comparison group. *Music Perception: An Interdisciplinary Journal*, *31*(5), 436–443. https://doi.org/10.1525/mp.2014.31.5.436

Kouo, J. L., & Egel, A. L. (2016). The effectiveness of interventions in teaching emotion recognition to children with autism spectrum disorder. *Review Journal of Autism and Developmental Disorders*, *3*(3), 252–265. https://doi.org/10.1007/s40489-016-0081-1

*Kujala, T., Lepistö, T., Nieminen-Von Wendt, T., Näätänen, P., & Näätänen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neuroscience Letters*, *383*(3), 260–265. https://doi.org/10.1016/j.neulet.2005.04.048

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. https://doi.org/10.1007/s10803-009-0700-0

Khawar, R., Malik, F., Maqsood, S., Yasmin, T., & Habib, S. (2013). Age and gender differences in emotion recognition ability and intellectual functioning. *Journal of Behavioural Sciences*, *23*(3), 53-71.

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. https://doi.org/10.1016/j.rasd.2014.05.012

*Law Smith, M. J., 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. https://doi.org/10.1016/j.neuropsychologia.2010.03.008

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00761

Le Sourn-Bissaoui, S., Aguert, M., Girard, P., Chevreuil, C., & Laval, V. (2013). Emotional speech comprehension in children and adolescents with autism spectrum disorders.

Journal of Communication Disorders, *46*(4), 309–320. https://doi.org/10.1016/j.jcomdis.2013.03.002
Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, *1066*(1–2), 147–157. https://doi.org/10.1016/j.brainres.2005.10.052
*Li, Y., Fang, H., Zheng, W., Qian, L., Xiao, Y., Wu, Q., … Ke, X. (2017). A fiber tractography study of social-emotional related fiber tracts in children and adolescents with autism spectrum disorder. *Neuroscience Bulletin*, *33*(6), 722–730. https://doi.org/10.1007/s12264-017-0155-9
Lindner, J. L., & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769–777. https://doi.org/10.1007/s10803-006-0105-2
*Lindström, R., Lepistö-Paisley, T., Makkonen, T., Reinvall, O., Nieminen-von Wendt, T., Alén, R., & Kujala, T. (2018). Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clinical Neurophysiology*, *129*(11), 2411–2420. https://doi.org/10.1016/j.clinph.2018.08.018
Livingston, L. A., & Happé, F. (2017). *Neuroscience and Biobehavioral Reviews Conceptualising compensation in neurodevelopmental disorders: Re fl ections from autism spectrum disorder*. *80*(May), 729–742. https://doi.org/10.1016/j.neubiorev.2017.06.005
*Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7
Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology*, *26*(4), 933–945. https://doi.org/10.1017/S0954579414000479
Martzoukou, M., Papadopoulou, D., & Kosmidis, M. H. (2017). The comprehension of syntactic and affective prosody by adults with autism spectrum disorder without accompanying cognitive deficits. *Journal of Psycholinguistic Research*, *46*(6), 1573–1595. https://doi.org/10.1007/s10936-017-9500-4
Mayes, S. D., Calhoun, S. L., Dickerson, S., Learning, S. L. C., Mayes, S. D., & Calhoun, S. L. (2007). *Learning , Attention , Writing , and Processing Speed in Typical Children and Children with ADHD , Autism , Anxiety , Depression , and Oppositional- Defiant Disorder OPPOSITIONAL-DEFIANT DISORDER*. *7049*. https://doi.org/10.1080/09297040601112773
Mazefsky, C. A., & Oswald, D. P. (2007). Emotion perception in Asperger's syndrome and high-functioning autism: The importance of diagnostic criteria and cue intensity. *Journal of Autism and Developmental Disorders*, *37*(6), 1086–1095. https://doi.org/10.1007/s10803-006-0251-6
McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*(3), 325-334. https://doi.org/10.1111/j.2044-8309.1995.tb01067.x
Milosavljevic, B., Leno, V. C., Simonoff, E., Baird, G., Pickles, A., Jones, C. R., ... & Happé, F. (2016). Alexithymia in adolescents with autism spectrum disorder: its relationship to internalising difficulties, sensory modulation and social cognition. *Journal of autism and developmental disorders*, *46*(4), 1354-1367. https://doi.org/10.1007/s10803-015-2670-8

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., ... & Clark, J. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Chinese edition). *Journal of Chinese Integrative Medicine*, *7*(9), 889-896.
Mohn, C., Argstatter, H., & Wilker, F. W. (2011). Perception of six basic emotions in music. *Psychology of Music*, *39*(4), 503–517. https://doi.org/10.1177/0305735610378183
Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences*, *1252*(1), 318–324. https://doi.org/10.1111/j.1749-6632.2012.06465.x
Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology*, *86*(1), 67–84. https://doi.org/10.1016/S0022-0965(03)00102-4
Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B., & Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care*, *28*(2), 138–144. https://doi.org/10.1017/S0266462312000086
Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry*, *41*(8), 1057–1065. https://doi.org/10.1111/1469-7610.00693
Neimann Rasmussen, L., & Montgomery, P. (2018). The prevalence of and factors associated with inclusion of non-English language studies in Campbell systematic reviews: A survey and meta-epidemiological study. *Systematic Reviews*, *7*(1), 129. https://doi.org/10.1186/s13643-018-0786-6
Nordström, H., & Laukka, P. (2019). The time course of emotion recognition in speech and music. *The Journal of the Acoustical Society of America*, *145*(5), 3058–3074. https://doi.org/10.1121/1.5108601
Nowell, K. P., Schanding, G. T., Kanne, S. M., & Goin-Kochel, R. P. (2015). Cognitive profiles in youth with autism spectrum disorder: An investigation of base rate discrepancies using the Differential Ability Scales—Second Edition. *Journal of Autism and Developmental Disorders*, *45*(7), 1978–1988. https://doi.org/10.1007/s10803-014-2356-7
Nuske, H. J., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion*, *27*(6), 1042–1061. https://doi.org/10.1080/02699931.2012.762900
*O'Connor, K. (2007). Brief report: Impaired identification of discrepancies between expressive faces and voices in adults with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *37*(10), 2008–2013. https://doi.org/10.1007/s10803-006-0345-1
*Oerlemans, A. M., Van Der Meer, J. M. J., Van Steijn, D. J., De Ruiter, S. W., De Bruijn, Y. G. E., De Sonneville, L. M. J., … Rommelse, N. N. J. (2014). Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European Child and Adolescent Psychiatry*, *23*(5), 257–271. https://doi.org/10.1007/s00787-013-0446-2
Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R. B., Martin, A., & Wallace, G. L. (2012). WISC-IV profile in high-functioning autism spectrum disorders: Impaired processing speed is associated with increased autism communication symptoms and decreased adaptive communication abilities. *Journal of Autism and Developmental Disorders*, *42*(5), 655–664. https://doi.org/10.1007/s10803-011-1289-7
*Otsuka, S., Uono, S., Yoshimura, S., Zhao, S., & Toichi, M. (2017). Emotion perception mediates the predictive relationship between verbal ability and functional outcome in high-functioning adults with autism spectrum disorder. *Journal of Autism and*

Developmental Disorders, *47*(4), 1166–1182. https://doi.org/10.1007/s10803-017-3036-1
Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry*, *31*(3), 343–361. https://doi.org/10.1111/j.1469-7610.1990.tb01574.x
Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233–240. https://doi.org/10.1111/jebm.12266
Page, M.J., Higgins, J.P.T., Sterne, J.A.C. (editors). Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook
Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 634–638. https://doi.org/10.3758/BF03206544
Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0068126
Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75–92. https://doi.org/10.1016/j.neuropsychologia.2006.04.025
Parron, C., Da Fonseca, D., Santos, A., Moore, D. G., Monfardini, E., & Deruelle, C. (2008). Recognition of biological motion in children with autistic spectrum disorders. *Autism*, *12*(3), 261-274. https://doi.org/10.1177/1362361307089520
Paulmann, S., & Pell, M. D. (2009). Facial expression decoding as a function of emotional meaning status: ERP evidence. *NeuroReport*, *20*(18), 1603–1608. https://doi.org/10.1097/WNR.0b013e3283320e3f
Pavlova, M. A., Guerreschi, M., Tagliavento, L., Gitti, F., Sokolov, A. N., Fallgatter, A. J., & Fazzi, E. (2017). Social cognition in autism: Face tuning. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-02790-1
Pell, M. D., & Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027256
*Pelphrey, K., Goldman, B. D., Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., … Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*(4), 249–261. https://doi.org/10.1023/A
Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(2), 111–141. https://doi.org/10.1016/S0010-0277(98)00043-2
Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*(2), 331-348. https://doi.org/10.1006/nimg.2002.1087
*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. https://doi.org/10.1017/S0033291709992364
Prinz, J. (2004). Which emotions are basic? In D. Evans & P. Cruse (Eds.), *Emotion, Evolution, and Rationality* (pp. 69–87). Oxford: Oxford University Press.

Prkachin, G. C. (2003). The effects of orientation on detection and identification of facial expressions of emotion. *British journal of psychology*, *94*(1), 45-62. https://doi.org/10.1348/000712603762842093

*Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2011). Emotion perception in music in high-functioning adolescents with autism spectrum Disorders. *Journal of Autism and Developmental Disorders*, *41*(9), 1240–1255. https://doi.org/10.1007/s10803-010-1146-0

Quintin, E. M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2013). Processing of musical structure by high-functioning adolescents with autism spectrum disorders. *Child Neuropsychology*, *19*(3), 250–275. https://doi.org/10.1080/09297049.2011.653540

*Rhodes, G., Burton, N., Jeffery, L., Read, A., Taylor, L., & Ewing, L. (2018). Facial expression coding in children and adolescents with autism : Reduced adaptability but intact norm-based coding. *British Journal of Psychology*, *109*(2), 204–218. https://doi.org/10.1111/bjop.12257

*Rigby, S. N., Stoesz, B. M., & Jakobson, L. S. (2018). Empathy and face processing in adults with and without autism spectrum disorder. *Autism Research*, *11*(6), 942–955. https://doi.org/10.1002/aur.1948

Rigoulot, S., Wassiliwizky, E., & Pell, M. D. (2013). Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition. *Frontiers in Psychology*, *4*(June), 1–14. https://doi.org/10.3389/fpsyg.2013.00367

Rommelse, N., Langerak, I., Van Der Meer, J., De Bruijn, Y., Staal, W., Oerlemans, A., & Buitelaar, J. (2015). Intelligence may moderate the cognitive profile of patients with ASD. *PLoS One*, *10*(10), e0138698. https://doi.org/10.1371/journal.pone.0138698

Rosset, D. B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., & Deruelle, C. (2008). Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(5), 919–925. https://doi.org/10.1007/s10803-007-0465-2

RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/

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. https://doi.org/10.1111/j.1467-8624.2009.01343.x

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russo-Ponsaran, N. M., Mckown, C., Johnson, J. K., Allen, A. W., Evans-Smith, B., & Fogg, L. (2015). Social-emotional correlates of early stage social information processing skills in children with and without autism spectrum disorder. *Autism Research*, *8*(5), 486–496. https://doi.org/10.1002/aur.1463

Rutherford, M. D., & McIntosh, D. N. (2007). Rules versus prototype matching: Strategies of perception of emotional facial expressions in the autism spectrum. *Journal of Autism and Developmental Disorders*, *37*(2), 187–196. https://doi.org/10.1007/s10803-006-0151-9

Rutherford, M. D., & Troje, N. F. (2012). IQ predicts biological motion perception in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(4), 557–565. https://doi.org/10.1007/s10803-011-1267-0

Ryan, C., Stafford, M., & King, R. J. (2016). Brief Report: Seeing the Man in the Moon: Do Children with Autism Perceive Pareidolic Faces? A Pilot Study. *Journal of Autism and Developmental Disorders*, *46*(12), 3838–3843. https://doi.org/10.1007/s10803-016-2927-x

Salomone, E., Bulgarelli, D., Thommen, E., Rossini, E., & Molina, P. (2019). Role of age and IQ in emotion understanding in Autism Spectrum Disorder: implications for educational interventions. *European journal of special needs education*, *34*(3), 383-392. https://doi.org/10.1080/08856257.2018.1451292

*Sasson, N. J., Pinkham, A. E., Weittenhiller, L. P., Faso, D. J., & Simpson, C. (2016a). Context effects on facial affect recognition in schizophrenia and autism: Behavioral and eye-tracking evidence. *Schizophrenia Bulletin*, *42*(3), 675–683. https://doi.org/10.1093/schbul/sbv176

*Sasson, N. J., Shasteen, J. R., & Pinkham, A. E. (2016b). Brief report: Reduced prioritization of facial threat in adults with autism. *Journal of Autism and Developmental Disorders*, *46*(4), 1471–1476. https://doi.org/10.1007/s10803-015-2664-6

Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition and Emotion*, *18*(5), 701–710. https://doi.org/10.1080/02699930341000176

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(6), 2408–2412. https://doi.org/10.1073/pnas.0908239106

*Sawyer, A. C. P., Williamson, P., & Young, R. L. (2012). Can gaze avoidance explain why individuals with Asperger's syndrome can't recognise emotions from facial expressions? *Journal of Autism and Developmental Disorders*, *42*(4), 606–618. https://doi.org/10.1007/s10803-011-1283-0

*Schaller, U. M., & Rauh, R. (2017). What difference does it make? Implicit, explicit and complex social cognition in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *47*(4), 961–979. https://doi.org/10.1007/s10803-016-3008-x

*Schelinski, S., & von Kriegstein, K. (2019). The relation between vocal pitch and vocal emotion recognition abilities in people with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders*, *49*(1), 68–82. https://doi.org/10.1007/s10803-018-3681-z

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *The Journal of the Acoustical Society of America*, *76*(5), 1346–1356. https://doi.org/10.1121/1.391450

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. *Trends in Cognitive Sciences*, *21*(3), 216–228. https://doi.org/10.1016/j.tics.2017.01.001

Schünemann HJ, Vist GE, Higgins JPT, Santesso N, Deeks JJ, Glasziou P, Akl EA, Guyatt GH. Chapter 15: Interpreting results and drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook

Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental retardation and developmental disabilities research reviews*, *10*(4), 234-247. https://doi.org/10.1002/mrdd.20038

*Shafritz, K. M., Bregman, J. D., Ikuta, T., & Szeszko, P. R. (2015). Neural systems mediating decision-making and response inhibition for social and nonsocial stimuli in autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *60*, 112–120. https://doi.org/10.1016/j.pnpbp.2015.03.001

*Shanok, N. A., Jones, N. A., & Lucas, N. N. (2019). The nature of facial emotion recognition impairments in children on the autism spectrum. *Child Psychiatry and Human Development*, *50*(4), 661–667. https://doi.org/10.1007/s10578-019-00870-z

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., … Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1735–1747. https://doi.org/10.1007/s10803-006-0307-7

*Sinzig, J., Morsch, D., & Lehmkuhl, G. (2008). Do hyperactivity, impulsivity and inattention have an impact on the ability of facial affect recognition in children with autism and ADHD? *European Child and Adolescent Psychiatry*, *17*(2), 63–72. https://doi.org/10.1007/s00787-007-0637-9

Sivathasan, S., Fernandes, T. P., Burack, J. A., & Quintin, E. M. (2020). Emotion processing and autism spectrum disorder: A review of the relative contributions of alexithymia and verbal IQ. *Research in Autism Spectrum Disorders*, *77*, 101608. https://doi.org/10.1016/j.rasd.2020.101608

Song, Y., Zhong, J., Jia, Z., & Liang, D. (2020). Emotional prosody recognition in children with high-functioning autism under the influence of emotional intensity: Based on the perspective of emotional dimension theory. *Journal of Communication Disorders*, *88*(July), 106032. https://doi.org/10.1016/j.jcomdis.2020.106032

Sowden, S., Schuster, B. A., Keating, C. T., Fraser, D. S., & Cook, J. L. (2021). The role of movement kinematics in facial emotion expression production and recognition. *Emotion (Washington, D.C.)*. https://doi.org/10.1037/emo0000835

Spineli, L. M., & Pandis, N. (2020). Prediction interval in random-effects meta-analysis. *American Journal of Orthodontics and Dentofacial Orthopedics*, *157*(4), 586–588. https://doi.org/10.1016/j.ajodo.2019.12.011

*Stephenson, K. G., Luke, S. G., & South, M. (2019). Separate contributions of autistic traits and anxious apprehension, but not alexithymia, to emotion processing in faces. *Autism*, *23*(7), 1830–1842. https://doi.org/10.1177/1362361319830090

*Stewart, M. E., Mcadam, C., Ota, M., Peppé, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*, *17*(1), 6–14.

Sun, X., Allison, C., Wei, L., Matthews, F. E., Auyeung, B., Wu, Y. Y., ... & Brayne, C. (2019). Autism prevalence in China is comparable to Western prevalence. *Molecular autism*, *10*(1), 7. https://doi.org/10.1186/s13229-018-0246-0

Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., Drew, A., Charman, T., … Wheelwright, S. (1998). The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(5), 747–753. https://doi.org/10.1017/S0021963098002595

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

Tanaka, J. W., & Sung, A. (2016). The “eye avoidance” hypothesis of autism face processing. *Journal of Autism and Developmental Disorders*, *46*(5), 1538–1552. https://doi.org/10.1007/s10803-013-1976-7

*Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., … Schultz, R. T. (2012). The perception and identification of facial emotions in individuals with autism spectrum disorders using the Let’s Face It! Emotion Skills Battery. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(12), 1259–1267. https://doi.org/10.1111/j.1469-7610.2012.02571.x

Tantam, D., Monaghan, L., Nicholson, H., & Stirling, J. (1989). Autistic children’s ability to interpret faces: A research note. *Journal of Child Psychology and Psychiatry*, *30*(4), 623–630. https://doi.org/10.1111/j.1469-7610.1989.tb00274.x

*Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language and Communication Disorders*, *50*(4), 452–466. https://doi.org/10.1111/1460-6984.12146

*Tell, D., Davidson, D., & Camras, L. A. (2014). Recognition of emotion from facial expressions with direct or averted eye gaze and varying expression intensities in children with autism disorder and typically developing children. *Autism Research and Treatment*, *2014*, 1–11. https://doi.org/10.1155/2014/816137

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*(11), 1559–1573. https://doi.org/10.1002/sim.1187

Tonks, J., Williams, W. H., Frampton, I., Yates, P., & Slater, A. (2007). Assessing emotion recognition in 9-15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain Injury*, *21*(6), 623–629. https://doi.org/10.1080/02699050701426865

*Tottenham, N., Hertzig, M. E., Gillespie-Lynch, K., Gilhooly, T., Millner, A. J., & Casey, B. J. (2014). Elevated amygdala response to faces and gaze aversion in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*, *9*(1), 106–117. https://doi.org/10.1093/scan/nst050

Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-010-1030-y

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. https://doi.org/10.1016/j.rasd.2016.08.004

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

*Uono, S., Sato, W., & Toichi, M. (2011). The specific impairment of fearful expression recognition and its atypical development in pervasive developmental disorder. *Social Neuroscience*, *6*(5–6), 452–463. https://doi.org/10.1080/17470919.2011.605593

*Uono, S., Sato, W., & Toichi, M. (2013). Common and unique impairments in facial-expression recognition in pervasive developmental disorder-not otherwise specified and Asperger's disorder. *Research in Autism Spectrum Disorders*, *7*(2), 361–368. https://doi.org/10.1016/j.rasd.2012.10.007

Van Lancker, D., Cornelius, C., & Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic, and normal children. *Developmental Neuropsychology*, *5*(2–3), 207–226. https://doi.org/10.1080/87565648909540433

*Vannetzel, L., Chaby, L., Cautru, F., Cohen, D., & Plaza, M. (2011). Neutral versus emotional human stimuli processing in children with pervasive developmental disorders not otherwise specified. *Research in Autism Spectrum Disorders*, *5*(2), 775–783. https://doi.org/10.1016/j.rasd.2010.09.005

Velikonja, T., Fett, A. K., & Velthorst, E. (2019). Patterns of Nonsocial and Social Cognitive Functioning in Adults with Autism Spectrum Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *76*(2), 135–151. https://doi.org/10.1001/jamapsychiatry.2018.3645

Vesterinen, H. M., Sena, E. S., Egan, K. J., Hirst, T. C., Churolov, L., Currie, G. L., … Macleod, M. R. (2014). Meta-analysis of data from animal studies: A practical guide. *Journal of Neuroscience Methods*, *221*, 92–102. https://doi.org/10.1016/j.jneumeth.2013.09.010

Vidas, D., Dingle, G. A., & Nelson, N. L. (2018). Children's recognition of emotion in music and speech. *Music & Science*, *1*, 205920431876265. https://doi.org/10.1177/2059204318762650
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03
Vogan, V. M., Leung, R. C., Safar, K., Martinussen, R., Smith, M. L., & Taylor, M. J. (2018). Longitudinal Examination of Everyday Executive Functioning in Children With ASD: Relations With Social, Emotional, and Behavioral Functioning Over Time. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01774
*Waddington, F., Hartman, C., de Bruijn, Y., Lappenschaar, M., Oerlemans, A., Buitelaar, J., … Rommelse, N. (2018). Visual and auditory emotion recognition problems as familial cross-disorder phenomenon in ASD and ADHD. *European Neuropsychopharmacology*, *28*(9), 994–1005. https://doi.org/10.1016/j.euroneuro.2018.06.009
*Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition and Emotion*, *22*(7), 1353–1380. https://doi.org/10.1080/02699930701782153
*Wang, J. E., & Tsao, F. M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, *37*(1), 162–170. https://doi.org/10.1016/j.ridd.2014.11.013
*Wang, S., & Adolphs, R. (2017). Reduced specificity in emotion judgment in people with autism spectrum disorder. *Neuropsychologia*, *99*(February), 286–295. https://doi.org/10.1016/j.neuropsychologia.2017.03.024
Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face Recognition Deficits in Autism Spectrum Disorders Are Both Domain Specific and Process Specific. *PLOS ONE*, *8*(9), e74541. https://doi.org/10.1371/journal.pone.0074541
*Wingenbach, T. S. H., Ashwin, C., & Brosnan, M. (2017). Diminished sensitivity and specificity at recognising facial emotional expressions of varying intensity underlie emotion-specific recognition deficits in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *34*, 52–61. https://doi.org/10.1016/j.rasd.2016.11.003
*Wong, N., Beidel, D. C., Sarver, D. E., & Sims, V. (2012). Facial emotion recognition in children with high functioning autism and children with social phobia. *Child Psychiatry Hum Dev*, *43*(5), 775–794. https://doi.org/10.1007/s10578-012-0296-z
World Health Organization. (2018). *International classification of diseases for mortality and morbidity statistics* (11th Revision). Retrieved from https://icd.who.int/browse11/l-m/en
*Wright, B., Clarke, N., Jordan, J., Young, A. W., Clarke, P., Miles, J., … Williams, C. (2008). Emotion recognition in faces and the use of visual context Vo in young people with high-functioning autism spectrum disorders. *Autism*, *12*(6), 607–626. https://doi.org/10.1177/1362361308097118
* Yeung, M. K., Han, Y. M. Y., Sze, S. L., & Chan, A. S. (2014). Altered right frontal cortical connectivity during facial emotion recognition in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *8*(11), 1567–1577. https://doi.org/10.1016/j.rasd.2014.08.013

Figure captions

Figure 1. Study selection process. PRISMA diagram of the combined initial and update literature search and screening process.

Figure 2. Risk of bias graph. Reviewers' judgements about each risk of bias item presented as percentages across the 72 included studies using the CASP (2017) checklist. A score was assigned for each criterion on the checklist using a three-point rating system developed by Duggleby et al. (2010). A score of 1 point denotes a high risk of bias and is given to papers that provide little to no justification for a particular issue; 2 points, a moderate risk of bias where the issue was addressed somewhat but not fully elaborated; 3 points, a low risk of bias with issues concerned being extensively justified and explained. The CASP scores were not used as a means of excluding papers but to provide indications of the quality across the included studies, as well as on the collective research evidence.

*Figure 3***.** Forest plot for six-emotion composite accuracy results. Twenty-two datasets contributed to this. The pooled SMDs indicated that the ASD groups performed worse than the TD groups, representing a large effect of -0.77 (95% CI [-1.03, -0.50]). The I^2 shows that there is significant heterogeneity between studies ($I^2 = 77.56\%$).

Figure 4. Bar graph with means and 95% confidence intervals error bars for the standardised mean difference (SMD) for each of the seven emotion categories.

Figure 5. Meta-regression scatter plot showing the SMD of the individual studies plotted against the mean age of the respective ASD group for the composite. Each point represents a study. The radius of the points is drawn proportional to the inverse of the standard errors (i.e., the precision of the effect estimates with larger/more precise studies shown as larger points). The predicted average standardised mean difference as a function of age is shown as the fitted regression line with corresponding 95% confidence interval bounds.

Figure 6. Funnel plot of studies included in the meta-analysis for the composite. Black dots indicate observed studies and white dots indicate imputed studies correcting for funnel plot asymmetry.

Figure 1

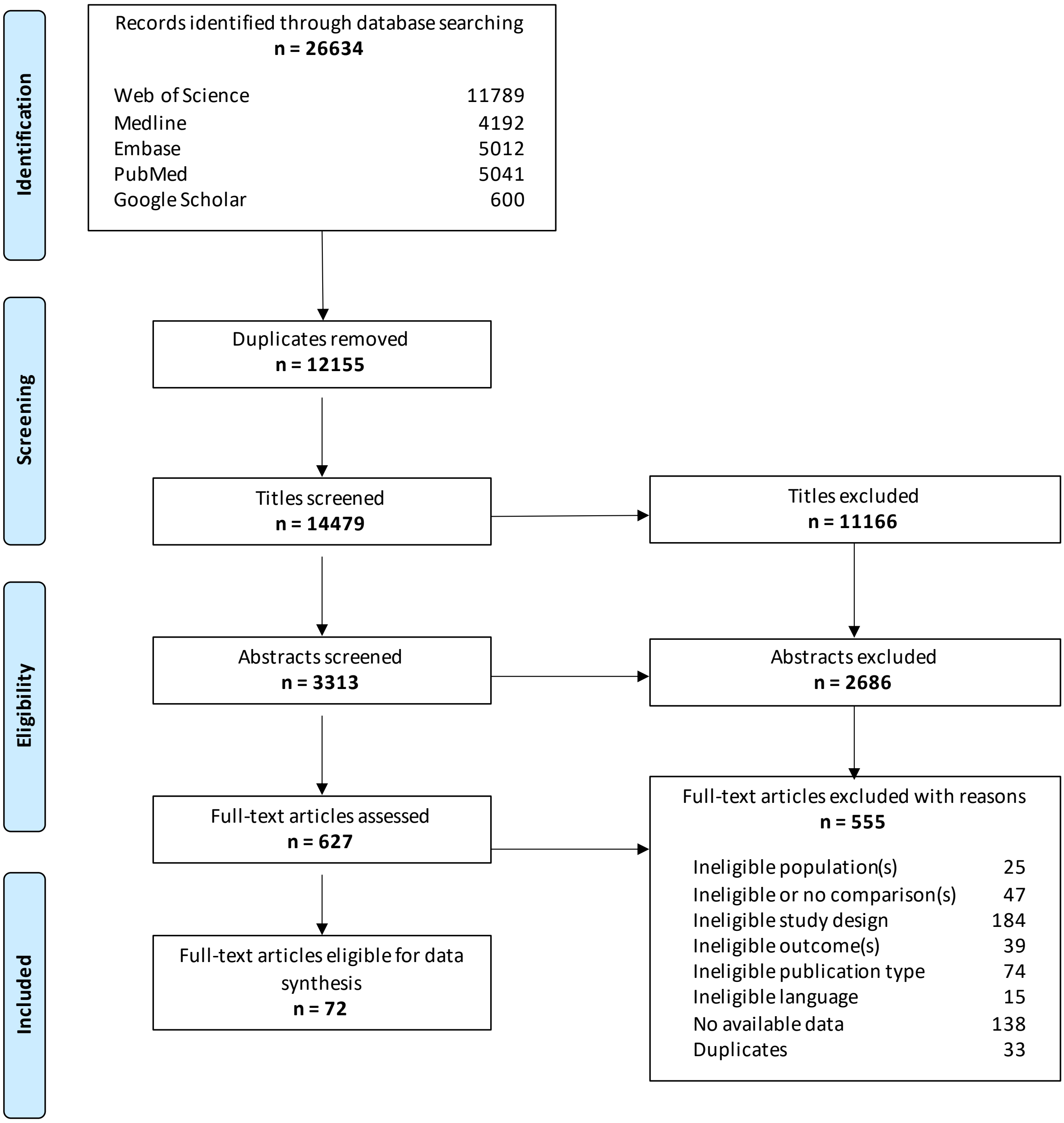
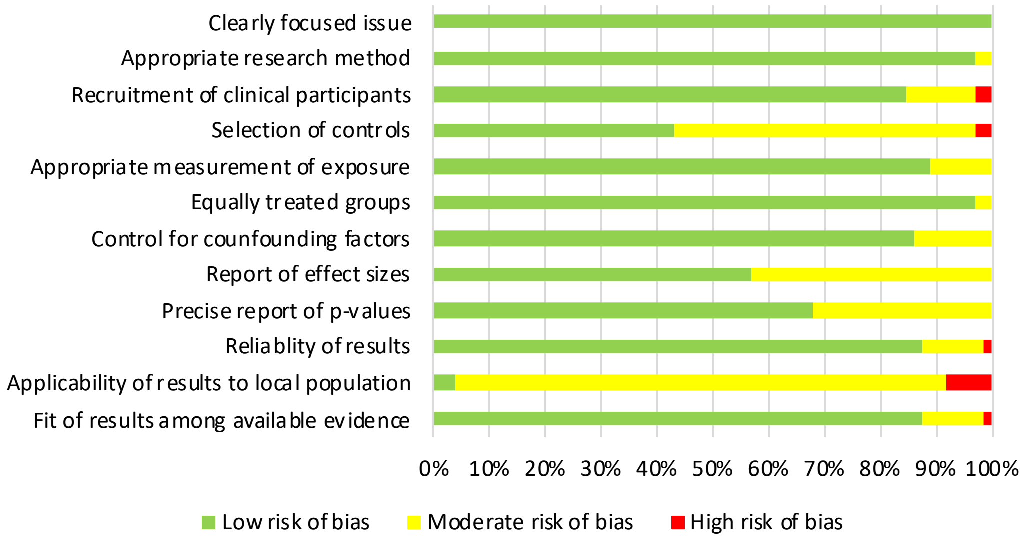
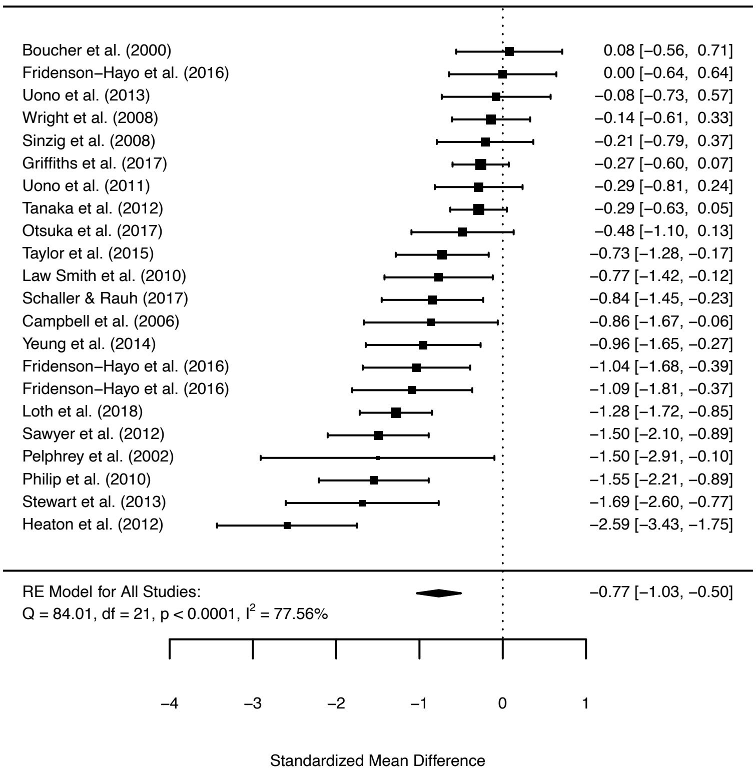

Identification
Screening
Eligibility
Included
Records identified through database searching
n = 26634
Web of Science 11789
Medline 4192
Embase 5012
PubMed 5041
Google Scholar 600
Duplicates removed
n = 12155
Titles screened
n = 14479
Titles excluded
n = 11166
Abstracts screened
n = 3313
Abstracts excluded
n = 2686
Full-text articles assessed
n = 627
Full-text articles excluded with reasons
n = 555
Ineligible population(s) 25
Ineligible or no comparison(s) 47
Ineligible study design 184
Ineligible outcome(s) 39
Ineligible publication type 74
Ineligible language 15
No available data 138
Duplicates 33
Full-text articles eligible for data synthesis
n = 72

Figure 2



Clearly focused issue
Appropriate research method
Recruitment of clinical participants
Selection of controls
Appropriate measurement of exposure
Equally treated groups
Control for counfounding factors
Report of effect sizes
Precise report of p-values
Reliablity of results
Applicability of results to local population
Fit of results among available evidence
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Low risk of bias
Moderate risk of bias
High risk of bias

Figure 3



Boucher et al. (2000) 0.08 [−0.56, 0.71]
Fridenson−Hayo et al. (2016) 0.00 [−0.64, 0.64]
Uono et al. (2013) −0.08 [−0.73, 0.57]
Wright et al. (2008) −0.14 [−0.61, 0.33]
Sinzig et al. (2008) −0.21 [−0.79, 0.37]
Griffiths et al. (2017) −0.27 [−0.60, 0.07]
Uono et al. (2011) −0.29 [−0.81, 0.24]
Tanaka et al. (2012) −0.29 [−0.63, 0.05]
Otsuka et al. (2017) −0.48 [−1.10, 0.13]
Taylor et al. (2015) −0.73 [−1.28, −0.17]
Law Smith et al. (2010) −0.77 [−1.42, −0.12]
Schaller & Rauh (2017) −0.84 [−1.45, −0.23]
Campbell et al. (2006) −0.86 [−1.67, −0.06]
Yeung et al. (2014) −0.96 [−1.65, −0.27]
Fridenson−Hayo et al. (2016) −1.04 [−1.68, −0.39]
Fridenson−Hayo et al. (2016) −1.09 [−1.81, −0.37]
Loth et al. (2018) −1.28 [−1.72, −0.85]
Sawyer et al. (2012) −1.50 [−2.10, −0.89]
Pelphrey et al. (2002) −1.50 [−2.91, −0.10]
Philip et al. (2010) −1.55 [−2.21, −0.89]
Stewart et al. (2013) −1.69 [−2.60, −0.77]
Heaton et al. (2012) −2.59 [−3.43, −1.75]
RE Model for All Studies: −0.77 [−1.03, −0.50]
Q = 84.01, df = 21, p < 0.0001, I2 = 77.56%
−4 −3 −2 −1 0 1
Standardized Mean Difference

Figure 4

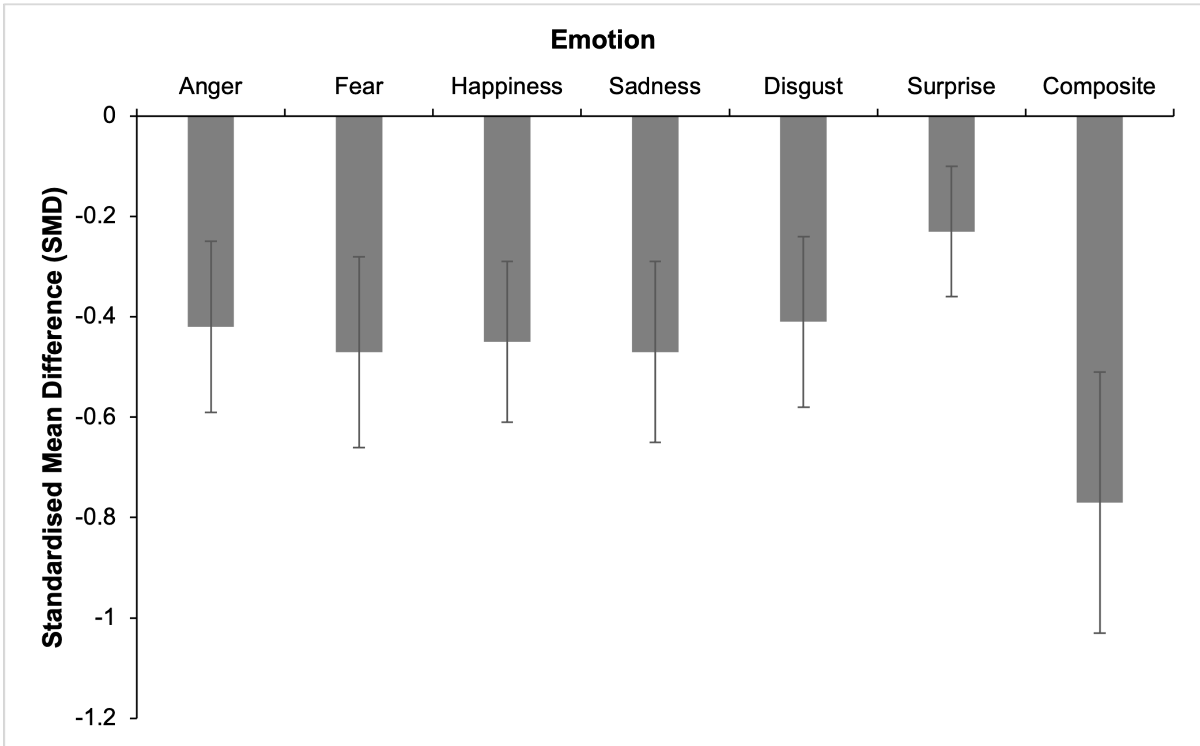
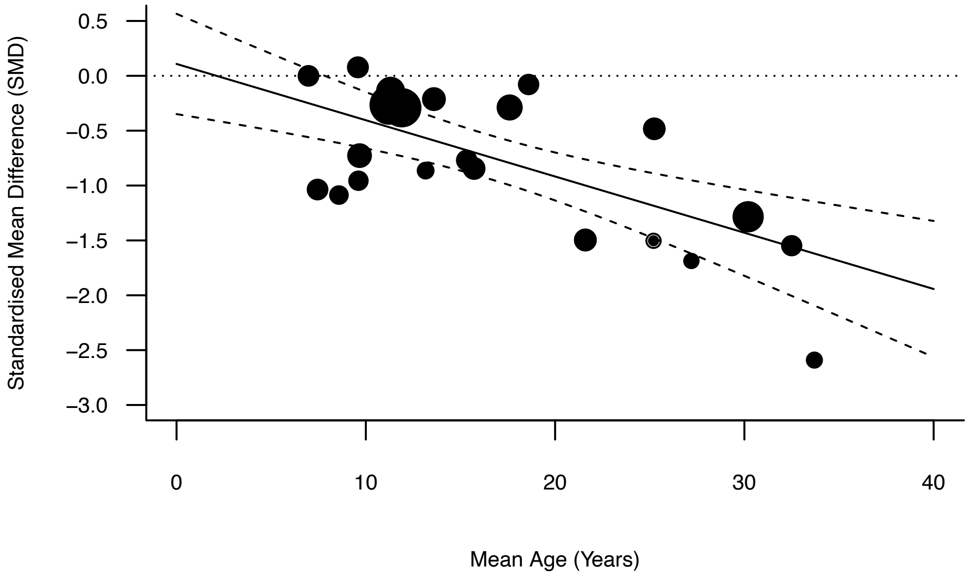
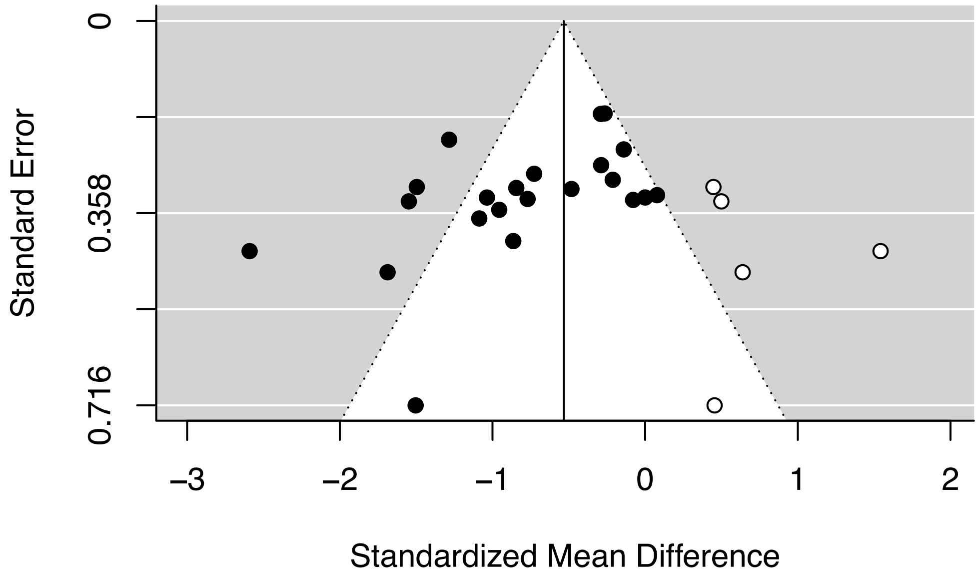

Emotion
Anger
Fear
Happiness
Sadness
Disgust
Surprise
Composite
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
Standardised Mean Difference (SMD)

Figure 5



0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
Standardised Mean Difference (SMD)
0
10
20
30
40
Mean Age (Years)

Figure 6



0
0.358
0.716
Standard Error
−3
−2
−1
0
1
2
Standardized Mean Difference

Tables

Table 1.

*Summary of included studies on emotion recognition in ASD. * indicates datasets included in the main analyses. - indicates information that was not available. X indicates unused datasets due to prioritisation and non-positive sampling variances.*

Row	Study	Number (males)	Diagnostic instrument	Mean age in years (SD/range)	IQ mean (SD/range)	Number (males)	Matching procedure	Mean age in years (SD range)	IQ mean (SD range)	Domain (Subdomain)	Type	Stimuli	Emotions	Mean accuracy ASD (SD)	Mean accuracy control (SD)
1	Akechi et al. (2009)	14 (10)	ASQ-J	12.10 (2.00)	FSIQ: 98.90 (16.10) VIQ: 10.50 (3.10) NVIQ: 9.10 (3.00)	14 (10)	Age Gender FSIQ VIQ NVIQ	11.90 (1.90)	FSIQ: 101.30 (12.80) VIQ: 11.40 (2.40) NVIQ: 9.10 (2.90)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful	-	-
2*	Akechi et al (2010)	14 (10)	ASQ-J	13.70 (2.30)	FSIQ: 96.80 (16.80) VIQ: 9.40 (3.40) NVIQ: 9.60 (2.70)	14 (8)	Age Gender FSIQ VIQ NVIQ	12.30 (2.10)	FSIQ: 104.10 (8.10) VIQ: 11.10 (2.90) NVIQ: 10.20 (2.30)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful	0.88 (0.16) 0.81 (0.14)	0.95 (0.09) 0.85 (0.11)
3*	Baker (2010)	19 (13)	ADI-R; ADOS; ASDS	12.80 (1.47)	FSIQ: - (85-115)	19 (13)	Age Gender	12.20 (1.43)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Nonsense passages with prosodic patterns implying different emotions	Angry Happy Sad	0.70 (0.22) 0.78 (0.20) 0.73 (0.24)	0.67 (0.21) 0.84 (0.15) 0.75 (0.14)
4*	Bast et al. (2019)	23 (18)	ADI-R; ADOS	15.90 (2.80)	FSIQ: 100.00 (16.10) VIQ: 102.40 (16.60) NVIQ: 02.70 (18.40)	24 (19)	Age Gender FSIQ VIQ NVIQ	15.80 (2.40)	FSIQ: 108/60 (14.40) VIQ: 103.30 (13.90) NVIQ: 110.80 (13.80)	Human faces (dynamic)	Verbal (forced-choice labelling)	Videos from the Movie for the Assessment of Social Cognition (Dziobek et al., 2006)	Angry Fearful Happy	0.39 (0.49) 0.59 (0.50) 0.50 (0.51)	0.66 (0.48) 0.92 (0.28) 0.67 (0.48)
5*	Berggren et al. (2016)	35 (18)	ADOS; ICD-10	11.60 (1.80)	FSIQ: 103.80 (11.90)	32 (18)	Age Gender FSIQ	11.70 (1.80)	FSIQ: 102.90 (8.50)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Swedish version; Bölte et al. 2002)	Angry Fearful Happy Sad Disgust Surprise	3.80 (1.80) 3.70 (1.80) 8.20 (1.40) 5.40 (1.60) 4.20 (1.50) 4.70 (1.80)	4.60 (1.90) 4.90 (2.20) 9.80 (1.30) 5.30 (1.30) 4.60 (1.00) 5.40 (1.80)
6*	Boggs et al. (2010)	17 (15)	ADI-R; DSM-IV-TR	15.47 (2.38)	FSIQ: 105.94 (10.12) VIQ: 105.35 99.71)	17 (15)	Age Gender FSIQ VIQ	15.12 (2.15)	FSIQ: 108.41 (10.06) VIQ: 108.59 (9.64)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of four females expressing the different emotions	Angry Fearful Happy Sad Surprise	0.90 (0.18) 0.49 (0.23) 0.99 (0.06) 0.94 (0.14) 0.91 (0.15)	0.96 (0.13) 0.54 (0.24) 1.00 (0) 0.96 (0.1) 0.91 (0.15)
7*	Boucher et al. (2000)	19 (16)	DSM-IV	9.58 (1.00)	-	19 (10)	Age Gender FSIQ	6.33 (0.80)	-	Speech prosody (excerpt)	Verbal (free-choice labelling)	Audio tapes of a woman reciting the days of the week or months of the year in each of the emotions	Overall	13.50 (7.79)	13.05 (1.84)

8a*	Brennand et al. (2011)	15 (14)	-	14.50 (2.70)	VIQ: 92.50 (21.70)	15 (12)	Age	13.30 (1.67)	VIQ: 117.60 (13.60)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Psuedo sentences consisting of phonemes and phonotactics common in European languages that elicited different emotions by German actors	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad	55.00 (19.36) 40.80 (19.36) 30.80 (15.88) 47.50 (17.82)	58.30 (19.36) 53.30 (19.36) 45.00 (15.88) 50.00 (17.82)
8b	Brennand et al. (2011)	15 (14)	-	14.50 (2.70)	VIQ: 92.50 (21.70)	32 (28)	Not matched	-	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Psuedo sentences consisting of phonemes and phonotactics common in European languages that elicited different emotions by German actors	Angry Fearful Happy Sad	55.00 (19.36) 40.80 (19.36) 30.80 (15.88) 47.50 (17.82)	65.20 (19.80) 48.00 (19.23) 48.00 (15.84) 65.20 (19.80)
9a*	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the standard condition of the production task (i.e. free production)	Angry Fearful Happy Sad Disgust Surprise	0.31 (0.18) 0.24 (0.16) 0.9 (0.12) 0.49 (0.14) 0.47 (0.14) 0.41 (0.15)	0.57 (0.25) 0.28 (0.18) 0.88 (0.17) 0.44 (0.1) 0.53 (0.21) 0.44 (0.16)
9b	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the communicate condition of the production task (i.e. experimenter guessed the posers' expressed emotions)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad ^{X} Disgust ^{X} Surprise	0.38 (0.24) 0.29 (0.13) 0.84 (0.13) 0.54 (0.21) 0.64 (0.14) 0.54 (0.14)	0.61 (0.22) 0.42 (0.25) 0.82 (0.2) 0.46 (0.18) 0.55 (0.15) 0.52 (0.20)
9c	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the mirror condition of the production task (i.e. production while watching own expressions in a camera)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad ^{X} Disgust ^{X} Surprise	0.34 (0.19) 0.34 (0.21) 0.93 (0.11) 0.53 (0.19) 0.41 (0.11) 0.62 (0.19)	0.64 (0.27) 0.45 (0.26) 0.85 (0.18) 0.53 (0.19) 0.50 (0.24) 0.55 (0.21)
10	Campbell et al. (2006)	13 (11)	ICD-10	13.16 (1.75)	VIQ: 96.07 (17.86)	13 (11)	Age Gender VIQ	13.32 (2.08)	VIQ: 95.92 (16.41)	Human faces (static)	Verbal (forced-choice labelling)	The Ekman-Friesen Test of Affect Recognition	Overall	71.39 (10.2)	80.69 (10.65)
11*	Ciaramidaro et al. (2018)	33 (31)	ADI-R; ADOS; ICD-10	18.76 (4.98)	NVIQ: 105.82 (13.75)	25 (21)	Gender VIQ	19.86 (3.45)	FSIQ: 109.00 (12.55) NVIQ: 109.00 (12.55)	Human faces (static)	Verbal (discrimination)	Photographs from the Karolinska Directed Emotional Faces	Angry Fearful	26.60 (3.18) -	28.76 (2.33) -
12*	Corden et al. (2008)	18 (13)	ADOS	32.90 (13.35)	FSIQ: 119.90 (11.10) VIQ: 116.30 (9.14) NVIQ: 117.10 (14.56)	17 (13)	Age Gender FSIQ VIQ NVIQ	31.90 (11.30)	FSIQ: 117.40 (8.26) VIQ: 115.10 (8.37) NVIQ: 115.90 (8.87)	Human faces (static)	Verbal (forced-choice labelling)	Halftone images of emotionally expressive faces	Angry Fearful Happy Sad Disgust Surprise	8.50 (1.34) 7.50 (1.41) 9.90 (0.24) 7.70 (1.19) 6.60 (2.57) 8.40 (2.20)	8.70 (1.10) 8.80 (1.33) 10.00 (0) 8.20 (1.25) 8.10 (2.26) 8.80 (1.29)

13*	Couture et al. (2010)	36 (29)	ADI-R	20.90 (5.70)	FSIQ: 101.20 (17.80)	41 (34)	Age Gender FSIQ	22.90 (5.60)	FSIQ: 109.40 (15.10)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of movie stills (Adolphs & Tranel, 2003)	Angry Fearful Happy Sad Disgust Surprise	0.52 (0.24) 0.62 (0.19) - 0.69 (0.27) - -	0.67 (0.19) 0.62 (0.17) - 0.80 (0.19) - -
14a*	Davidson et al. (2019)	23 (19)	-	11.08 (1.75)	FSIQ: 97.00 (16.30) VIQ: 46.00 (13.30) NVIQ: 49.00 (13.10)	23 (19)	Age Gender NVIQ	11.50 (2.08)	FSIQ: 107.50 (10.50) VIQ: 58.00 (8.00) NVIQ: 50.10 (6.20)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	0.91 (0.28) 0.54 (0.49) 0.63 (0.49) 0.60 (0.37)	0.94 (0.25) 0.52 (0.49) 0.76 (0.43) 0.90 (0.38)
14b	Davidson et al. (2019)	23 (19)	-	11.08 (1.75)	FSIQ: 97.00 (16.30) VIQ: 46.00 (13.30) NVIQ: 49.00 (13.10)	23 (19)	Age Gender NVIQ	11.50 (2.08)	FSIQ: 107.50 (10.50) VIQ: 58.00 (8.00) NVIQ: 50.10 (6.20)	Non-human faces (static)	Verbal (forced-choice labelling)	Canine faces from the internet	Angry Fearful Happy Sad	0.80 (0.24) 0.68 (0.29) 0.89 (0.29) 0.82 (0.32)	0.83 (0.24) 0.70 (0.39) 0.89 (0.21) 0.78 (0.22)
15*	Doi et al. (2013)	20 (20)	AQ-J; DSM-IV	32.10 (7.30)	FSIQ: 104.20 (15.30) VIQ: 109.90 (15.50) NVIQ: 92.70 (19.90)	20 (20)	Age Gender FSIQ VIQ	33.50 (4.70)	FSIQ: 107.20 (13.90) VIQ: 104.10 (13.80) NVIQ: 109.90 (12.40)	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs at from the ATR DB99 and ATR-Promotions database	Angry Happy Sad	80.80 (17.90) 87.90 (15.70) 84.40 (19.60)	91.20 (7.20) 91.30 (5.70) 90.50 (7.50)
15b	Doi et al. (2013)	20 (20)	AQ-J; DSM-IV	32.10 (7.30)	FSIQ: 104.20 (15.30) VIQ: 109.90 (15.50) NVIQ: 92.70 (19.90)	20 (20)	Age Gender FSIQ VIQ	33.50 (4.70)	FSIQ: 107.20 (13.90) VIQ: 104.10 (13.80) NVIQ: 109.90 (12.40)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Spoken words uttering a common family name in Japan in different emotions	Angry Happy Sad	90.80 (16.60) 53.30 (30.90) 87.50 (14.20)	88.80 (18.90) 85.00 (17.00) 92.50 (11.80)
16*	Eack et al. (2015)	45 (40)	ADOS	24.64 (5.72)	FSIQ: 112.60 (15.74)	30 (22)	Age Gender	26.40 (5.80)	FSIQ: 105.53 (7.01)	Human faces (static)	Verbal (forced-choice labelling)	Penn Emotion Recognition Test – 40 (Kohler et al. 2003)	Angry Fearful Happy Sad	4.90 (1.55) 6.59 (1.64) 7.37 (0.68) 6.04 (1.70)	5.18 (1.3) 7.22 (1.05) 7.91 (0.31) 6.77 (1.14)
17*	Evers et al. (2014)	22 (22)	ADI-R; DSM-IV-TR	7.85 (0.88)	FSIQ: 94.36 (11.93)	22(22)	Age Gender FSIQ	7.95 (0.68)	FSIQ: 98.50 (7.78)	Human Faces (static)	Verbal (forced-choice labelling)	California Facial Expressions Database	Angry Fearful Happy Sad	0.71 (0.22) 0.76 (0.21) 0.92 (0.11) 0.52 (0.23)	0.67 (0.26) 0.75 (0.21) 0.98 (0.04) 0.56 (0.26)
18*	Fink et al. (2014)	114 (76)	DSM-IV	10.65 (1.23)	VIQ: 103.58 (14.440	145 (94)	Age Gender	10.32 (1.32)	VIQ: 110.56 (15.78)	Human Faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Angry Fearful Happy Sad	3.43 (0.86) 2.79 (0.91) 3.71 (0.78) 2.40 (1.04)	3.52 (0.78) 2.94 (0.68) 3.76 (0.52) 2.69 (1.00)
19a*	Fridenson-Hayo et al. (2016)	20 (18)	ADOS-2; DSM-IV-TR; ICD-10	7.45 (1.31)	VIQ: 11.15 (4.26) NVIQ: 12.50 (2.96)	22 (19)	Age Gender VIQ NVIQ	7.50 (1.47)	VIQ: 11.82 (2.99) NVIQ: 11.55 (2.30)	Human faces (Dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading	Overall	0.70 (0.18)	0.86 (0.12)
19b *	Fridenson-Hayo et al. (2016)	16 (15)	ADOS-2; DSM-IV-TR; ICD-10	8.58 (1.03)	VIQ: 11.38 (3.56) NVIQ: 11.44 (2.48)	18 (13)	Age Gender VIQ NVIQ	7.80 (1.42)	VIQ: 12.22 (2.71) NVIQ: 9.72 (3.12)	Human faces (Dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading emotions obtained from the Mindreading database	Overall	0.69 (0.14)	0.84 (0.13)
19c*	Fridenson-Hayo et al. (2016)	19 (15)	ADOS-2; DSM-IV-TR; ICD-10	6.97 (1.03)	VIQ: 9.05 (1.90) NVIQ: 11.00 (2.79)	18 (15)	Age Gender VIQ NVIQ	7.36 (1.20)	VIQ: 10.11 (1.76) NVIQ: 11.83 (2.70)	Human faces (dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading	Overall	0.74 (0.13)	0.74 (0.12)
19d	Fridenson-Hayo et al. (2016)	20 (18)	ADOS-2; DSM-IV-TR; ICD-10	7.45 (1.31)	VIQ: 11.15 (4.26) NVIQ: 12.50 (2.96)	22 (19)	Age Gender VIQ NVIQ	7.50 (1.47)	VIQ: 11.82 (2.99) NVIQ: 11.55 (2.30)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.68 (0.22)	0.73 (0.13)
19e	Fridenson-Hayo et al. (2016)	16 (15)	ADOS-2; DSM-IV-TR; ICD-10	8.58 (1.03)	VIQ: 11.38 (3.56) NVIQ: 11.44 (2.48)	18 (13)	Age Gender VIQ NVIQ	7.80 (1.42)	VIQ: 12.22 (2.71) NVIQ: 9.72 (3.12)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.64 (0.14)	0.74 (0.16)

19f	Fridenson-Hayo et al. (2016)	19 (15)	ADOS-2; DSM-IV-TR; ICD-10;	6.97 (1.03)	VIQ: 9.05 (1.90) NVIQ: 11.00 (2.79)	18 (15)	Age Gender VIQ NVIQ	7.36 (1.21)	VIQ: 10.11 (1.76) NVIQ: 11.83 (2.70)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.69 (0.13)	0.69 (0.13)
20a*	Griemel et al. (2014)	38 (38)	ADI-R; ADOS; DSM-IV; ICD-10	21.10 (9.50)	FSIQ: 107.70 (13.20)	37 (37)	Age Gender FSIQ	20.60 (7.00)	FSIQ: 113.00 (10.20)	Human faces (static)	Nonverbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychological Task battery (De Sonneville. 2001)	Angry Fearful Happy Sad	5.58 (4.50) 4.00 (3.04) 1.53 (1.96) 8.21 (5.52)	2.46 (2.27) 1.89 (2.11) 1.00 (1.08) 4.95 (3.53)
20b	Griemel et al. (2013)	38 (38)	ADI-R; ADOS; DSM-IV; ICD-10	21.10 (9.50)	FSIQ: 107.70 (13.20)	18(18)	Gender FSIQ	10.5 (1.3)	FSIQ: 111.7 (15.6)	Human faces (static)	Nonverbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychological Task battery (De Sonneville. 2001)	[X] Angry [X] Fearful [X] Happy [X] Sad	5.58 (4.50) 4.00 (3.04) 1.53 (1.96) 8.21 (5.52)	6.28 (3.86) 4.39 (2.62) 1.11 (1.08) 9.72 (4.88)
21*	Griffiths et al. (2019)	66 (58)	-	11.24 (2.91)	VIQ: 127.14 (24.26) NVIQ: 37.08 (11.61)	70 (35)	Age VIQ NVIQ	11.24 (2.49)	VIQ: 135.17 (24.99) NVIQ: 40.16 (9.72)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Disgust Surprise Overall	0.60 (0.29) 0.46 (0.31) 0.81 (0.24) 0.70 (0.27) 0.49 (0.32) 0.58 (0.28) 0.61 (0.31)	0.7 (0.25) 0.44 (0.32) 0.87 (0.18) 0.8 (0.23) 0.68 (0.26) 0.63 (0.26) 0.69 (0.29)
22*	Grossman et al. (2000)	13 (13)	ICD-10	11.80 (3.27)	FSIQ: 106.40 (18.42) VIQ: 115.80 (15.63) NVIQ: 97.20 (21.59)	13 (12)	Age FSIQ VIQ	11.50 (1.90)	FSIQ: 116.20 (12.75) VIQ: 115.00 (11.92) NVIQ: 111.80 (14.22)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Surprise	0.75 (0.29) 0.48 (0.31) 0.96 (0.09) 0.81 (0.23) 0.75 (0.31)	0.89 (0.17) 0.50 (0.23) 1.00 (0) 0.87 (0.17) 0.89 (0.17)
23*	He et al. (2019)	21 (17)	DSM-5	5.09 (0.95)	-	21 (17)	Age Gender	4.94 (0.90)	-	Human faces (dynamic)	Verbal (forced-choice labelling)	Short film scenes from CASIA Chinese Natural Emotional Audio-Visual Database	Happy Sad	t = -2.52 t = -3.44	p = .021 p = .003
24*	Heaton et al. (2012)	20 (15)	AQ; DSM	33.70 (12.77)	FSIQ: 109.10 (18.43) VIQ: 106.40 (17.45) NVIQ: 109.50 (17.95)	20 (15)	Age Gender FSIQ VIQ NVIQ	33.60 (12.06)	FSIQ: 109.50 (15.11) VIQ: 109.00 (12.84) NVIQ: 105.15 (12.89)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Vocal recordings of four actors expressing the different emotions verbally	Overall	60.42 (9.78)	85.16 (8.92)
25*	Heikkinen et al (2010)	12 (9)	ADI-R; ADOS-G; ICD-10	14.50 (-)	VIQ: 107.00 (-) NVIQ: 105.00 (-)	15 (8)	Age	14.30 (-)	-	Speech prosody (excerpt)	Verbal (forced-choice labelling)	Emotional speech data in Finnish from the MediaTeam, University of Oulu's emotional speech corpus database	Angry Happy Sad	13.08 (2.28) 16.58 (1.68) 15.17 (2.21)	13.67 (2.47) 17.07 (1.28) 16.40 (1.12)
26*	Hubbard et al. (2017)	22 (20)	ADOS; DSM-IV; DSM-5	25.91 (5.34)	FSIQ: 111.32 (11.20)	30 (10)	Not matched	22.53 (7.37)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Phrases portraying each of the emotions recorded by TD talkers during the production task	Angry Happy Sad	0.50 (0.16) 0.29 (0.15) 0.37 (0.20)	0.49 (0.18) 0.42 (0.14) 0.39 (0.20)
27a*	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 2 distractors drawn in black against a white background	Angry Happy	97.50 (6.10) 95.90 (7.50)	99.30 (2.40) 98.00 (3.80)

27b	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 5 distractors drawn in black against a white background	X Angry X Happy	98.30 (4.40) 97.70 (4.00)	99.60 (1.70) 98.00 (6.70)
27c	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 11 distractors drawn in black against a white background	X Angry X Happy	99.60 (1.90) 99.20 (2.60)	99.60 (1.70) 98.20 (4.30)
28*	Isomura et al. (2014)	10 (8)	AQ-J; DSM-IV; ICD-10	10.47 (1.10)	FSIQ: 103.40 (13.85)	14 (11)	Age FSIQ	10.09 (1.30)	103.30 (9.28)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures drawn in black against a white background	X Angry X Happy	0.71 (0.05) 0.66 (0.05)	0.69 (0.04) 0.58 (0.05)
29*	Jaervinen et al. (2016)	17 (13)	ADI-R; ADOS	10.60 (-)	-	20 (8)	Age NVIQ	10.70 (-)	-	Music (segment)	Verbal (forced-choice labelling)	Novel musical pieces eliciting each of the emotions	Fearful Happy Sad	77.21 (27.68) 94.85 (9.94) 66.18 927.87)	91.88 (10.94) 99.38 (2.8) 85 (14.96)
30*	Ketelaars et al. (2016)	31 (0)	DSM-IV-TR	41.35 (11.22)	FSIQ: 105.80 (15.44)	28 (0)	Age Gender	39.89 (13.20)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Speech samples of semantically neutral sentences from the Amsterdam Neuropsychological Task battery	Angry Fearful Happy Sad	91.40 (9.97) 50.00 (23.27) 83.33 (12.91) 87.10 (15.64)	93.45 (9.97) 47.32 (22.8) 88.39 (10.72) 88.1 (15.62)
31*	Kim et al (2015)	19 (13)	ASSQ; SCQ; SRS	11.10 (2.50)	FSIQ: 110.60 (15.30) VIQ: 114.40 (16.30) NVIQ: 107.20 (15.80)	23 (16)	Age Gender FSIQ VIQ NVIQ	11.50 (2.30)	FSIQ: 115.20 (10.30) VIQ: 117.60 (10.30) NVIQ: 110.60 (13.00)	Human faces (dynamic)	Verbal (forced-choice labelling)	Avatar recordings eliciting each of the emotions	Angry Fearful Happy Sad Disgust Surprise	62.40 (10.9) 52.00 (15.00) 50.70 (7.10) 56.00 (9.30) 60.40 (13.50) 54.40 (6.10)	56.00 (18.80) 52.20 (23.40) 83.20 (21.50) 57.60 (18.00) 58.20 (22.50) 57.10 (24.10)
32*	Kliemann et al. (2012)	16 (16)	ADI-R; AQ; ASDI; DSM-IV	30.44 (6.34)	VIQ: 108.06 (7.38) NVIQ: 128.47 (10.82)	17 (17)	Age Gender FSIQ VIQ NVIQ	30.47 (6.23)	VIQ: 108.12 (14.76) NVIQ: 126.40 (8.94)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Fearful Happy	0.92 (0.07) 0.95 90.05)	0.97 (0.03) 0.98 (0.01)
33*	Kliemann et al. (2010)	17 (12)	ADI; ADOS; AQ	32.70 (8.20)	VIQ: 104.50 (15.60)	19 (14)	Age Gender VIQ	30.40 (5.90)	VIQ: 110.40 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Fearful Happy	9039 (10.75) 94.37 (5.07)	96.44 (4.11) 96.58 (3.56)
34*	Król & Król (2019)	21 (19)	ADOS; ICD-10	16.27 (4.84)	FSIQ: 109.43 (17.67) VIQ: 110.14 (18.26) NVIQ: 107.24 (18.58)	23 (18)	Age Gender FSIQ VIQ NVIQ	16.31 (2.69)	FSIQ: 112.30 (10.59) VIQ: 10.00 (11.77) NVIQ: 113.17 (11.77)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from FACES database (Ebner et al., 2010)	Angry Fearful Happy Sad Disgust	0.86 (0.23) 0.88 (0.32) 0.95 (0.16) 0.71 (0.34) 0.83 (0.34)	0.81 (0.33) 0.91 (0.25) 1.00 (0.00) 0.82 (0.29) 1.00 (0.00)
35a*	Kujala et al. (2005)	8 (4)	DSM-IV; ICD-10	33.00 (22-43)	FSIQ: 114.00 (99-140)	8 (4)	Age Gender	32 (-)	-	Speech Prosody (utterance)	Verbal (forced-choice labelling)	Finnish word ('Saara') uttered by a female speaker with different emotional connotations	Sad	25.00 (38.00)	63.00 (42.00)
35b	Kujala et al. (2005)	8 (4)	DSM-IV; ICD-10	33.00 (22-43)	FSIQ: 114.00 (99-140)	8 (4)	Age Gender	32 (-)	-	Speech Prosody (utterance)	Nonverbal (detection)	Finnish word ('Saara') uttered by a female speaker with different emotional connotations	Sad	93.00 (14.00)	95.00 (8.00)

36*	Law Smith et al. (2010)	21 (21)	ADI-R; ADOS; DSM-IV	15.33 (2.20)	NVIQ: 100.67 (12.22)	16 (16)	Age Gender NVIQ	14.76 (2.08)	NVIQ: 100.56 (11.69)	Human faces (dynamic)	Verbal (forced-choice labelling)	Video clips of actors depicting each of the emotions	Angry Fearful Happy Sad Disgust Surprise Overall	- - - - F = -8.357 - F = -5.67	- - - - p = .007 - p = .023
37*	Li et al. (2017)	34 (30)	DSM-5; ADI-R	9.27 (2.23)	FSIQ: 109.94 (20.82)	39 (29)	Age Gender FSIQ	10.05 (3.20)	FSIQ: 113.03 (16.83)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Chinese facial affective picture system	Angry Fearful Happy Sad	0.47 (0.24) 0.34 (0.22) 0.81 (0.20) 0.50 (0.24)	0.66 (0.21) 0.47 (0.3) 0.9 (0.19) 0.77 (0.19)
38	Lindström et al. (2018)	15 (15)	ADI-R; DSM-IV; DSM-5; ICD-10	10.40 (-)	VIQ: 108.00 (14.72) NVIQ: 98.00 (12.89)	16 (16)	Age Gender VIQ	10.10 (-)	VIQ: 116.00 (15.30) NVIQ: 108.00 (12.90)	Speech prosody (utterance)	Nonverbal (discrimination)	Finnish word ('Saara') uttered by a female speaker with different emotional connotations	Sad	0.98 (0.04)	0.97 (0.06)
39*	Loth et al. (2018)	46 (34)	AQ	30.20 (9.40)	FSIQ: 116.00 (87-135) VIQ: 113.90 (85-160)	53 (33)	Age Gender FSIQ VIQ NVIQ	27.50 (7.80)	FSIQ: 115.50 (85-143) VIQ: 114.00 (74-146)	Human faces (static)	Verbal (matching)	Images sourced from films made in non-English speaking countries	Overall	0.74 (0.14)	0.89 (0.09)
40a*	O'Connor (2007)	18 (16)	DSM-IV; Gillberg & Gillberg (1989) criteria	26.90 (7.80)	-	18 (16)	Age Gender	25.20 (6.50)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs selected from the Mind Reading Emotions Library (Baron-Cohen et al. 2003)	Angry Happy Sad	0.94 (0.09) 0.97 (0.06) 0.87 (0.11)	0.97 (0.05) 0.99 (0.03) 0.91 (0.08)
40b	O'Connor (2007)	18 (16)	DSM-IV; Gillberg & Gillberg (1989) criteria	26.90 (7.80)	-	18 (16)	Age Gender	25.20 (6.50)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Semantically neutral sentences ("I want to go to the other movies") spoken in each of the emotions by six female and eight male actors	Angry Happy Sad	0.84 (0.15) 0.77 (0.18) 0.86 (0.13)	0.84 (0.11) 0.81 (0.15) 0.89 (0.09)
41a*	Oerlemans et al. (2014)	90 (73)	ADI-R; CRS-R; SCQ	10.60 (2.05)	FSIQ: 103.20 (13.70) VIQ: 101.55 (15.70) NVIQ: 105.20 (17.25)	139 (62)	-	9.20 (1.90)	FSIQ: 107.40 (11.70) VIQ: 108.80 (13.10) NVIQ: 106.40 (14.70)	Human faces (static)	Verbal (discrimination)	Emotional faces eliciting each of the emotions	Angry Fearful Happy Sad	82.50 (1.20) 82.60 (0.70) 93.70 (0.70) 79.20 (1.20)	85.70 (1.00) 84.80 (1.70) 95.70 (0.60) 81.60 (0.90)
41b	Oerlemans et al. (2014)	66 (54)	ADI-R; CRS-R; SCQ	11.55 (1.35)	FSIQ: 102.56 (13.30) VIQ: 1.1.20 (16.15) NVIQ: 104.25 (15.85)	72 (33)	-	10.70 (1.10)	FSIQ: 106.10 (11.30) VIQ: 108.20 (12.50) NVIQ: 104.10 (14.60)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Sentences with neutral content spoken in the tone of each emotion (Vingerhoets et al, 2003)	Angry Fearful Happy Sad	82.30 (1.80) 33.40 (2.20) 77.40 (1.90) 72.60 (2.70)	84.80 (1.70) 31.10 (2.00) 77.40 (1.80) 68.60 (2.50)
42a*	Otsuka et al. (2017)	21 (14)	AQ; DSM-IV	25.24 (5.75)	FSIQ: 112.00 (9.92) VIQ: 115.05 (11.16) NVIQ: 105.38 (12.46)	21 (14)	Age Gender FSIQ VIQ NVIQ	24.90 (6.32)	FSIQ: 113.57 (11.58) VIQ: 113.43 (12.35) NVIQ: 110.81 (12.38)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988) presented for 2000ms	Angry Fearful Happy Sad Disgust Surprise Overall	70.24 (23.21) 55.95 (31.53) 97.62 (7.52) 72.62 (30.52) 50 (36.23) 97.62 (7.52) 73.41 (12.11)	75 (19.36) 71.43 (24.09) 100 (0) 80.95 (23.59) 45.24 (28.08) 100 (0) 78.57 (8.57)

42b	Otsuka et al. (2017)	21 (14)	AQ; DSM-IV	25.24 (5.75)	FSIQ: 112.00 (9.92) VIQ: 115.05 (11.16) NVIQ: 105.38 (12.46)	21 (14)	Age Gender FSIQ VIQ NVIQ	24.90 (6.32)	FSIQ: 113.57 (11.58) VIQ: 113.43 (12.35) NVIQ: 110.81 (12.38)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988) presented for 50ms	[X] Angry [X] Fearful [X] Happy [X] Sad [X] Disgust [X] Surprise [X] Overall	32.14 (22.56) 19.05 (22.23) 89.29 (28.03) 55.95 (31.53) 10.71 (18.66) 88.1 (18.74) 49.6 (13.94)	41.67 (21.41) 19.05 (17.51) 94.05 (10.91) 64.29 (29.12) 23.81 (30.08) 94.05 (17.51) 56.15 (8.4)
43*	Pelphrey et al. (2002)	5 (5)	ADI-R; ADOS; DSM-IV; ICD-10	25.20 (-)	FSIQ: 100.75 (7.69) VIQ: 117.00 (23.12) NVIQ: 86.50 (9.57)	5 (5)	Gender	28.20 (-)	-	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Disgust Surprise Overall	0.60 (0.29) 0.65 (0.22) 0.95 (0.11) 0.85 (0.22) 0.70 (0.21) 0.80 (0.27) 0.76 (0.12)	0.90 (0.14) 0.95 (0.11) 1.00 (0.00) 0.95 (0.09) 0.80 (0.33) 0.95 (0.32) 0.93 (0.08)
44a*	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Disgust Surprise Overall	61.30 (25.99) 58.26 (26.05) 94.78 (11.63) 65.65 (24.09) 65.65 (27.44) 79.57 (22.66) 70.74 (14.79)	90.43 (7.06) 80.87 (16.49) 100 (0.00) 82.61 (14.21) 84.75 (17.29) 90.87 (9.96) 88.3 (5.47)
44b	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988)	[X] Angry [X] Fearful [X] Happy [X] Sad [X] Disgust	62.17 (34.14) 82.65 (23.12) 98.17 (4.82) 83.3 (26.08) 72.09 (29.39)	88.17 (13.51) 86.96 (26.25) 100 (0) 95.04 (13.34) 89.48 (18.85)
44c	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Nonverbal (matching)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988)	Angry Fearful Happy Sad Disgust	85.65 (17.84) 85.78 (14.970 96.91 (7.42) 77.04 (23.16) 72.7 (21.5)	95.09 (8.15) 93.17 (10.51) 96.91 (14.8) 94.43 (11.2) 86.35 (20.4)
44d	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Strings of numbers spoken in an emotional tone from Calder Vocal Emotion (Calder et al., 2004)	Angry Fearful Happy Sad Disgust	63.48 (20.14) 59.57 (25.85) 57.83 (20.66) 75.22 (15.34) 48.26 930.55)	83.48 (14.96) 77.83 (11.66) 76.96 (14.6) 80 (14.14) 76.52 (17.48)
45	Quintin et al. (2011)	26 (20)	DSM-IV	13.58 (1.92)	FSIQ: 97.00 (15.00) VIQ: 94.00 (19.00) NVIQ: 101.00 (13.00)	26 (12)	Age VIQ	13.50 (2.17)	FSIQ: 108.00 (12.00) VIQ: 107.00 (13.00) NVIQ: 107.00 (15.00)	Music (excerpt)	Verbal (forced-choice labelling)	Music clips eliciting each of the emotions	Fearful Happy Sad	4.38 (1.06) 3.77 91.21) 3.65 (1.41)	4.69 (0.55) 3.85 (0.97) 4.23 (0.82)
46*	Rhodes et al. (2018)	19 (17)	ADOS-2; DSM-IV	12.25 (1.92)	FSIQ: 107.30 (12.40) VIQ: 101.90 (11.40) NVIQ: 112.50 (15.30)	19 (14)	Age FSIQ VIQ NVIQ	12.25 (1.83)	FSIQ: 107.80 (5.00) VIQ: 103.10 (6.30) NVIQ: 110.40 (5.40)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	0.92 (0.17) 0.83 (0.15) 0.96 (0.09) 0.80 (0.24)	0.95 (0.10) 0.86 (0.15) 0.99 (0.06) 0.78 (0.22)
47a*	Rigby et al. (2018)	16 (11)	-	27.80 (7.80)	FSIQ: 106.30 (10.80) VIQ: 107.80 (13.90) NVIQ: 103.90 (15.70)	16 (11)	Age Gender FSIQ VIQ NVIQ	27.30 (7.50)	FSIQ: 113.40 (11.40) VIQ: 110.70 (9.50) NVIQ: 113.30 (11.40)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of female Caucasian actors from Pilz et al. (2006)	Angry Surprise	93.30 (7.10) 96.10 (6.60)	97.70 (3.00) 96.30 (4.70)
47b	Rigby et al. (2018)	16 (11)	-	27.80 (7.80)	FSIQ: 106.30 (10.80) VIQ: 107.80 (13.90) NVIQ: 103.90 (15.70)	16 (11)	Age Gender FSIQ VIQ NVIQ	27.30 (7.50)	FSIQ: 113.40 (11.40) VIQ: 110.70 (9.50) NVIQ: 113.30 (11.40)	Human faces (dynamic)	Verbal (forced-choice labelling)	Photographs of female Caucasian actors from Pilz et al. (2006)	[X] Angry [X] Surprise	97.50 (2.70) 95.50 (5.10)	95.50 (4.80) 98.00 (2.80)

48*	Sasson et al. (2016a)	21 (18)	ADOS	23.43 (4.36)	FSIQ: 101.48 (16.97)	39 (23)	Gender FSIQ	35.87 (9.33)	FSIQ: 100.56 (14.87)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of an individual's face from Kohler et al., 2003	^{X} Angry Fearful ^{X} Happy Sad	1.00 (0) 0.71 (0.46) 1.00 (0) 0.92 (0.28)	1.00 (0.00) 0.79 (0.41) 1.00 (0.00) 0.97 (0.26)
49*	Sasson et al. (2016b)	21 (19)	ADOS; AQ	23.81 (4.58)	FSIQ: 11.56 (12.09)	28 (25)	Age Gender FSIQ	23.75 (6.60)	FSIQ: 116.71 (10.38)	Human faces (static)	Nonverbal (discrimination)	Faces eliciting each of the emotions	Angry Happy	84.66 (13.82) 91.01 (11.58)	86.31 (10.85) 96.63 (5.31)
50*	Sawyer et al. (2012)	30 (20)	DSM	21.60 (9.80)	FSIQ: 108.10 (17.90) VIQ: 109.70 (19.10) NVIQ: 104.30 (18.20)	24 (7)	Age FSIQ VIQ NVIQ	24.00 (9.20)	FSIQ: 114.10 (13.00) VIQ: 113.40 (12.80) NVIQ: 111.40 (12.80)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Overall	69.97 (11.09)	86.77 (11.04)
51a*	Schaller & Rauh (2017)	23 (23)	ADI-R; ADOS	15.72 (1.25)	NVIQ: 105.65 (11.47)	22 (22)	Age Gender NVIQ	15.85 (0.97)	NVIQ: 103.77 (11.09)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	Angry Fearful Overall	F = -6.48 F = -6.11 0.81 (0.07)	p = .015 p = .018 0.86 (0.05)
51b	Schaller & Rauh (2017)	23 (23)	ADI-R; ADOS	15.72 (1.25)	NVIQ: 105.65 (11.47)	22 (22)	Age Gender NVIQ	15.85 (0.97)	NVIQ: 103.77 (11.09)	Human faces (dynamic)	Verbal (forced-choice labelling)	Facially Expressed Emotion Labelling (FEEL; Kessler et al., 2002)	^{X} Overall	0.83 (0.09)	0.79 (0.11)
52*	Schelinski & von Kriegstein (2019)	16 (13)	ADI-R; ADOS; AQ; ICD-10	33.75 (10.12)	FSIQ: 110.31 (13.79) VIQ: 110.75 (12.35) NVIQ: 107.38 (17.55)	16 (13)	Age Gender FSIQ VIQ NVIQ	33.69 (9.58)	FSIQ: 111.50 (10.97) VIQ: 108.75 (12.59) NVIQ: 112.69 (9.59)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Two-syllabic semantically neutral German nouns spoken in each one of the emotions by one female and one male actor	Angry Fearful Happy Sad Disgust	84.98 (13.15) 68.37 927.82) 64.10 (30.93) 51.45 (32.05) 50.18 (22.77)	92.48 (6.58) 89.85 (6.71) 80.54 (9.93) 81.64 (10.76) 65.97 (21.78)
53*	Shafritz et al. (2015)	15 (12)	ADI; ADOS	18.10 (-)	FSIQ: 101.50 (18.60) VIQ: 105.70 (18.80) NVIQ: 103.50 (17.40)	15 (12)	Age Gender VIQ NVIQ	18.40 (12-23)	FSIQ: 115.20 (9.30) VIQ: 118.80 (14.90) NVIQ: 108.00 (8.10)	Human faces (static)	Verbal (detection)	Ekman facial affect set	Fearful Happy	92.40 (7.68) 95.00 (6.45)	93.21 (8.05) 96.14 (3.42)
54a*	Shanok et al. (2019)	12 (9)	GARS-2	5.75 (0.97)	-	16 (10)	-	5.50 (1.41)	-	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	77.77 (16.42) 86.11 (17.17) 94.44 (12.98) 69.44 (22.49)	91.67 (14.99) 93.75 (13.44) 100.00 (0.00) 95.83 (11.39)
54b	Shanok et al. (2019)	12 (9)	GARS-2	5.75 (0.97)	-	16 (10)	-	5.50 (1.41)	-	Human faces (static)	Verbal (forced-choice labelling)	Mothers replicating photographs from the NimStim Face Stimulus Set	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad	97.22 (14.91) 83.33 (22.48) 100.00 (0.00) 86.11 (17.17)	91.67 (14.99) 93.75 (13.44) 100.00 (0.00) 93.75 (13.44)
55a*	Sinzig et al. (2008)	19 (17)	ASD: ADI-R; ADOS DSM-IV	13.60 (3.40)	FSIQ: 111.00 (19.10)	29 (22)	Age FSIQ	12.80 (2.90)	FSIQ: 109.00 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	Angry Fearful Happy Sad Disgust Surprise Overall	5.40 (1.60) 3.20 (1.30) 7.60 (1.60) 5.80 (1.70) 4.40 (1.70) 4.90 (1.30) 0.76 (0.09)	6.1 (1.10) 2.9 (1.30) 7.7 (1.10) 5.9 (1.60) 4.1 (1.50) 5.2 (0.90) 0.78 (0.07)
55b	Sinzig et al. (2008)	21 (20)	ASD+ADHD: ADI-R; ADOS; DCL-TES; DSM-IV	11.60 (3.70)	FSIQ: 102.00 (13.10)	29 (22)	Age FSIQ	12.80 (2.90)	FSIQ: 109.00 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad ^{X} Disgust ^{X} Surprise ^{X} Overall	5.10 (2.00) 3.50 (2.80) 6.70 (2.20) 4.90 (2.30) 4.00 (1.80) 3.60 (1.80) 0.72 (0.10)	6.1 (1.10) 2.9 (1.30) 7.7 (1.10) 5.9 (1.60) 4.1 (1.50) 5.2 (0.90) 0.78 (0.07)

56*	Stephenson et al. (2019)	30 (-)	ADOS-2; AQ; DSM-5	24.52 (6.04)	FSIQ: 112.36 (10.63)	46 (-)	FSIQ	20.93 (2.03)	111.95 (8.21)	Human faces (dynamic)	Verbal (forced-choice labelling)	Dynamic faces from the Amsterdam Dynamic Facial Expression Set; van der Schalk et al., 2011)	Angry Fearful Happy	0.93 (0.26) 0.99 (0.09) 1.00 (0.05)	0.95 (0.22) 0.99 (0.10) 1.00 (0.05)
57*	Stewart et al. (2013)	11 (7)	DSM-IV	27.20 (7.50)	VIQ: 14.90 (6.20)	14 (8)	Age Gender VIQ	26.40 (5.60)	VIQ: 18.10 (4.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from JAFFE database (Lyons et al 1999)	Overall	0.55 (0.08)	0.68 (0.07)
58a*	Tanaka et al. (2012)	66 (56)	ADI-R; ADOS-G; DSM-IV	11.90 (4.00)	FSIQ: 106.80 (20.90)	68 (43)	Age FSIQ	11.90 (3.10)	FSIQ: 106.80 (7.80)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad Disgust Surprise Overall	82.40 (21.20) 67.50 (22.00) 98.10 (6.70) 91.70 (13.50) 87.50 (18.00) 86.60 (18.60) F = -2.86	89.30 (13.40) 69.20 (20.80) 97.50 (5.80) 94.70 (9.70) 92.80 (12.20) 87.80 (14.60) p = .09
58b	Tanaka et al. (2012)	67 (57)	ADI-R; ADOS-G; DSM-IV	12.00 (4.00)	FSIQ: 106.80 (20.90)	66 (42)	Age FSIQ	11.90 (3.10)	FSIQ: 106.80 (7.80)	Human faces (static)	Nonverbal (matching)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad Disgust	84.30 (15.30) 74.10 (22.20) 96.80 (9.70) 65.90 (22.40) 66.70 (20.10)	94.20 (9.90) 86.90 (16.90) 99.50 (2.90) 82.30 (16.80) 75.80 (16.60)
59a*	Taylor et al. (2015)	17 (12)	ADOS-G; DSM-IV	9.67 (2.25)	VIQ: 91.72 (15.33) NVIQ: 10.41 (2.94)	54 (26)	Age NVIQ	8.94 (1.92)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs posed by both children and adults from movie files featured on Mind Reading DVD (Baron-Cohen, 2002)	Overall	62.50 (11.97)	70.75 (10.99)
59b	Taylor et al. (2015)	17 (12)	ADOS-G; DSM-IV	9.67 (2.25)	VIQ: 91.72 (15.33) NVIQ: 10.41 (2.94)	54 (26)	Age NVIQ	8.94 (1.92)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	The sentence "Oh I'm going out of the room now but I'll be back later" spoken each of the emotions with Australian accents	Overall	64.22 (13.26)	71.06 (12.71)
60*	Tell et al. (2014)	22 (17)	ADI-R; ADOS-G; DSM-IV	10.31 (-)	FSIQ: 102.68 (6.64) VIQ: 9.32 (1.69)	22 (17)	Age Gender	9.80 (-)	VIQ: 10.64 (1.67)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	63.30 (4.90) 49.30 (7.10) 80.20 (4.80) 45.00 (5.90)	65.80 (4.70) 73.20 (6.40) 79.60 (4.70) 54.50 (6.90)
61*	Tottenham et al. (2014)	33 (30)	ADOS; AQ	15.00 (6.00)	FSIQ: 111.00 (-)	53 (35)	FSIQ VIQ	16.00 (8.00)	FSIQ: 103.00 (-)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry	0.93 (0.15)	0.90 (0.24)
62*	Uono et al. (2011)	28 (23)	DSM-IV-TR	17.60 (5.20)	FSIQ: 103.30 (13.40) VIQ: 105.20 (14.70) NVIQ: 100.10 (13.30)	28 (24)	Age Gender	18.00 (4.00)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry Fearful Happy Sad Disgust Surprise Overall	60.70 (21.70) 32.10 (26.46) 98.70 (4.23) 79.90 (16.93) 39.30 (24.34) 95.30 (10.05) 67.70 (8.47)	53.80 (22.75) 51.80 (19.58) 98.20 (4.76) 76.30 (19.58) 46.40 (22.75) 94.20 (9.00) 70.10 (7.94)
63a*	Uono et al. (2013)	18 (15)	AS: CARS; DSM-IV	18.60 (6.50)	FSIQ: 106.00 (11.90) VIQ 108.30 (11.90) NVIQ: 101.90 (13.90)	18 (14)	Age Gender	18.80 (3.60)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry Fearful Happy Sad Disgust Surprise Overall	59.00 (24.18) 40.30 (30.12 97.90 (4.67) 84.70 (14.42) 48.60 (25.46) 97.20 (5.52) 71.30 (8.06)	54.2 (22.06) 54.9 (16.12) 97.9 (4.67) 76.4 (18.24) 52.1 (22.06) 95.8 97.64) 71.9 (6.79)

63b	Uono et al. (2013)	18 (12)	PDD-NOS: CARS; DSM-IV	19.80 (4.70)	FSIQ: 101.00 (14.10) VIQ: 103.10 (17.70) NVIQ: 98.20 (11.30)	18 (14)	Age Gender	18.80 (3.60)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	[X] Angry [X] Fearful [X] Happy [X] Sad [X] Disgust [X] Surprise [X] Overall	62.50 (16.97) 31.30 (22.06) 99.30 (2.97) 75.70 (16.12) 32.60 (22.91) 93.80 (11.46) 65.90 (7.64)	54.2 (22.06) 54.9 (16.12) 97.9 (4.67) 76.4 (18.24) 52.1 (22.06) 95.8 97.64) 71.9 (6.79)
64*	Vannetzel et al. (2011)	10 (9)	ADI-R; CARS; DSM-IV	9.60 (1.70)	FSIQ: above 70	35 (30)	Age	8.40 (1.80)	-	Human faces (static)	Nonverbal (discrimination)	Ekman facial affect set	Angry Happy Sad	33.30 (20.54) 70.00 (15.48) 53.30 (19.28)	80.00 (26.55) 85.70 (23.01) 84.80 (20.06)
65a*	Waddington et al. (2018)	89 (69)	ADI-R	12.32 (2.48)	FSIQ: 101.51 (14.67)	220 (110)	Gender	13.11 (2.35)	FSIQ: 105.5 (12.42)	Human faces (static)	Verbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychologica l Task battery (DeSonneville, 1999)	Angry Fearful Happy Sad	85.24 (10.77) 84.52 (12.67) 94.58 (6.12) 79.98 (12.98)	88.24 (9.48) 89.85 (10.20) 96.00 (4.58) 84.40 (10.19)
65b	Waddington et al. (2018)	89 (69)	ADI-R	12.32 (2.48)	FSIQ: 101.51 (14.67)	220 (110)	Gender	13.11 (2.35)	FSIQ: 105.5 (12.42)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Affective Prosodytask from the Amsterdam Neuropsychologica l Task battery (DeSonneville, 1999)	Angry Fearful Happy Sad	84.93 (16.76) 36.47 (21.84) 79.00 (17.02) 72.73 (22.68)	87.20 (14.61) 40.36 (20.73) 81.06 (15.69) 72.77 (21.24)
66a*	Wallace et al. (2008)	26 (23)	ADI-R; ICD-10	32.00 (9.00)	VIQ: 148.00 (13.00) NVIQ: 101.00 (18.00)	26 (23)	Age Gender FSIQ VIQ NVIQ	31.00 (9.00)	VIQ: 153.00 (9.00) NVIQ: 98.00 (12.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry Fearful Happy Sad Disgust Surprise	0.62 (0.17) 0.62 (0.24) 0.93 (0.13) 0.68 (0.22) 0.72 (0.23) 0.68 (0.22)	0.73 (0.22) 0.85 (0.17) 0.99 (0.70) 0.86 (0.11) 0.88 (0.12) 0.84 (0.22)
66b	Wallace et al. (2008)	26 (23)	ADI-R; ICD-10	32.00 (9.00)	FSIQ: 122.00 (10.00) VIQ: 118.00 (14.00) NVIQ: 122.00 (7.00)	26 (23)	Age Gender FSIQ VIQ NVIQ	31.00 (9.00)	FSIQ: 117.00 (13.00) VIQ: 115.00 (14.00) NVIQ: 116.00 (13.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion presented in a piecemeal fashion starting from the eyes or the mouth	[X] Angry [X] Fearful [X] Disgust [X] Surprise	0.70 (0.22) 0.76 (0.22) 0.72 (0.30) 0.65 (0.27)	0.82 (0.27) 0.92 (0.11) 0.91 (0.13) 0.75 (0.25)
67a*	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Neutral sentences spoken in each one of the emotions	Angry Happy Sad	83.72 (14.07) 88.01 (14.54) 96.29 (10.13)	85.72 (14.45) 95.55 (6.25) 99.15 (2.35)
67b	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Emotional sentences spoken in matching emotional prosody	[X] Angry [X] Happy [X] Sad	84.95 (15.59) 74.47 (16.42) 97.71 (4.79)	86.94 (11.11) 81.33 (19.24) 98.57 (2.57)
67c	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Neutral words spoken in each one of the emotions	[X] Angry [X] Happy [X] Sad	84.90 (17.83) 53.78 (30.05) 94.23 (12.46)	79.12 (19.6) 63.12 (27.18) 94.23 (3.68)
67d	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Emotional words spoken in matching emotional prosody	[X] Angry [X] Happy [X] Sad	73.00 (17.54) 73.01 (15.44) 94.23 (7.38)	69.45 (18.62) 83.00 (10.5) 97.11 (3.35)

68*	Wang & Adolphs (2017)	18 (18)	ADI-R; ADOS-2; DSM-5; ICD-10	30.80 (7.40)	FSIQ: 105.00 (13.30)	15 (11)	Age FSIQ VIQ NVIQ	35.10 (11.40)	FSIQ: 107.00 (8.69)	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs obtained from the STOIC database (Roy et al., 2007)	Fearful Happy	0.97 (0.03) 0.96 (0.05)	0.99 (0.03) 0.98 (0.04)
69*	Wingenbach et al. (2017)	12 (9)	AQ; SCQ	17.30 (0.75)	-	12 (9)	Age Gender	16.90 (0.29)	-	Human faces (dynamic)	Verbal (forced-choice labelling)	Facial emotional video from The Amsterdam Dynamic Facial Expression Set (van der Schalk et al., 2011)	Angry Fearful Happy Sad Disgust Surprise	0.74 (0.08) 0.47 (0.08) 0.96 (0.02) 0.78 (0.04) 0.77 (0.07) 0.94 (0.03)	0.89 (0.08) 0.78 (0.05) 0.95 (0.05) 0.88 (0.02) 0.84 (0.05) 0.93 (0.03)
70*	Wong et al. (2012)	19 (16)	ADI-R	11.28 (1.48)	FSIQ: 118.21 (14.93)	21 (15)	Age FSIQ	10.24 (1.81)	FSIQ: 113.43 (11.21)	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli from the Standardized Penn Emotion Recognition Set (Gur et al., 2002)	Angry Fearful [X] Happy Sad Disgust	2.89 (0.66) 3.05 (0.91) 4.00 (0) 3.16 (0.96) 1.16 (0.96)	2.71 (0.96) 2.9 (1.14) 4.00 (0) 3.19 (0.98) 1.52 (1.25)
71*	Wright et al. (2008)	35 (33)	ADI-R; ADOS; AQ; ICD-10	11.31 (2.17)	FSIQ: 104.63 (17.99) VIQ: 105.66 (21.01) NVIQ: 103.03 (16.09)	35 (33)	Age Gender FSIQ VIQ NVIQ	11.57 (1.94)	FSIQ: 103.86 (16.26) VIQ: 105.74 (16.31) NVIQ: 100.94 (16.39)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Ekman's Facial Expressions of Emotion: Stimuli and Tests (FEEST)	Angry Fearful Happy Sad Disgust Surprise Overall	5.57 (2.05) 4.77 (2.38) 9.51 (1.17) 6.91 (2.42) 5.40 (2.98) 8.09 (2.17) 40.26 (9.56)	6.57 (1.80) 4.74 (2.97) 9.91 (0.28) 6.43 (2.71) 5.03 (2.74) 8.77 (2.13) 41.51 (7.99)
72*	Yeung et al. (2014)	18 (15)	ADI-R; DSM-IV	9.61 (3.13)	FSIQ: 101.33 (10.85)	18 (11)	Age Gender FSIQ	10.72 (3.61)	FSIQ: 107.06 (9.35)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Angry Fearful Happy Sad Disgust Surprise Overall	0.29 (0.16) 0.12 (0.11) 0.95 (0.28) 0.44 (0.19) 0.10 (0.11) 0.53 (0.21) 0.41 (0.14)	0.40 (0.17) 0.19 (0.14) 1.12 (0.17) 0.57 (0.17) 0.23 (0.19) 0.67 (0.14) 0.53 (0.12)

Abbreviations: ADI-R, Autism Diagnostic Instrument Revised; ADOS, Autism Diagnostic Observation Schedule; ADI-R-III, AADOS-G, Autism Diagnostic Observation Schedule—Generic; ADOS-2, Autism Diagnostic Observation Schedule, Second Edition; ASDI, Autism Spectrum Diagnostic Interview; ASDS, Asperger Syndrome Diagnostic Scale; ASQ-J, Autism Screening Questionnaire, Japanese Version; AQ, Autism Spectrum Quotient Questionnaire; AQ-J, Autism Spectrum Quotient, Japanese Version; ASSQ, Autism Spectrum Screening Questionnaire; CARS, Childhood Autism Rating Scale; DCL-TES, Diagnostic Checklist for Pervasive Developmental Disorders (*Diagnostik Chekliste fr Tiefgreifende Entwicklungsstörungen*); CRS-R, Conners' Rating Scales-Revised; DSM, The Diagnostic and Statistical Manual of Mental Disorders; DSM-IV, The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; DSM-IV-TR, The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision; DSM-5, The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; GARS-2, Gilliam Autism Rating Scale, Second Edition; ICD-10, International Statistical Classification of Diseases, Tenth Revision; SCQ, Social Communication Questionnaire; SRS, Social Responsiveness Scale.